CBPF-NF-017/93

Generalized Entropy-Based Criterion for Consistent Nonparametric Testing

by

Constantino Tsallis

Centro Brasileiro de Pesquisas Físicas — CBPF/CNPq
Rua Dr. Xavier Sigaud, 150
22290-180 – Rio de Janeiro – RJ, Brasil

Abstract

Through the use of a recently generalized entropy, we propose, along the lines of Robinson 1991, a consistent criterion for testing relevant econometric hypothesis such as independence of time series quantities. This criterion recovers that of Kullback-Leibler as a particular case, and yields a satisfactory alternative to the criterion proposed by Brock et al 1987.

Key-words: Entropy; Nonparametric testing; Independence criterion; Time series.

The problem of consistent testing (i.e., discrimination between two hypothesis) is a central one in Econometrics, as well as in many other areas of knowledge. Nonparametric testing is of course a very neat one, and has been proposed, on an entropy basis, by Dmitriev and Tarasenko 1973, Ahmad and Lin 1976 and Vasicek 1976, among others. Very recently, Robinson 1991 has used the Kullback and Leibler 1951 measure of information (in turn based on Shannon 1948 entropy) to make an elegant discussion of independence versus dependence in time series of quantities of financial interest. He applied this procedure to analyse the daily, weekly and monthly exchange rates for the Deutschmark, Japanese yen, Swiss franc and Pound sterling against the US dollar, using data of the Bank of England covering the period 2 January 1978 through 28 June 1985.

On a quite different background, we have proposed (Tsallis 1988) a generalization of Shannon entropy. This generalized entropy (hereafter noted S_q, where q is a fixed arbitrary real number; $q \to 1$ yields Shannon entropy) has enabled various satisfactory generalizations such as that of the Boltzmann-Gibbs equilibrium distribution (Tsallis 1988), the Thermodynamics (Curado and Tsallis 1991), the Boltzmann H-theorem (Mariz 1992, Ramshaw 1993 (a,b)), the Ehrenfest theorem (Plastino and Plastino 1993 a), the von Neumann equation (Plastino and Plastino 1993 b), the Langevin and Fokker-Planck equations (Stariolo 1993), the Schroedinger equation (Silver and Tsallis 1993), the Variational Method in Statistical Mechanics (Plastino and Tsallis 1993), among others. Furthermore, the generalized entropy S_q has enabled (Plastino and Plastino 1993c) a satisfactory solution of an old paradox in Astrophysics, namely that related to Chandrasekhar 1958 polytropic theory for stellar dynamics (see also Binney and Tremaine 1987).

The aim of the present work is to show how these ideas can be used to generalize the Robinson 1991 proposal for consistent nonparametric testing.

Let us first adapt the generalized entropy to a (normalized) probability distribution function $f(x)$ where x is a continuous random variable defined in a d-dimensional Euclidean space. We define (Tsallis 1988).

$$S_q(f) \equiv \int dx\, f(x) \frac{1-[f(x)]^{q-1}}{q-1} \tag{1}$$

with $\int dx\, f(x) = 1$ and $q \in \Re$. Several interesting properties of $S_q(f)$ are presented in Tsallis 1988 and Curado and Tsallis 1991. Let us retain now that, by using the fact that $[f(x)]^{q-1} \sim 1 + (q-1)\ln f(x)$ in the $q \to 1$ limit, we promptly obtain Shannon entropy $S^S(f)$ as a particular case. In other words,

$$\lim_{q\to 1} S_q(f) = -\int dx\ f(x) \ln f(x) \equiv S^S(f) \tag{2}$$

Let us next establish an interesting property concerning *independent* random variables respectively described by $f_1(x_1)$ and $f_2(x_2)$ (assumed normalized). Definition (1) straightforwardly leads to

$$S_q(f_1 f_2) = S_q(f_1) + S_q(f_2) + (1-q) S_q(f_1) S_q(f_2) \tag{3}$$

i.e., S_q is a *pseudo-additive* (or *pseudo-extensive*) quantity; it is *additive* (or *extensive*) if and only if $q = 1$. This property gives a hint on the statistical interpretation of the entropic parameter q : $(q-1)$ measures the *lack of extensivity* of the information on the system. Also, it is worthy mentionning at this point that Jumarie 1988 (a,b) suggests, for a quantity analogous to q, that $q = 1$, $q < 1$ and $q > 1$ respectively correspond to "no knowledge", "prior knowledge" and "prior misknowledge" (see also Losee 1988).

Let us now tackle with Kullback and Leibler 1951 measure of information $I(f,g)$ (see Eq. 1.1 of Robinson 1991) for discriminating between two hypothesis respectively characterized by the generic normalized distributions $f(x)$ and $g(x)$:

$$I(f,g) \equiv \int dx\, f(x) \ln \frac{f(x)}{g(x)} \tag{4}$$

This quantity satisfies (see, for instance, Robinson 1991)

$$\begin{aligned} I(f,g) &\geq 0\ \forall (f,g) && (5.a)\\ I(f,g) &= 0 \text{ if and only if } f = g \text{ almost everywhere} && (5.b) \end{aligned}$$

This property is strongly emphasized by Robinson 1991 (see his Eq. (1.2)) since it constitutes the very basis for *consistency* of his nonparametric testing. Definition (1) naturally leads (see also Ramshaw 1993a and Silver and Tsallis 1993) to the following generalization of Kullback and Leibler measure of information (or *cross entropy*):

$$I_q(f,g) \equiv \int dx\, f(x) \frac{[f(x)/g(x)]^{q-1} - 1}{q-1} \tag{6}$$

We can straightforwardly verify that $\lim_{q \to 1} I_q(f,g) = I(f,g),\ \forall (f,g)$.

Let us now generalize (by following along the lines of Plastino and Tsallis 1993) the very important property (5). If $r \geq 0$, we have that

$$\begin{aligned} \frac{r^{q-1}-1}{q-1} &\geq 1 - \frac{1}{r} \quad \text{if} \quad q > 0 && (7.a)\\ &= 1 - \frac{1}{r} \quad \text{if} \quad q = 0 && (7.b)\\ &\leq 1 - \frac{1}{r} \quad \text{if} \quad q < 0 && (7.c) \end{aligned}$$

(for $q \neq 0$, the equality holds if and only if $r = 1$). Consequently, for say $q > 0$,

$$\frac{[\frac{f(x)}{g(x)}]^{q-1} - 1}{q-1} \geq 1 - \frac{g(x)}{f(x)} \tag{8}$$

hence

$$\int dx\, f(x) \frac{[f(x)/g(x)]^{q-1} - 1}{q-1} \geq \int dx\, f(x) \left[1 - \frac{g(x)}{f(x)}\right] = 1 - 1 = 0 \tag{9}$$

But the left-side member of this inequality precisely is $I_q(f,g)$. Consistenly Eqs. (7) yield

$$I_q(f,g) \geq 0 \quad \text{if} \quad q > 0 \tag{10.a}$$
$$= 0 \quad \text{if} \quad q = 0 \tag{10.b}$$
$$\leq 0 \quad \text{if} \quad q < 0 \tag{10.c}$$

The equality holds, for $q \neq 0$, if and only if $f = g$ almost everywhere. Equations (5) are thus generalized for arbitrary q. Let us mention that it is also possible to establish Eqs. (10) by starting from inequalities proved in Plastino and Tsallis 1993 and performing the transformation $q - 1 \leftrightarrow 1 - q$.

By performing a different transformation, namely $q - 1/2 \leq 1/2 - q$, into Eqs. (10), we can prove that

$$\frac{I_{1-q}(f,g)}{1-q} = \frac{I_q(g,f)}{q} \tag{11}$$

Consequently, as a family of nonparametric entropy-based testings, it is enough to consider $q \geq 1/2$ with

$$I_q(f,g) \geq 0 \tag{12}$$

the equality holding if and only if $f = g$ almost everywhere. The particular case $q = 1$ has been extensively discussed by Robinson 1991. The criterion for the particular case $q = 1/2$ becomes

$$\int dx \sqrt{f(x)g(x)} \leq 1 \tag{13}$$

The criterion for the particular case $q = 2$ becomes

$$\int dx [f(x)]^2 / g(x) \leq 1 \tag{14}$$

Let us now adapt the main results of this paper to the problem of independence of a time series $\{X_t\}$ $(t = 0, 1, 2, \cdots) : X_t$ could be the daily (or weekly or monthly) enchange rate of some currency, or alternatively, following Robinson 1991, it could be given by $X_t \equiv \ln(Y_t/Y_{t-1})$ where Y_t would be the just mentionned exchange rate. The relevant random variable would be $z \equiv (x,y) \equiv (X_t, X_{t-1})$, (i.e., a $d = 2$ problem), and the corresponding distribution function would be $f(x,y)$ such that $\int dxdy\, f(x,y) = 1$. The marginal distribution functions are given by

$$h_1(x) \equiv \int dy\, f(x,y) \tag{15.a}$$
$$h_2(y) \equiv \int dx\, f(x,y) \tag{15.b}$$

The discrimination criterion for independence of course concerns the comparison of $f(x,y)$ with $g(x,y) \equiv h_1(x)h_2(y)$. We now address the particular case $h_1 = h_2 \equiv h$, which is if course the most frequent case for financial quantities. Criterion (12) becomes

$$\int dxdy\, f(x,y)\frac{[\frac{f(x,y)}{h(x)h(y)}]^{q-1}-1}{q-1} \geq 0 \quad \left(q \geq \frac{1}{2}\right) \tag{16}$$

The evaluation if this quantity gives a satisfactory measure of the degree of dependence between x and y; when and only when it vanishes, x and y can be considered independent. In the $q \to 1$ limit, Eq. (16) becomes

$$\int dxdy\, f(x,y) \ln\, f(x,y) - 2\int dx\, h(x) \ln h(x) \geq 0 \tag{17}$$

thus recovering the expression (2.3) of Robinson 1991. For the $1/2 \leq q < 1$ cases, Eq. (16) becomes

$$\int dxdy [f(x,y)]^q [h(x)h(y)]^{1-q} \leq 1 \tag{18}$$

The $q = 1/2$ particular case becomes

$$\int dxdy \sqrt{f(x,y)h(x)h(y)} \leq 1 \tag{19}$$

For the $q > 1$ cases, Eq. (16) becomes

$$\int dxdy \frac{[f(x,y)]^q}{[h(x)h(y)]^{q-1}} \geq 1 \tag{20}$$

The $q = 2$ particular case becomes

$$\int dxdy [f(x,y)]^2/[h(x)h(y)] \geq 1 \tag{21}$$

In some sense, we can say that this is a satisfactory "quadratic", criterion, in contrast with the quantity ((5.3) of Robinson 1991)

$$\int dxdy [f(x,y)]^2 - (\int dx [h(x)]^2)^2 \tag{22}$$

basically introduced by Brock et al 1987. Indeed, as properly criticised by Robinson 1991, the quantity (22) has no definite sign and its zero value does not guarantee independence of x and y. In other words, it cannot be considered an optimal criterion, and could conveniently be replaced by the present criterion (21).

To conclude, let us stress that, for the $q \neq 0$ *generic* case, $I_q(f,g) \neq I_q(g,f)$ if $f \neq q$. Consequently, if we want, for some reason, to think of a reciprocal "distance" between f and g, it might be convenient to define a symetrized quantity, for instance

$$I_q^S(f,g) \equiv \frac{1}{2}[I_q(f,g) + I_q(g,f)] \tag{23}$$

hence $I_q^S(f,g) = I_q^S(g,f)$, $\forall (f,g)$. Adapting this to the problem of *measuring* the degree of dependence between x and y, we finally propose the use, for practical purposes, of the following family of criteria:

$$I_q^S(f(x,y), \{[\int dy\, f(x,y)][\int dx\, f(x,y)]\}) \geq 0 \quad \left(q \geq \frac{1}{2}\right) \tag{24}$$

The generalization for $d > 2$ is straightforward, namely

$$I_q^S(f(x_1, x_2, \cdots, x_d), g(x_1, x_2, \cdots, x_d)) \geq 0 \quad \left(q \geq \frac{1}{2}\right) \tag{25}$$

with

$$\begin{aligned} g(x_1, x_2, \cdots, x_d) \quad &\equiv \quad [\int dx_2 dx_3 \cdots dx_d f(x_1, \cdots, x_d)] \times \\ &\times \quad [\int dx_1 dx_3 \cdots dx_d f(x_1, \cdots, x_d)] \times \\ &\cdots \quad [\int dx_1 dx_2 \cdots dx_{d-1} f(x_1, \cdots, x_d)] \end{aligned} \tag{26}$$

The equality in (25) holds if and only if $(x_1, x_2, \cdots, x_d)$ are all independent. The study of the q-dependence of specific results (e.g., on financial time series) could be very enlightening.

I am indebted to R. Guenzburger for calling my attention onto Robinson 1991 work.

References

AHMAD, I.A. and LIN, P.-I. (1976), "A Nonparametric Estimate of the Entropy for Absolutely Continuous Distributions", IEEE Transactions on Information Theory, IT **22**, 372-375.

BINNEY, J. and TREMAINE, S. (1987), "Galactic Dynamics" (Princeton University Press, Princeton).

BROCK, W., DECHERT, D. and SCHEINKMAN, J. (1987), "A Test for Independence Based on the Correlation Dimension" (preprint, University of Wisconsin).

CHANDRASEKHAR, S. (1958), "An Introduction to the Theory of Stellar Structure" (University of Chicago Press, Chicago).

CURADO, E.M.F. and TSALLIS, C. (1991), "Generalized Statistical Mechanics: Connection with Thermodynamics", Journal of Physics A **24**, L 69-72; Corrigenda: A **24**, 3187 (1991) and A **25**, 1019 (1992).

DMITRIEV, Y.G. and TARASENKO, F.P. (1973), "On the Estimation of Functionals of the Probability Density and its Derivatives", Theory of Probability and its Applications, **18**, 628-633.

JUMARIE, G. (1988a), "Renyi Entropy of Maps: Applications to Fuzzy Sets, Pattern Recognition and Chaotic Dynamics", Cybernetics and Systems **19**, 169-192.

JUMARIE, G. (1988b), " New Results on the Entropy of Deterministic Maps: Applications to Liapunov Exponent, Divergence, Compositional Inference and Pattern Encoding" Cybernetics and Sysyems **19**, 311-338.

KULLBACK, S. and LEIBLER, R.A. (1961), "On Information and Sufficiency", Annals of Mathematical Statistics", **22**, 79-86.

LOSEE Jr. R.M. (1988), "The Science of Information-Measurement and Applications" (Academic Press).

MARIZ, A.M. (1992), "On the Irreversible Nature of the Tsallis and Renyi Entropies" Physics Letters A **165**, 409-411.

PLASTINO, A.R. and PLASTINO, A., (1993a), "Tsallis Entropy, Ehrenfest Theorem and Information Theory", to appear in Physics Letters A.

PLASTINO, A.R. and PLASTINO, A. (1993b), "Dynamical Aspects of Tsallis Entropy", preprint (University of La Plata).

PLASTINO, A.R. and PLASTINO, A. (1993c), "Stellar Polytropes and Tsallis Entropy" Physics Letters A **174**, 384-386.

PLASTINO, A. and TSALLIS, C. (1993), "Variational Method in Generalized Statistical Mechanics", preprint (Centro Brasileiro de Pesquisas Físicas).

RAMSHAW, J.D. (1993a), "H-Theorem for the Tsallis and Renyi Entropies", Physics Letters A **175**, 169-170.

RAMSHAW, J.D. (1993b), "Irreversibility and Generalized Entropies", Physics Letters A**175**, 171-172.

ROBINSON, P.M. (1991), "Consistent Nonparametric Entropy-Based Testing", Review of Economic Studies, **58**, 437-453.

SHANNON, C.E. (1948), "A Mathematical Theory of Communication", Bell System Technical Journal, **27**, 379-423, 623-656.

SILVER, R.N. and TSALLIS, C. (1993), "Generalized Entropy as a Possible Basis for a Nonlinear Schroedinger-Like Equation", preprint (Centro Brasileiro de Pesquisas Físicas).

STARIOLO, D.A. (1993), "The Langevin and Fokker-Planck Equations in the Framework of a Generalized Statistical Mechanics", preprint (Centro Brasileiro de Pesquisas Físicas).

TSALLIS, C. (1988), "Possible Generalization of Boltzmann-Gibbs Statistics", Journal of Statistical Physics **52**, 479-487.

VASICEK, O. (1976), "A Test for Normality Based on Sample Entropy", Journal of the Royal Statistical Society, Series B, **38**, 54-59.