Práctico 5

Regresión no paramétrica

1. Considere la función

$$r(x) = 10\sin(x) + x^2\,, \quad x \in [0,6]$$

   y el modelo $Y_i = r(x_i) + \epsilon_i$.

   (a) A partir del siguiente código en el lenguaje R, simule 100 valores de la v.a Y_i.

```
r<-function(x){10*sin(x) +x^2}
x<-runif(100,0,6)
epsilon<-rnorm(100,0,3)
```

   (b) Escriba un código en R para hallar la estimación no-paramétrica de la función de regresión r a través de los diferentes métodos expuestos en clase.

2. Considere el conjunto de datos *Cosmic Microwave Background (CMB)* empleado en [1], disponible en
   `http://www.stat.cmu.edu/~larry/all-of-nonpar/data.html`.

   (a) Estimar la función de regresión en el rango de valores $[0, 400]$ de la variable de entrada momento multipolar (referirse al panel inferior de la Figura 5.1, página 62, [1]) a través de los siguientes métodos no paramétricos:

      i. Regresograma.
      ii. Promedios locales.
      iii. Núcleos.
      iv. k-ésimo vecino más cercano. [1]
      v. Regresión polinomial local de orden 1.

   (b) Estimar la función de regresión en todo el recorrido de valores de la variable de entrada momento multipolar (referirse al panel superior de la Figura 5.1, página 62, [1]) teniendo en cuenta que no se cumple el supuesto de homocedasticidad.

[1] Ver el paquete de R "**FNN**"

3. Considere la función de *Doppler*:

$$r(x) = \sqrt{x(1-x)}\sin\left(\frac{2,1\pi}{x+0,05}\right) , \quad x \in [0,1] .$$

Generar 1000 observaciones del modelo $Y_i = r(x_i) + \sigma\epsilon_i$ donde $x_i = i/n$, y $\epsilon_i \sim N(0,1)$. Estimar la función r, para $\sigma = 1/10, 1\, y\, 3$ a través de:

(a) i. k-ésimo vecino más cercano. [2]
ii. Núcleos.[3]
iii. Regresión polinomial local de orden 1.

(b) Graficar los datos y las funciones estimadas para cada σ.

Bibliografía

[1] Larry Wasserman (2006). *All of Nonparametric Statistics*, Springer.

[2] Ver el paquete de R "**FNN**"
[3] Ver la función en R llamada **ksmooth()**