4

Local Weighting of Functional Variables

In the finite dimensional case, the local weighting techniques are very popular in the community of nonparametricians because they are very well adapted to nonparametric models. The aim of this chapter is to explain how the concept of local smoothing can be extended to the functional data case. Clearly, local approaches need to have at hand some topological ways for measuring proximity between functional data, and therefore this chapter will be directly linked with the semi-metric modelling discussed in Chapter 3.

In the finite dimensional case, one of the most common approaches among these local weighting methods is certainly the kernel one. It is impossible to give an exhaustive bibliography about nonparametric methods for finite dimensional variables, but the state of art in this field is well summarized in [S00] and [AP03] while a large number of references can be found in [SV00] concerning the kernel methods especially. We will see in this chapter how kernel smoothing ideas can be adapted to infinite dimensional variables.

The chapter is organized as follows. In Section 4.1 we give a basic discussion on kernel method, explaining how (and why) what is classically done for finite dimensional variables can be adapted to functional setting. The second aim (Section 4.2) consists in seeing how the local weighting is in relation to the notion of small ball probabilities. As we will see, small ball probabilities can be viewed as a tool for describing some local behaviours of functional data and the kernel approach allows us to take into account these kinds of local properties. Section 4.3 proposes some general theoretical aspects concerning kernel weighting. Finally, note that Sections 4.1 and 4.2 is of interest to a large public whereas Section 4.3 is meant for statisticians interested in theoretical aspects.

4.1 Why Use Kernel Methods for Functional Data?

Kernel methods are well-known and intensively used by the community of nonparametricians because they are a useful way to do local weighting. We start

by recalling shortly what is kernel local weighting in the real and multivariate cases before extending it to the functional context.

4.1.1 Real Case

As it is well known, kernel local weighting is based on a kernel function (classically denoted by K) and on a smoothing parameter, which is called bandwidth and usually denoted by h. If x is a fixed real number, the kernel local weighting transforms n r.r.v. $X_1, X_2, \ldots, X_n$ into $\Delta_1, \Delta_2, \ldots, \Delta_n$ such that:

$$\Delta_i \ = \ \Delta_i(x, h, K) \ = \frac{1}{h} K \left(\frac{x - X_i}{h} \right).$$

The main idea of the local weighting around x is to attribute at each r.r.v. X_i a weight taking into account the distance between x and X_i; the more X_i is distant from x, the smaller is the weighting.

Before going on, let us recall what is a kernel function exactly in this simplest situation. In fact, there exists a large variety of kernels. Any density function can be considered as a kernel, but even unnecessary positive functions can be used ([GM84]). A large literature exists on this field (see [MN89] and [B93] for interesting advances and [HVZ] for a presentation of the state of art). To simplify our purpose, we consider at this stage only positive and symmetrical kernels which are the most classical ones. Figure 4.1 displays various kernel functions which are analytically defined as follows:

(a) *Box* kernel: $K(u) = \frac{1}{2} 1_{[-1,+1]}(u)$,
(b) *Triangle* kernel: $K(u) = (u + 1)1_{[-1,0]}(u) + (1 - u)1_{[0,+1]}(u)$,
(c) *Quadratic* kernel: $K(u) = \frac{3}{4}(1 - u^2)1_{[-1,+1]}(u)$,
(d) *Gaussian* kernel: $K(u) = \frac{1}{\sqrt{2\pi}} \exp\{-\frac{u^2}{2}\}$.

To precise the notion of kernel local weighting, let us consider the Box kernel and rewrite the Δ_i's as follows:

$$\Delta_i \ = \frac{1}{h} 1_{[x-h,x+h]}(X_i).$$

In this situation, the local feature of the weighting appears obvious since the r.r.v. outside the range $[x-h, x+h]$ are ignored. In addition, the normalization $1/h$ is proportional to the size of the set $[x - h, x + h]$ on which the X_i's are taken into account. These points are not only true for the Box kernel, but are shared by any compactly supported kernels.