Kernel Density Estimation : the General Case

by

C.C.Y. Dorea
Departamento de Matemática
Universidade de Brasilia
70910-900 Brasilia-DF - Brazil
cdorea@mat.unb.br

Abstract

For estimating a density function $p(x)$ the kernel weighted average over the empirical distribution $F_n(\cdot)$ constitute a general class of consistent and asymptotically normal estimators ,

$$p_n(x) = \int_{-\infty}^{+\infty} \frac{1}{h} K(\frac{x-y}{h})\, dF_n(y) = \frac{1}{nh} \sum_{k=1}^{n} K(\frac{x-X_k}{h})$$

where the kernel $K(\cdot)$ and the window $h = h_n > 0$ are suitably chosen, and $X_1, X_2, \cdots, X_n$ are independent random variables with a common density $p(x)$.

In this note, we consider the problem of kernel estimates in the general case. Let $p(x)$ be a density with respect to a σ-finite measure ν on $(E, \mathcal{E})$ where $E \subset R^d$ and $\mathcal{E}$ is a σ-field of subsets of E. For each $h > 0$ and each $x \in E$ let $W(h, x, .)$ be defined on E and define the estimates

$$p_n(x) = \frac{1}{nh} \sum_{k=1}^{n} W(h, x, X_k) \ , \ h = h_n$$

where $X_1, X_2, \cdots, X_n$ are independent and have a common density $p(x)$.

We present sufficient conditions for $p_n(x)$ to be consistent in quadratic mean, strongly consistent and asymptotically normal. Also, it is shown that our results include the classical results for continuous densities on R^d and extend some of the results of kernel estimates for discrete distributions.