

CORRECTION

A CLASS OF RÉNYI INFORMATION ESTIMATORS FOR MULTIDIMENSIONAL DENSITIES

Ann. Statist. **36** (2008) 2153–2182

BY NIKOLAI LEONENKO AND LUC PRONZATO

Cardiff University and CNRS/Université de Nice

In a recent paper [1], Leonenko, Pronzato and Savani consider the estimation of the Rényi and Tsallis entropies, respectively, $H_q^* = \frac{1}{1-q} \log \int_{\mathbb{R}^m} f^q(x) dx$ and $H_q = \frac{1}{q-1} (1 - \int_{\mathbb{R}^m} f^q(x) dx)$, $q \neq 1$, of an unknown probability measure with density f on \mathbb{R}^m with respect to the Lebesgue measure though k th nearest-neighbor distances in a sample X_1, \dots, X_N i.i.d. with the density f . The results in [1] about the asymptotic unbiasedness and consistency of the estimator proposed for $I_q = \mathbb{E}\{f^{q-1}(X)\} = \int_{\mathbb{R}^m} f^q(x) dx$ are correct for $q > 1$ but, for $q < 1$, convergence in distribution should be complemented by additional arguments to obtain the required convergence of moments.

Following [1], define $\hat{I}_{N,k,q} = \frac{1}{N} \sum_{i=1}^N (\zeta_{N,i,k})^{1-q}$, with $\zeta_{N,i,k} = (N-1)C_k \times V_m(\rho_{k,N-1}^{(i)})^m$, where $V_m = \pi^{m/2} / \Gamma(m/2 + 1)$ is the volume of the unit ball $\mathcal{B}(0, 1)$ in \mathbb{R}^m , $C_k = [\frac{\Gamma(k)}{\Gamma(k+1-q)}]^{1/(1-q)}$ and $\rho_{k,N-1}^{(i)}$ denote the k th nearest-neighbor distance from a given X_i to some other X_j in the sample X_1, \dots, X_N . Also define $r_c(f) = \sup\{r > 0: \int_{\mathbb{R}^m} |x|^r f(x) dx < \infty\}$, so that $\mathbb{E}|X_i|^r < \infty$ if $r < r_c(f)$ and $\mathbb{E}|X_i|^r = \infty$ if $r > r_c(f)$ (see [2]).

A correct version of the convergence results of $\hat{I}_{N,k,q}$ to I_q for $0 < q < 1$ and f having unbounded support can be obtained by using results on the sub-additivity of Euclidean functionals (see [3]) and can be formulated as follows [2]: if $I_q < \infty$ and $r_c(f) > m \frac{1-q}{q}$, then $\mathbb{E}[\hat{I}_{N,k,q}] \rightarrow I_q$, $N \rightarrow \infty$ (asymptotic unbiasedness; see Theorem 3.1 of [1]); if $I_q < \infty$, $q > \frac{1}{2}$ and $r_c(f) > 2m \frac{1-q}{2q-1}$, then $\mathbb{E}[\hat{I}_{N,k,q} - I_q]^2 \rightarrow 0$, $N \rightarrow \infty$ (L_2 convergence; see Theorem 3.2 of [1]). The situation is simpler when f has bounded support \mathcal{S}_f . When f is bounded away from 0 and infinity on \mathcal{S}_f and \mathcal{S}_f consists of a finite union of convex bounded sets with

Editors' note: The standard *Annals of Statistics* policy is that brief author responses and corrigenda are only reviewed for style and appropriateness; authors themselves are responsible for their correctness.

Received November 2009.

AMS 2000 subject classifications. 94A15, 62G20.

Key words and phrases. Entropy estimation, estimation of statistical distance, estimation of divergence, nearest-neighbor distances, Rényi entropy, Havrda–Charvát entropy, Tsallis entropy.

nonempty interiors, then $\hat{I}_{N,k,q}$ converges a.s. and in the L_2 sense to I_q for $q < 1$ (see [2]). However, when $\inf_{x \in S_f} f(x) = 0$ and $q < 1$, the results obtained indicate that k should be selected in relation with the growth rate of f close to 0 (details are available from the authors and will be presented elsewhere).

Acknowledgments. The authors express their deep thanks to J. Yukich who pointed out the mistake in [1] when $q < 1$ and to M. Penrose for helpful discussions around the recent paper [2].

REFERENCES

- [1] LEONENKO, N., PRONZATO, L. and SAVANI, V. (2008). A class of Rényi information estimators for multidimensional densities. *Ann. Statist.* **36** 2153–2182. [MR2458183](#)
- [2] PENROSE, M. D. and YUKICH, J. E. (2011). Laws of large numbers and nearest neighbor distances. *Advances in Directional and Linear Statistics, A Festschrift for Sreenivasa Rao Jammalamadaka* (M. Wells and A. Sengupta, eds.). Springer. To appear. Available at [arXiv:0911.0331v1](#).
- [3] YUKICH, J. E. (1998). *Probability Theory of Classical Euclidean Optimization Problems*. Springer, Berlin. [MR1632875](#)

CARDIFF SCHOOL OF MATHEMATICS
CARDIFF UNIVERSITY
SENGHENYDD ROAD
CARDIFF CF24 4AG
UK
E-MAIL: leonenkon@Cardiff.ac.uk

LABORATOIRE I3S
CNRS/UNIVERSITÉ DE NICE–SOPHIA ANTIPOLIS
LES ALGORITHMES, 2000 ROUTE DES LUCIOLES
BP 121, 06903 SOPHIA-ANTIPOLIS CEDEX
FRANCE
E-MAIL: pronzato@i3s.unice.fr