Journal of Official Statistics
Vol. 1, No. 1, 1985, pp. 83–86

Letters to the Editor

Letters to the Editor will be confined to discussions of papers which have appeared in the Journal of Official Statistics and of important issues facing the statistical community.

All Those Journals

Dear Editor,

I am grateful for this opportunity to publicize my proposal that I first expressed in May '84 in a letter addressed to friends at Statistics Sweden, Statistics Canada and the IASS. The proposal is incomplete, imperfect, premature and unconventional. Perhaps you and others can shape it into a practical solution for a growing problem.

The *Journal of Official Statistics* in 1985 replaced the Statistical Review as "the official methodological journal of Statistics Sweden. It will deal exclusively with problems faced by government statisticians all over the world". It will accept contributions chiefly from Statistics Sweden but also from elsewhere (in English).

Survey Methodology has been for several years a publication (in English and French) of and chiefly for the technical staff of Statistics Canada. However it has been gaining attention, readers and some contributors in government and survey institutions. It has been very good, but informal, not typeset, and with the policy that republication in other journals is permitted.

The two journals differ somewhat and each has a primary duty to its own staff, problems and office. But both have much to offer to all governmental and survey statisticians. Both want to avoid (we hope) becoming still another academic, English, mathematical-statistical journal.

We see two journals devoted to government statistical and survey methods. How soon will Australia, the UK, the USA, and others also issue their journals? Also in English? Who will receive and who will read them all? Meanwhile we have requests for the IASS to have such a publication that we cannot meet. I propose some kind of loose cooperation between the two journals and the IASS, with the following features. 1) The journals would maintain autonomy and differences, with primary attention to their primary home clients. 2) Any changes would be gradual and voluntary. 3) Each journal would send some copies (20 or 50 or 200) to the other office for distribution. 4) The two journals would be staggered over the year, each with 2 or 3 issues. 5) Other offices may join into this cooperation. The chief role of IASS will be to promote this cooperation. 6) Some exchanges of ideas, dialogue and "replies" may develop freely between the two journals. 7) IASS would offer a label like, "Endorsed by the IASS". It would also advertise for subscribers. Also perhaps buy subscriptions for each IASS Country Representative and perhaps for each director of national statistical office. 8) IASS would also stimulate contributions from survey and governmental statisticians. But acceptance of them, and general responsibilities, would remain with the separate editorial offices. 9) If the new International Association of Official Statisticians is formed in '85, it may wish to cosponsor together with IASS.

This unprecendented plan certainly has problems and faces obstacles. But it deserves discussions, suggestions and perhaps a modified implementation, because the present situation is chaotic and needs directed developments. The plan offers opportunities without great risks since autonomy is retained by the respective editorial offices; and withdrawal to the *status quo* will be easy if the plan fails. Meanwhile the IASS can help to coordinate and foster its seemingly impossible task: an international, multilingual organization of survey

statisticians. And perhaps a future IAOS will join IASS in these efforts.

Sincerely and hopefully,
Leslie Kish
President IASS ('83–'85)
University of Michigan
Ann Arbor, MI. 48106
USA

Reply

Dear Professor Kish,

You and I have exchanged letters concerning the new journal for some time. Your suggestions have been very stimulating and interesting but this time I will abstain from detailed comments and instead give the prospective readers a chance to discuss the matter.

I will, however, make some general comments.

1) Our endeavors certainly mean the birth of a new journal, but at the same time another journal ceases to appear, so we have not contributed to an increase in the number of journals.

2) It is not true that the Journal of Official Statistics will accept contributions chiefly from Statistics Sweden. Quite simply, there is not enough people at Statistics Sweden who have the kind of work or the inclination to submit papers to scholarly journals. We would be very happy if we would receive one or two contributions from Statistics Sweden to each issue.

3) The basic idea with the redesign of our methodological journal has been to make it an international one, which is also indicated by the subtitle: "An International Review published by Statistics Sweden". Thus, we hope that professionals in Australia, the UK, the USA and other countries will choose to submit contributions to our journal. In that way the statistical offices in these countries do not have to issue their own journals. There is already a suitable refereed journal for them, albeit published in Sweden by its statistical office, Statistics Sweden. However, we are of course interested in any cooperation and help that we can get from IASS and IAOS.

Yours sincerely,

Lars Lyberg
Editor

A Note on Entropy

Dear Sir,

Since many years it is customary to use the term "Entropy"[1] in such disciplines as economics, sociology, politics, regional sciences, mathematical statistics, which have nothing to do with the theory of thermodynamics. It is only a question of terminology, but nevertheless of some importance for a clear scientific language. A correction might be advisable.

In classical thermodynamics the entropy S of a system is defined by $dS = dQ/T$, where Q is the heat supplied in a reversible way to a system, and T stands for absolute temperature. Only differences in S can be measured. The dimension of entropy is [energy/temperature].

A second notion of entropy is met in the kinetic theory of gases. In a simple form and apart from constants the entropy of a system is defined by $S = k \log W = -\int_0^\infty f(v) \log f(v) dv = kH$, where W is the probability of the state of the system and $f(v)$ is the frequency distribution of the velocities of the molecules. Moreover k is Boltzmann's constant with the dimension [energy/temperature]. It should be emphasized that the second definition of S is based on the function $H = -\sum_{i=1}^{N} p_i \log p_i$ $(i = 1, \ldots, N)$, where p_i are probabilities or fractions, which sum up to 1.

The function H is a measure of the probability of the state of a system (e.g. a frequency distribution), or a measure for information, or a measure of inequality like the one introduced by Theil. The function H as such is not a thermodynamical notion, because it can be generated by pure statistical reasoning. Given M items and N cells, let there be m_i items in cell i with the probability $p_i = m_i/M$. A simple calculation shows that the probability W of this state is given by $\log W = -\sum_{i=1}^{N} p_i \log p_i = H$ (apart from constants). H is dimensionless and shows analogy with entropy. Obviously S and H are not the same, but H is often called entropy. Some quotations may be presented here.

[1] From the Greek: en = in and tropein = to turn. Perhaps the word may be derived with reference to the Second Law of Thermodynamics. (The word should not be related to the Greek "entropiai", which means cunning tricks.)

*

Hart, P. E. (1971): Entropy and Other Measures of Concentration. Journal of the Royal Statistical Society, Series A, 134, p. 73.

ENGINEERS have made important contributions to economics and to economic statistics, though a long time lag has often tended to elapse before economists have recognized their value. This is certainly true for the study of business concentration, for the work of the French engineer Gibrat (1931) was neglected by economists for many years. In recent years, economists have begun to use the concept of entropy, originally formulated in statistical thermodynamics and subsequently used in information theory by communications engineers. In particular the entropy function has been used to measure the extent to which the output or employment of an industry is concentrated in the control of a few large firms in an industry.

Karlin, S. and Rinott, Y. (1981): Entropy Inequalities for Classes of Probability Distributions I. The Univariate Case. Advances in Applied Probability, 13, p. 93.

Entropy functionals of probability densities feature importantly in classifying certain finite-state stationary stochastic processes, in discriminating among competing hypotheses, in characterizing Gaussian, Poisson, and other densities, in describing information processes, and in other contexts.

The entropy functional for a probability density

$$H(p) = -\int_0^\infty p(x)\log p(x)\, dx$$

or in discrete form

$$H(p) = -\Sigma\, p_i \log p_i$$

plays a prominent role in many branches of science. Indeed, $H(p)$ is used as a measure of uncertainty in statistical theory and information processing, as a measure of disorder for thermodynamic systems (entropy equals the Boltzmann H-function), as a measure of diversity in ecological structures, and as part of a set of criteria for apportioning races and species in population classifications.

Lev, B. and Theil, H. (1978): A Maximum Entropy Approach to the Choice of Asset Depreciation. Journal of Accounting Research, 16, p. 286.

The entropy of a distribution with density function $f(t)$ is:

$$H = -\int_0^\infty f(t)\log f(t)\, dt.$$

The choice of a distribution most consistent with available information thus proceeds by maximizing the entropy (degree of uninformativeness) subject to the available information.

Possible extensions of the use of the maximum entropy criterion include the derivation of additional depreciation schemes based on different kinds of information available.

Nijkamp, P. and Paelinck, J. H. P. (1974): A Dual Interpretation and Generalization of Entropy Maximization in Regional Sciences. Papers of the Regional Science Association, 23, p. 13.

The major part of modern science is concerned with systems, in the sense of entities consisting of specialized, interdependent parts; see Berry (5). Therefore, it is understandable that the concept of entropy is frequently being used in order to obtain insights into the uncertainty of systems. In an analogous manner, the notion of entropy is introduced into the social sciences; hence the term "social physics'. Recently the concept of entropy has also found applications in regional systems (for example, trip distributions, freight flows, migration flows, and activity allocations). These rather complex spatial processes are overloaded with uncertainties. The entropy concept attempts to determine the most probable spatial configurations of a system which is capable of adopting numerous uncertain spatial states. This implies that entropy in regional analysis is a probability concept, describing the outcome of a stochastic process.

Theil, H. (1967): Economics and Information Theory. North Holland, Amsterdam.
Theil, H. (1972): Statistical Decomposition Analysis. North Holland, Amsterdam.

The expected information of a distribution is frequently called the *entropy* of that distribution. As stated in Section 1.8, this term is of physical origin. It is a measure of "disorder'. The closer the n probabilities x, are to $1/n$, and the larger n is, the less order there is in the system. The second law of thermodynamics states that there is an inherent tendency of the entropy to increase. There is no need to dwell on physical analogies here; we shall merely use 'entropy' as a synonym for the longer expression 'expected information'.

Wilson, A. G. (1970): The Use of the Concept of Entropy in System Modelling. Operational Research Quarterly, 21, p. 247.

The concept of 'entropy' is widely used in the physical sciences and, recently, has aroused interest in the social sciences, especially in building models of urban and regional systems to be used for planning purposes.

There are considerable difficulties in writing about entropy in an operational research context at the present time unless the context is made absolutely plain. The problem is that the concept of entropy is often misused and this leads many people to reject it as a useful concept. The purpose of this paper is to demonstrate types of circumstances where it is very fruitful indeed.

*

As for suggestions of a more appropriate denomination there are various possibilities. If the word 'entropy' (S) is reserved for thermodynamics a more specific indication is necessary when using 'entropy' (H) in other disciplines, e.g. 'information entropy', 'statistical entropy', 'numerical entropy', 'inequality coefficient'.

However, the word 'entropy' (*sec*) has become so rooted in all kinds of disciplines in addition to thermodynamics that a change may not be accepted by scientists.

Recently F.R. Povel (in a thesis for the Doctorate, Utrecht 1984) made an appropriate suggestion: In thermodynamics 'entropy' (S) should be termed 'thermodynamical entropy' while in all other disciplines simply used 'entropy' (H).

The above asks for some adaptation in the thermodynamical terminology which hopefully may be acceptable.

J. H. C. Lisman and R. de Man
Blauwe Kamerlaan 19
2594 BK The Hague
Netherlands