Ann. Inst. Statist. Math.
Vol. 41, No. 4, 683–697 (1989)

ESTIMATION OF ENTROPY AND OTHER FUNCTIONALS OF A MULTIVARIATE DENSITY*

HARRY JOE

Department of Statistics, University of British Columbia, 2021 West Mall, Vancouver, B.C., Canada V6T 1W5

(Received August 25, 1987; revised March 7, 1989)

Abstract. For a multivariate density f with respect to Lebesgue measure μ, the estimation of $\int J(f) f \, d\mu$, and in particular $\int f^2 d\mu$ and $-\int f \log f \, d\mu$, is studied. These two particular functionals are important in a number of contexts. Asymptotic bias and variance terms are obtained for the estimators $\hat{I} = \int J(\hat{f}) dF_n$ and $\tilde{I} = \int J(\hat{f}) \hat{f} \, d\mu$, where $\hat{f}$ is a kernel density estimate of f and F_n is the empirical distribution function based on the random sample $X_1, \ldots, X_n$ from f. For the two functionals mentioned above, a first order bias term for $\hat{I}$ can be made zero by appropriate choices of non-unimodal kernels. Suggestions for the choice of bandwidth are given; for $\hat{I} = \int \hat{f} dF_n$, a study of optimal bandwidth is possible.

Key words and phrases: Kernel density estimation, multivariate density, empirical process, entropy.