José Antonio Álvarez Ocete
77553417Q
joseantonio.alvarezo@estudiante.uam.es

Ejercicio 6.

Enunciado. Sean $X_1, \dots, X_n$ variables aleatorias *i.i.d.* de una distribución con densidad f. Se considera el estimador del núcleo $\hat{f}$ con núcleo rectangular $K(x) = \mathbb{I}_{[-1/2,1/2]}(x)$ y parámetro de suavizado h.

a) Calcula el sesgo y la varianza de $\hat{f}(x)$, para un valor de x fijo.

b) Demuestra que tanto el sesgo como la varianza tienden a cero si $h \to 0$ y $nh \to \infty$.

Al estudiar el sesgo y la varianza de $\hat{f}(x)$ para un valor de x fijo, estudiamos dichos valores respecto a la muestra tomada de $X_1, \dots, X_n$. En primer lugar, calculemos la esperanza del núcleo rectangular dado por la función indicatriz:

$$\mathbb{E}\left[K\left(\frac{x-t}{h}\right)\right] = \int_{-\infty}^{+\infty} f(t)\ K\left(\frac{x-t}{h}\right)\ dt$$

Utilizamos el cambio de variable $w = \frac{x-t}{h}, dw = -\frac{dt}{h}$ e invirtiendo los límites de integración obtenemos:

$$\begin{aligned}
\mathbb{E}\left[K\left(\frac{x-t}{h}\right)\right] &= \int_{-\infty}^{+\infty} f(t)\ K\left(\frac{x-t}{h}\right)\ dt \\
&= \int_{+\infty}^{-\infty} f(x-wh)\ K(w)\ (-h)\ dw \\
&= h\int_{-\infty}^{+\infty} f(x-wh)\ \mathbb{I}_{[-1/2,1/2]}(w)\ dw \\
&= h\int_{-\frac{1}{2}}^{\frac{1}{2}} f(x-wh) dw
\end{aligned} \tag{1}$$

Calculemos el sesgo de nuestro estimador del núcleo haciendo uso de la expresión anterior. Para ello calculamos su esperanza:

$$\begin{aligned}
\mathbb{E}[\hat{f}(x)] &= \mathbb{E}\left[\frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x-t}{h}\right)\right] \\
&\overset{\text{iid}}{=} \frac{1}{nh}\ n\mathbb{E}\left[K\left(\frac{x-t}{h}\right)\right] \\
&\overset{(1)}{=} \frac{1}{h}\ h\int_{-\frac{1}{2}}^{\frac{1}{2}} f(x-wh) dw \\
&= \int_{-\frac{1}{2}}^{\frac{1}{2}} f(x-wh) dw
\end{aligned}$$

Este valor únicamente depende del punto fijado x, el parámetro de suavizado h y la función de densidad f. Su valor equivale al área bajo la gráfica de f en el intervalo $[x - h/2, x + h/2]$, como se puede apreciar en la figura 1.

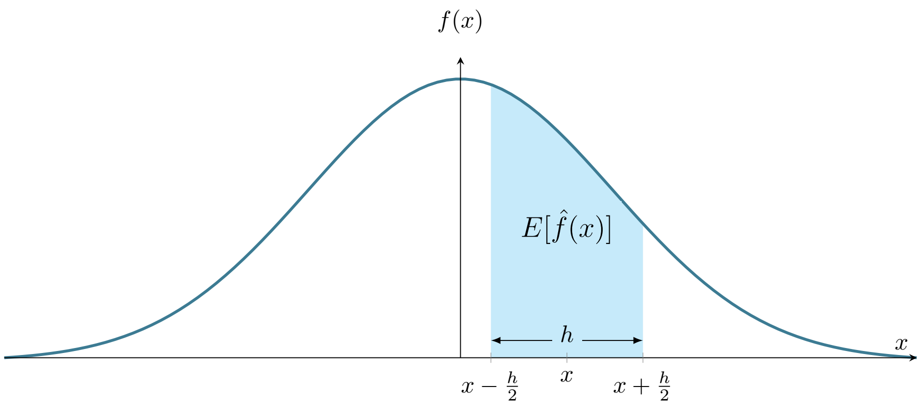

Figure 1: Representación gráfica del valor $\mathbb{E}[\hat{f}(x)]$.

Es decir, estamos aproximando el valor de una función en un punto por su integral en un intervalor centrado en dicho punto. Naturalmente, al tomar $h \to 0$, dicho valor tiende al valor del punto. Es decir, el sesgo tiende a 0. Analíticamente:

$$\begin{aligned}
\text{Sesgo}(f) &= E[\hat{f}(x)] - f(x) \\
&= \int_{-\frac{1}{2}}^{\frac{1}{2}} f(x-wh)dw - f(x) \\
&\xrightarrow{h\to\infty} \int_{-\frac{1}{2}}^{\frac{1}{2}} f(x)dw - f(x) \\
&= f(x) \underbrace{\int_{-\frac{1}{2}}^{\frac{1}{2}} dw}_{=1} - f(x) \\
&= f(x) - f(x) = 0
\end{aligned}$$

Por otro lado, podemos darnos cuenta de lo siguiente:

$$K(x)^2 = K(x) \quad \forall x \in \mathbb{R}, \tag{2}$$

pues la función indicatriz tiene por imagen el conjunto $\{0, 1\}$ y la función $x \mapsto x^2$ sobre este conjunto es la identidad. Calculemos la varianza del estimador del núcleo:

$$\begin{aligned}
\text{Var}[\hat{f}(x)] &= \text{Var}\left[\frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x-t}{h}\right)\right] \\
&\overset{\text{iid}}{=} \frac{1}{n^2h^2}\, n\text{Var}\left[K\left(\frac{x-t}{h}\right)\right] \\
&= \frac{1}{nh^2}\left(\mathbb{E}\left[K\left(\frac{x-t}{h}\right)^2\right] - \mathbb{E}\left[K\left(\frac{x-t}{h}\right)\right]^2\right) \\
&\overset{(2)}{=} \frac{1}{nh^2}\left(\mathbb{E}\left[K\left(\frac{x-t}{h}\right)\right] - \mathbb{E}\left[K\left(\frac{x-t}{h}\right)\right]^2\right) \\
&\overset{(1)}{=} \frac{1}{nh^2}\left(h\int_{-\frac{1}{2}}^{\frac{1}{2}} f(x-wh)dw - h^2\left(\int_{-\frac{1}{2}}^{\frac{1}{2}} f(x-wh)dw\right)^2\right) \\
&= \frac{1}{nh}\left(\int_{-\frac{1}{2}}^{\frac{1}{2}} f(x-wh)dw - h\left(\int_{-\frac{1}{2}}^{\frac{1}{2}} f(x-wh)dw\right)^2\right)
\end{aligned}$$

Finalmente probamos que la varianza tiende a 0 si $h \to 0$ y $nh \to \infty$:

$$\text{Var}[\hat{f}(x)] = \underbrace{\frac{1}{nh}}_{\hookrightarrow 0} \left(\underbrace{\int_{-\frac{1}{2}}^{\frac{1}{2}} f(x-wh)dw}_{\hookrightarrow f(x)} - \underbrace{h \left(\int_{-\frac{1}{2}}^{\frac{1}{2}} f(x-wh)dw \right)^2}_{\hookrightarrow 0} \right) \longrightarrow 0$$

Aunque ya probamos que el error cuadrático medio para un núcleo cualquiera tiende a 0, en este ejercicio lo hemos probado para el caso del núcleo indicatriz sin el uso de aproximaciones.