Institute of Mathematical Statistics

LECTURE NOTES — MONOGRAPH SERIES

FIXED DESIGN REGRESSION UNDER ASSOCIATION

George G. Roussas
University of California, Davis

Abstract

For $n = 1, \ldots, n$, let $x_{ni}, i = 1, \ldots, n$, be points in a compact subset in $\Re^d, d \geq 1$, at which observations Y_{ni} are taken. It is assumed that these observations have the structure $Y_{ni} = g(x_{ni}) + \varepsilon_{ni}$, where g is a real-valued unknown function, and the errors $(\varepsilon_{n1}, \ldots, \varepsilon_{nn})$ coincide with the segment $(\xi_1, \ldots, \xi_n)$ of a strictly stationary sequence of random variables $\xi_1, \xi_2, \ldots$. For each $x \in \Re^d$, the function $g(x)$ is estimated by $g_n(x; \boldsymbol{x}_n) = \sum_{i=1}^n w_{ni}(x; \boldsymbol{x}_n) Y_{ni}$, where $\boldsymbol{x}_n = (x_{n1}, \ldots, x_{nn})$ and $w_{ni}(\cdot; \cdot)$ are weight functions. Under suitable conditions on the underlying stochastic process $\xi_1, \xi_2, \ldots$ and the weights $w_{ni}(\cdot; \cdot)$, it is shown that the estimate $g_n(x; \boldsymbol{x}_n)$ is asymptotically unbiased, and consistent in quadratic mean. By adding the assumption of (positive or negative) association of the sequence $\xi_1, \xi_2, \ldots$, it is shown that $g_n(x; \boldsymbol{x}_n)$, properly normalized, is also asymptotically normal.

Key words and phrases: Fixed design regression, stationarity, weights, fixed design regression estimate, asymptotic unbiasedness, consistency in quadratic mean, association, asymptotic normality.

1 Introduction

For each natural number n, consider the design points x_{ni}, $i = 1, \ldots, n$ in $\Re^d$, $d \geq 1$, which, through a real-valued (Borel) function g defined on $\Re^d$, produce observations Y_{ni}, subject to errors ε_{ni}, $1 \leq i \leq n$. That is,

$$Y_{ni} = g(x_{ni}) + \varepsilon_{ni}, \qquad 1 \leq i \leq n. \tag{1.1}$$

It is eventually assumed that, for each n, $(\varepsilon_{n1}, \ldots, \varepsilon_{nn})$ is equal in distribution to $(\xi_1, \ldots, \xi_n)$, where $\{\xi_n\}$, $n \geq 1$, is a (strictly) stationary and (positively or negatively) associated (see Definition 1.1) sequence of random variables (r.v.s). The problem we are faced with here is that of estimating