The Annals of Statistics
2010, Vol. 38, No. 6, 3837–3838
DOI: 10.1214/10-AOS773

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 kth 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 kth 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 subadditivity 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 $\mathbb{E}[\hat{I}_{N,k,q}] \to I_q$, $N \to \infty$ (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 $\mathbb{E}[\hat{I}_{N,k,q} - I_q]^2 \to 0$, $N \to \infty$ (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.