

**Sobre a utilização de supernúcleos na  
estimação duma densidade de  
probabilidade**

Ana Raquel Carraco Marques





# Sobre a utilização de supernúcleos na estimação duma densidade de probabilidade

Ana Raquel Carraco Marques

Dissertação para a obtenção do Grau de **Mestre em Matemática**  
Área de Especialização em **Estatística, Optimização e Matemática Financeira**

## Júri

**Presidente:** Professor Doutor Paulo Eduardo Aragão Aleixo e Neves de Oliveira

**Orientador:** Professor Doutor Carlos Manuel Rebelo Tenreiro da Cruz

**Vogal:** Professora Doutora Maria Emília de Mesquita Nogueira

**Data:** Agosto de 2013



# Resumo

Neste trabalho sobre a utilização de supernúcleos na estimação da densidade de probabilidade, depois de uma breve introdução aos estimadores do núcleo da densidade de probabilidade, estabelecemos a convergência em média quadrática integrada de tais estimadores e obtemos expressões assintóticas para o erro quadrático médio integrado nos casos em que o estimador do núcleo é baseado num núcleo de ordem finita  $k \geq 2$  ou num supernúcleo. Tendo em conta que os estimadores baseados num supernúcleo não são densidades de probabilidade pois podem tomar valores negativos, estudamos a seguir uma correcção proposta por Glad, Hjort e Ushakov (*Scandinavian J. Statist.* 30, 415–247, 2003) que permitirá transformá-los em estimadores próprios da densidade tendo o estimador corrigido um erro quadrático médio integrado não superior ao do estimador original. Finalmente, apresentamos um estudo de simulação com o objectivo de comparar os estimadores do núcleo baseados no núcleo normal (núcleo 2.<sup>a</sup> ordem) e no núcleo trapezoidal (supernúcleo).

**Palavras Chave:** Estimação da densidade de probabilidade, núcleos de 2.<sup>a</sup> ordem, supernúcleos, erro quadrático médio integrado

# Abstract

In this work on the use of superkernels in the estimation of a probability density function, after a brief introduction to the probability density kernel estimator we establish its mean integrated square error consistency and we obtain asymptotic expressions for the mean integrated square error of kernel estimators based on either finite order kernels or superkernels. Taking into account that superkernel based estimators are not proper density estimators because they may assume negative values, we study a correction proposed by Glad, Hjort and Ushakov (*Scandinavian J. Statist.* 30, 415–247, 2003), that turns any density estimator which integrates to 1 into one which is a proper density estimator with an inferior integrated square error. Finally, we present a simulation study in order to compare the estimators based on the normal kernel (second order kernel) and on the trapezoidal kernel (superkernel).

**Keywords:** Probability density estimation, second order kernels, superkernels, mean integrated squared error



# Agradecimentos

*Ao meu orientador, Professor Doutor Carlos Tenreiro, pela sua dedicação, paciência e conselhos transmitidos ao longo deste ano, que foram muito importantes para o desenvolvimento deste trabalho.*

*Agradeço também aos meus familiares e amigos pelo seu acompanhamento e motivação ao longo do meu curso.*





# Conteúdo

<b>1</b>	<b>Introdução</b>	<b>1</b>
<b>2</b>	<b>Comportamento assintótico do Erro Quadrático Médio Integrado</b>	<b>5</b>
2.1	Convergência em média quadrática integrada . . . . .	6
2.2	Desenvolvimento assintótico do EQMI . . . . .	8
2.2.1	Núcleos de 2 <sup>a</sup> ordem . . . . .	8
2.2.2	Núcleos de ordem finita . . . . .	10
2.2.3	Supernúcleos . . . . .	11
<b>3</b>	<b>A correcção de Glad-Hjort-Ushakov</b>	<b>15</b>
3.1	O estimador $\tilde{f}_n$ está bem definido . . . . .	16
3.2	Erro quadrático integrado de $\tilde{f}_n$ . . . . .	18
3.3	Outra correcção . . . . .	23
<b>4</b>	<b>Estudo comparativo: Núcleo de 2.<sup>a</sup> ordem vs. Supernúcleo</b>	<b>25</b>
4.1	Comparação baseada no EQMI . . . . .	25
4.1.1	Desenvolvimento exacto do EQMI . . . . .	25
4.1.2	Densidades de teste . . . . .	27
4.1.3	Resultados numéricos . . . . .	28
4.2	Comparação baseada no EQI . . . . .	28
4.2.1	Escolha da janela pelo método <i>plug-in</i> . . . . .	30
4.2.2	Escolha da janela pelo método de validação cruzada . . . . .	35
	<b>Bibliografia</b>	<b>41</b>



# Capítulo 1

## Introdução

Dada uma amostra  $X_1, \dots, X_n$ , de variáveis aleatórias reais independentes e identicamente distribuídas com densidade de probabilidade  $f$ , desconhecida, o estimador do núcleo da densidade, introduzido por Rosenblatt (1956) e por Parzen (1962), é definido, para  $x \in \mathbb{R}$ , por

$$f_n(x) = \frac{1}{n} \sum_{i=1}^n K_{h_n}(x - X_i),$$

com

$$K_h(x) = \frac{1}{h} K\left(\frac{x}{h}\right),$$

onde  $K$  é uma função real de variável real integrável, satisfazendo

$$\int K(x) dx = 1,$$

e  $(h_n)$  uma sucessão de números reais estritamente positivos convergindo para zero quando  $n \rightarrow +\infty$ . A função  $K$  é designada por núcleo e  $h_n$  é designada por janela.

Os núcleos mais utilizados na prática são o núcleo normal standard definido por

$$K(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}},$$

e o núcleo de Epanechnikov definido por

$$K(x) = \frac{3}{4} (1 - x^2) I(|x| \leq 1).$$

Este último núcleo goza de certas propriedades de optimalidade descritas por Epanechnikov (1969).

Os núcleos apresentados anteriormente são densidades de probabilidade simétricas, satisfazendo as condições

$$m_1(K) = 0 \text{ e } m_2(K) > 0$$

onde  $m_j(K)$ , para  $j = 1, 2, \dots$ , é o momento de ordem  $j$  de  $K$  definido por

$$m_j(K) = \int x^j K(x) dx,$$

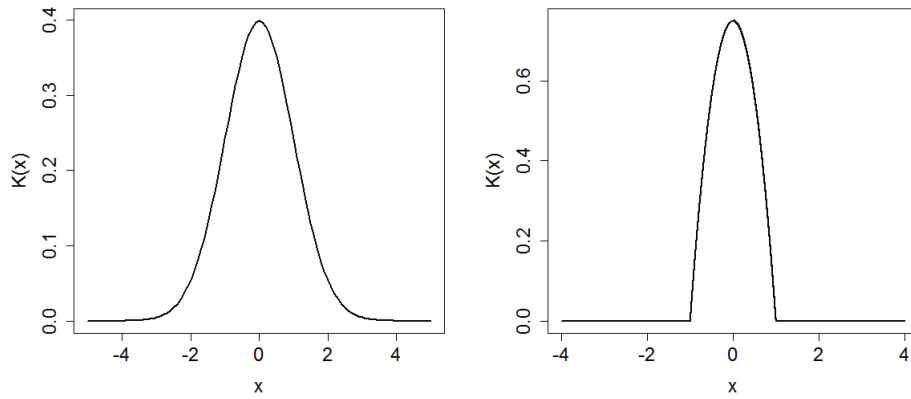


Figura 1.1: Núcleos Normal e de Epanechnikov.

sempre que  $\int |x|^j |K(x)| dx < \infty$ . Dizemos, neste caso, que os núcleos anteriores são núcleos de 2.<sup>a</sup> ordem. Atendendo ao facto de as propriedades do núcleo serem transferidas para o estimador do núcleo, quando o núcleo  $K$  é uma densidade o estimador do núcleo,  $f_n$ , também o é.

É, no entanto, possível tomar para núcleo funções que admitem momentos nulos até à ordem  $k > 2$ , isto é,

$$m_j(K) = 0 \text{ para } j = 1, \dots, k - 1 \text{ e } m_k(K) \neq 0.$$

Tais núcleos, designados por núcleos de ordem  $k$ , tomam necessariamente valores negativos quando  $k > 2$ , o que faz com que o estimador do núcleo neles baseado não seja mais uma densidade de probabilidade.

Existem vários métodos para construir núcleos de ordem superior à segunda (por exemplo, ver Ruppert e Wand, 1992; Jones e Foster, 1993). Um exemplo de núcleo de ordem 4, construído a partir do núcleo normal, é dado por

$$K(x) = \frac{1}{2} (3 - x^2) \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

que admite a seguinte representação gráfica apresentada na Figura 1.2.

Neste trabalho, estamos particularmente interessados numa classe de núcleos que admitem momentos nulos de todas as ordens (Devroye, 1987, pg. 100). Tais núcleos, designados supernúcleos, são caracterizados pelo facto de possuírem uma transformada de Fourier,

$$\varphi_K(t) = \int e^{itx} K(x) dx, \quad t \in \mathbb{R},$$

constante numa vizinhança da origem (ver Chácon et al., 2007a; Chácon et al., 2007b).

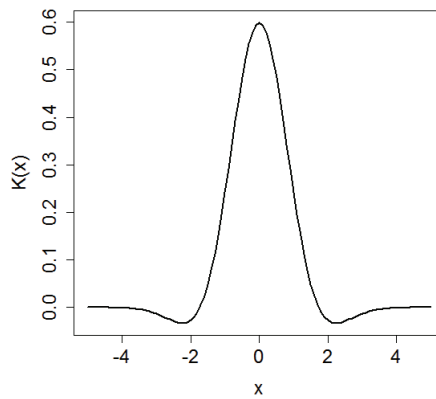


Figura 1.2: Núcleo de ordem 4 baseado no núcleo Normal.

Um exemplo de supernúcleo é o chamado supernúcleo trapezoidal

$$K(x) = \begin{cases} \frac{\cos x - \cos(2x)}{\pi x^2} & , x \neq 0 \\ \frac{3}{2\pi} & , x = 0 \end{cases}$$

com transformada de Fourier

$$\varphi_K(t) = I(0 \leq |t| < 1) + (2 - |t|)I(1 \leq |t| < 2).$$

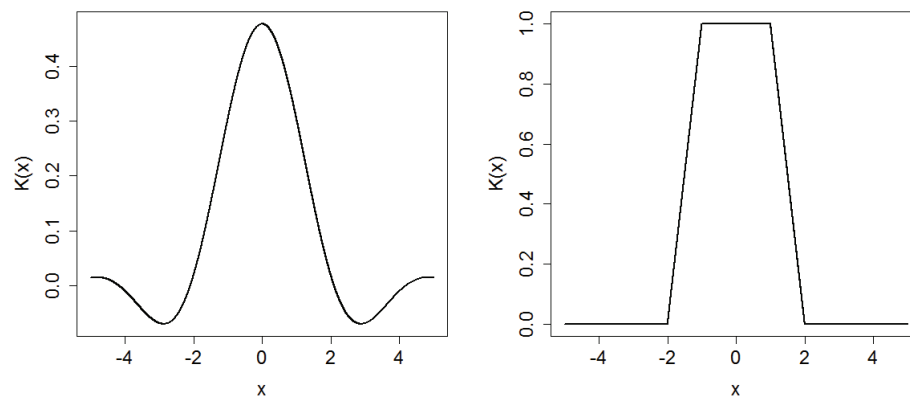


Figura 1.3: Supernúcleo trapezoidal e respectiva transformada de Fourier.

Neste trabalho, sobre a utilização de supernúcleos na estimação duma densidade de probabilidade, centraremos a nossa atenção no erro quadrático integrado

$$\text{EQI}(f_n) = \int \{f_n(x) - f(x)\}^2 dx$$

e no erro quadrático médio integrado

$$\text{EQMI}(f_n) = E \int \{f_n(x) - f(x)\}^2 dx$$

como medidas globais da qualidade de  $f_n$  enquanto estimador de  $f$ .

No Capítulo 2, começaremos por estabelecer a convergência em média quadrática integrada do estimador do núcleo sob condições gerais sobre o núcleo  $K$  e a densidade desconhecida  $f$ . De seguida, impondo condições de regularidade sobre a densidade  $f$ , apresentamos um desenvolvimento assintótico para o erro quadrático médio integrado que nos permite obter uma expressão para a janela que minimiza o erro quadrático médio integrado assintótico. No caso dos estimadores baseados num núcleo de ordem  $k > 2$ , iremos observar que estes têm potencialmente um viés,  $E(f_n(x)) - f(x)$ , inferior ao do estimador do núcleo baseado num núcleo de 2.<sup>a</sup> ordem.

De seguida, no Capítulo 3, tendo em conta que os núcleos de ordem  $k > 2$  e os supernúcleos não são densidades de probabilidades, pois podem tomar valores negativos, vamos estudar uma correcção proposta por Glad, Hjort e Ushakov (2003), que permitirá transformar os estimadores neles baseados em densidades de probabilidades, com um erro quadrático integrado inferior ao do estimador original.

Finalmente, no Capítulo 4, apresentamos um estudo de simulação, com o objectivo de comparar o estimador do núcleo baseado no núcleo normal standard com o estimador do núcleo baseado no supernúcleo trapezoidal. Iniciamos este estudo comparando os EQMI associados às janelas óptimas, no sentido da minimização do EQMI, cuja existência é estabelecida por Chácon et al. (2007b), e às janelas assintoticamente óptimas de cada um dos estimadores, num sentido que precisaremos mais à frente. Estas comparações teóricas são a seguir ilustradas através de um estudo de simulação em que a escolha da janela será feita de forma automática, recorrendo a dois métodos: o método *plug-in* a duas etapas e o método de validação cruzada.

## Capítulo 2

# Comportamento assintótico do Erro Quadrático Médio Integrado

Uma das medidas mais usadas para avaliar a qualidade global dum estimador  $f_n$  da densidade  $f$  é o erro quadrático médio integrado, definido por

$$\text{EQMI}(f_n) = \text{E} \int \{f_n(x) - f(x)\}^2 dx$$

(ver Rosenblatt, 1956). Esta medida de discrepância entre  $f_n$  e  $f$ , não é mais do que o integral, estendido a  $\mathbb{R}$ , do erro quadrático médio de  $f_n$ , no ponto  $x$ , dado por

$$\begin{aligned} \text{EQM}(f_n(x)) &= \text{E}(f_n(x) - f(x))^2 \\ &= \text{Var}(f_n(x)) + \text{Viés}(f_n(x))^2, \end{aligned}$$

onde

$$\text{Viés}(f_n(x)) = \text{E}(f_n(x)) - f(x).$$

Assim, reescrevendo a expressão do erro quadrático médio integrado obtemos

$$\text{EQMI}(f_n) = \text{IVAR}(f_n) + \text{IVIES}(f_n) \quad (2.1)$$

onde

$$\text{IVAR}(f_n) = \int \text{Var}(f_n(x)) dx \quad (2.2)$$

e

$$\text{IVIES}(f_n) = \int \text{Viés}(f_n(x))^2 dx. \quad (2.3)$$

Neste capítulo, começaremos por estudar a convergência em média quadrática integrada do estimador do núcleo sob condições gerais sobre o núcleo  $K$  e a densidade  $f$ . De seguida, e já sob condições de regularidade sobre a densidade  $f$ , apresentamos desenvolvimentos assintóticos para o erro quadrático médio integrado do estimador do núcleo, a partir dos quais deduziremos expressões teóricas para a janela óptima (num sentido a precisar) do estimador do núcleo.

No caso dos núcleos de ordem finita, seguiremos de perto a exposição de Tenreiro (2010), enquanto que no caso dos supernúcleos os resultados apresentados são devidos a Chácon et al. (2007a).

## 2.1. Convergência em média quadrática integrada

Nesta secção mostramos que o erro quadrático médio integrado do estimador do núcleo, definido por (2.1), converge para zero, quando  $n \rightarrow +\infty$ , sempre que  $f$  e  $K$  sejam de quadrado integrável e a janela  $h_n$  satisfaça as condições  $h_n \rightarrow 0$  e  $nh_n \rightarrow +\infty$ . É de salientar que todos os núcleos apresentados anteriormente são de quadrado integrável.

Sendo  $g$  e  $h$  funções de quadrado integrável, ao longo deste trabalho, usaremos a notação

$$R(g) = \int g(x)^2 dx$$

e, representamos por  $g * h$  o produto de convolução

$$g * h(x) = \int g(x - y)h(y) dy.$$

No resultado seguinte obtemos expressões exactas para os termos  $\text{IVAR}(f_n)$  e  $\text{IVIES}(f_n)$ , definidos por (2.2) e (2.3), respectivamente, que utilizaremos posteriormente para provar a convergência em média quadrática integrada de  $f_n$  para  $f$ .

**Teorema 2.1.1.** *Sejam  $K$  e  $f$  de quadrado integrável. Para  $h_n > 0$  temos*

$$\text{IVAR}(f_n) = \frac{1}{nh_n} R(K) - \frac{1}{n} \int K_{h_n} * f(x)^2 dx$$

e

$$\text{IVIES}(f_n) = \int K_{h_n} * f(x)^2 dx - 2 \int K_{h_n} * f(x)f(x) dx + R(f).$$

**Demonstração:** Tendo em conta que o estimador do núcleo tem por média

$$\mathbb{E}(f_n(x)) = \int K_{h_n}(x - y)f(y) dy = K_{h_n} * f(x) \quad (2.4)$$

(ver Wand e Jones, 1995, pg. 14), o viés do núcleo é dado por

$$\text{Viés}(f_n(x)) = K_{h_n} * f(x) - f(x) \quad (2.5)$$

$$= \int K(z) \left( f(x - zh_n) - f(x) \right) dz. \quad (2.6)$$

Relativamente à variância, esta admite o desenvolvimento

$$\begin{aligned} \text{Var}(f_n(x)) &= \frac{1}{n} \text{Var}(K_{h_n}(x - X_1)) \\ &= \frac{1}{n} \left\{ \mathbb{E}(K_{h_n}(x - X_1)^2) - (\mathbb{E}(K_{h_n}(x - X_1)))^2 \right\} \\ &= \frac{1}{nh_n} \int K(z)^2 f(x - zh_n) dz - \frac{1}{n} (\mathbb{E}(f_n(x)))^2. \end{aligned} \quad (2.7)$$



Tendo em conta as expressões (2.2), (2.4) e (2.7), obtém-se a expressão pretendida para o  $\text{IVAR}(f_n)$ . A expressão para  $\text{IVIES}(f_n)$  é consequência imediata de (2.3) e (2.5). ■

Estamos agora em condições de demonstrar a convergência em média quadrática integrada de  $f_n$  para  $f$ . Como veremos, as condições que a janela,  $h_n$ , satisfaz terão um papel importante nesta convergência. A condição  $h_n \rightarrow 0$  implica uma redução do termo de viés, ou seja,  $\text{IVIES}(f_n) \rightarrow 0$ . No entanto, uma tal convergência para zero não pode ocorrer de forma violenta uma vez que é necessário que  $nh_n \rightarrow +\infty$  para que o termo de variância convirja para zero, isto é,  $\text{IVAR}(f_n) \rightarrow 0$ .

**Teorema 2.1.2.** *Sejam  $K$  e  $f$  de quadrado integrável. Se  $h_n \rightarrow 0$  e  $nh_n \rightarrow +\infty$  então  $f_n$  converge para  $f$  em média quadrática integrada, ou seja,*

$$\text{EQMI}(f_n) \rightarrow 0, \text{ quando } n \rightarrow +\infty.$$

**Demonstração:** Começemos por notar que o integral  $\int K_{h_n} * f(x)^2 dx$ , que surge nas expressões do  $\text{IVAR}(f_n)$  e  $\text{IVIES}(f_n)$ , dadas no teorema anterior, pode ser escrito na forma

$$\int \varphi_{h_n} * f(x)f(x)dx$$

com  $\varphi = \bar{K} * K$ , onde  $\bar{g}(u) = g(-u)$ , sendo  $g$  de quadrado integrável. Com efeito,

$$\begin{aligned} \int K_{h_n} * f(x)^2 dx &= \iiint K_{h_n}(x-u)f(u)K_{h_n}(x-z)f(z)dudzdx \\ &= \iiint \bar{K}_{h_n}(u-z-y)K_{h_n}(y)f(u)f(z)dudzdy \\ &= \iint \bar{K}_{h_n} * K_{h_n}(u-z)f(z)f(u)dudz \\ &= \int (\bar{K} * K)_{h_n} * f(u)f(u)du. \end{aligned} \tag{2.8}$$

Provemos agora que, sendo  $\varphi$  uma função integrável, é válida a convergência

$$\int \varphi_{h_n} * f(x)f(x)dx \rightarrow R(f) \int \varphi(z)dz, \tag{2.9}$$

quando  $n \rightarrow +\infty$ . Atendendo a que

$$\begin{aligned} \int \varphi_{h_n} * f(x)f(x)dx &= \iint \varphi_{h_n}(x-y)f(y)f(x)dx dy \\ &= \iint \varphi(z)f(x-zh_n)f(x)dx dz \\ &= \int \varphi(z)\bar{f} * f(zh_n)dz, \end{aligned}$$

usando o Teorema da Convergência Dominada de Lebesgue (ver Fernandez, 1976, pg. 75) e o facto de  $\bar{f} * f$  ser limitada e contínua na origem, pois é o produto de convolução de funções de quadrado integrável, concluimos o desejado.

Usando (2.9), concluimos finalmente que, quando  $n \rightarrow +\infty$ , temos

$$\text{IVAR}(f_n) = \frac{1}{nh_n}R(K) - \frac{1}{n} \left( R(f) \int \bar{K} * K(x)dx + o(1) \right) \rightarrow 0$$

e

$$\text{IVIES}(f_n) \rightarrow R(f) \int \bar{K} * K(x)dx - 2R(f) \int K(x)dx + R(f) = 0.$$

■

## 2.2. Desenvolvimento assintótico do EQMI

Enquanto que a convergência em média quadrática integrada do estimador do núcleo foi obtida sob condições gerais sobre  $f$ , os desenvolvimentos assintóticos do EQMI( $f_n$ ) que apresentamos nesta secção necessitam de condições de regularidade sobre a densidade desconhecida  $f$ .

A fórmula de Taylor com resto integral (ver Lima, 1992, pg. 261–2), que recordamos de seguida, será utilizada nos parágrafos 2.2.1 e 2.2.2.

**Teorema 2.2.1.** *Seja  $g$  uma função real de variável real admitindo derivada até à ordem  $p - 1$  numa vizinhança do ponto  $x \in \mathbb{R}$ . Sendo  $u_n$  uma qualquer sucessão de números reais convergente para zero, se  $g$  admite derivada de ordem  $p$  contínua numa vizinhança do ponto  $x \in \mathbb{R}$ , temos*

$$g(x + u_n) = \sum_{i=0}^{p-1} \frac{u_n^i}{i!} g^{(i)}(x) + R_n(x),$$

onde

$$R_n(x) = u_n^p \int_0^1 \frac{(1-t)^{(p-1)}}{(p-1)!} g^{(p)}(x + tu_n) dt.$$

### 2.2.1. Núcleos de 2ª ordem

Neste parágrafo, iremos obter o desenvolvimento assintótico para o erro quadrático médio integrado no caso em que o núcleo  $K$  é um núcleo de 2.ª ordem. Em consequência de tal desenvolvimento, obtemos também uma expressão para a janela que minimiza o erro quadrático médio integrado assintótico.

**Teorema 2.2.2.** *Sejam  $K$  um núcleo de quadrado integrável de 2.<sup>a</sup> ordem,  $f$  com derivada de segunda ordem contínua e,  $f$  e  $f''$  de quadrado integrável. Se  $h_n \rightarrow 0$  então*

$$\begin{aligned} \text{IVAR}(f_n) &= \frac{1}{nh_n}R(K) + O\left(\frac{1}{n}\right), \\ \text{IVIES}(f_n) &= \frac{h_n^4}{4}m_2^2(K)R(f'') + o(h_n^4) \end{aligned}$$

e

$$\text{EQMI}(f_n) = \frac{1}{nh_n}R(K) + \frac{h_n^4}{4}m_2^2(K)R(f'') + O\left(\frac{1}{n}\right) + o(h_n^4).$$

**Demonstração:** Tendo em conta a demonstração do Teorema 2.1.2, o desenvolvimento para  $\text{IVAR}(f_n)$  é obtido de imediato.

Atendendo à expressão (2.6), pela fórmula de Taylor com resto integral, temos

$$\text{Viés}(f_n(x)) = \int K(z) \left( -zh_n f'(x) + (-zh_n)^2 \int_0^1 (1-t) f''(x - th_n z) dt \right) dz.$$

Como  $K$  é um núcleo de 2.<sup>a</sup> ordem, temos que  $m_1(K) = 0$ , portanto,

$$\text{Viés}(f_n(x)) = h_n^2 \int \int_0^1 z^2 K(z) (1-t) f''(x - th_n z) dt dz$$

e, assim, ficamos com

$$\begin{aligned} \text{IVIES}(f_n) &= h_n^4 \int \int \int \int_0^1 \int_0^1 u^2 K(u) (1-s) f''(x - suh_n) \\ &\quad \times v^2 K(v) (1-t) f''(x - tvh_n) ds dt dudv dx \\ &= h_n^4 \int \int \int_0^1 \int_0^1 u^2 v^2 K(u) K(v) \\ &\quad \times (1-s)(1-t) \bar{f}'' * f''((su - tv)h_n) ds dt dudv. \end{aligned}$$

Usando agora o Teorema da Convergência Dominada de Lebesgue, analogamente ao que foi feito na demonstração do Teorema 2.1.2, obtemos

$$\text{IVIES}(f_n) = \frac{h_n^4}{4}m_2^2(K)R(f'') + o(h_n^4).$$

O desenvolvimento de  $\text{EQMI}(f_n)$  é obtido tendo em conta as expressões anteriores. ■

De seguida, apresentamos uma expressão para a janela assintoticamente óptima, no sentido da minimização do erro quadrático médio integrado assintótico definido como sendo a soma das parcelas mais significativas do  $\text{EQMI}(f_n)$ , dado no teorema anterior,

$$\text{EQMIA}(h_n) = \frac{1}{nh_n}R(K) + \frac{h_n^4}{4}m_2^2(K)R(f'').$$

Para calcular a janela,  $h_n$ , que minimiza a expressão anterior, utilizaremos o seguinte resultado (ver Parzen, 1962):

**Lema 2.2.1.** *Seja  $A, B, \alpha$  e  $\beta$  quantidades positivas. A função*

$$\psi(h) = Ah^\alpha + Bh^{-\beta},$$

com  $h > 0$  atinge um mínimo absoluto no ponto

$$h^* = \left( \frac{\beta B}{\alpha A} \right)^{1/(\alpha+\beta)}.$$

**Teorema 2.2.3.** *Se  $f$  é tal que  $R(f'') > 0$ , o valor de  $h_n$  que minimiza o EQMIA é dado por*

$$h_{\text{EQMIA}} = c_K R(f'')^{-1/5} n^{-1/5},$$

onde

$$c_K = \left( \frac{R(K)}{m_2^2(K)} \right)^{1/5}.$$

## 2.2.2. Núcleos de ordem finita

Neste parágrafo, iremos apresentar os desenvolvimentos do erro quadrático médio integrado assintótico e da janela assintoticamente ótima quando  $K$  é um núcleo de ordem finita  $k > 2$ . Analogamente ao que foi feito no parágrafo anterior (veja-se a demonstração do Teorema 2.2.2), pela fórmula de Taylor com resto integral, ficamos com

$$\text{Viés}(f_n(x)) = (-1)^k h_n^k \int_0^1 z^k K(z) \frac{(1-t)^{k-1}}{(k-1)!} f^{(k)}(x - th_n z) dt dz,$$

o que permite obter o desenvolvimento para o termo de viés,  $\text{IVIES}(f_n)$  e, por consequência, a expressão para o  $\text{EQMI}(f_n)$ , dados no teorema seguinte.

**Teorema 2.2.4.** *Sejam  $K$  um núcleo de quadrado integrável de ordem  $k$ ,  $f$  com derivada de ordem  $p$  contínua em  $\mathbb{R}$  e,  $f$  e  $f^{(p)}$  de quadrado integrável. Se  $h_n \rightarrow 0$  então*

$$\text{IVAR}(f_n) = \frac{1}{nh_n} R(K) + O\left(\frac{1}{n}\right),$$

$$\text{IVIES}(f_n) = \frac{h_n^{2r}}{(r!)^2} m_r^2(K) R(f^{(r)}) + o(h_n^{2r})$$

e

$$\text{EQMI}(f_n) = \frac{1}{nh_n} R(K) + \frac{h_n^{2r}}{(r!)^2} m_r^2(K) R(f^{(r)}) + O\left(\frac{1}{n}\right) + o(h_n^{2r}),$$

onde  $r = \min\{k, p\}$ .

Analogamente ao que foi feito no parágrafo anterior, o desenvolvimento assintótico do erro quadrático médio integrado é dado por

$$\text{EQMIA}(h_n) = \frac{1}{nh_n} R(K) + \frac{h_n^{2r}}{(r!)^2} m_r^2(K) R(f^{(r)}),$$

e a expressão para a janela assintoticamente óptima é dada por:

**Teorema 2.2.5.** *Se  $f$  é tal que  $R(f^{(r)}) > 0$ , o valor  $h_n$  que minimiza o EQMIA é dado por*

$$h_{\text{EQMIA}} = c_K R(f^{(r)})^{-1/(2r+1)} n^{-1/(2r+1)},$$

onde

$$c_K = \left( \frac{(r!)^2 R(K)}{2r m_r^2(K)} \right)^{1/(2r+1)}.$$

### 2.2.3. Supernúcleos

Ao fixarmos a ordem do núcleo, estamos de certa forma a limitar a ordem de convergência do termo de viés do estimador,  $\text{IVIES}(f_n)$ , para densidades que admitem derivadas contínuas para além da ordem  $k$ . De facto, mesmo que a densidade  $f$  possua derivadas contínuas de ordem superior a  $k$ , a ordem do termo mais significativo do  $\text{IVIES}(f_n)$  é determinada pela ordem do núcleo  $K$ .

Ao considerarmos para  $K$  um supernúcleo, e atendendo a que os seus momentos de todas as ordens são nulos, encontramos

$$\text{IVIES}(f_n) = o(h_n^{2p}),$$

onde  $f$  admite derivada de ordem  $p$  contínua em  $\mathbb{R}$  e  $f^{(p)}$  é de quadrado integrável. Assim, a ordem de convergência do termo de viés,  $\text{IVIES}(f_n)$ , depende apenas da regularidade de  $f$ .

No que se segue, apresentamos expressões alternativas para o  $\text{IVAR}(f_n)$  e para o  $\text{IVIES}(f_n)$  (ver Chácon et al., 2007a) em termo das transformadas de Fourier da densidade  $f$  e do núcleo  $K$ .

**Teorema 2.2.6.** *Sejam  $K$  e  $f$  de quadrado integrável. Para  $h_n > 0$  temos*

$$\text{IVAR}(f_n) = \frac{1}{2\pi} \left( \frac{1}{nh_n} \int |\varphi_K(t)|^2 dt - \frac{1}{n} \int |\varphi_K(th_n)|^2 |\varphi_f(t)|^2 dt \right)$$

e

$$\text{IVIES}(f_n) = \frac{1}{2\pi} \int |\varphi_f(t)|^2 |\varphi_K(th_n) - 1|^2 dt.$$

**Demonstração:** Tendo em conta a expressão do  $\text{IVAR}(f_n)$  presente no Teorema 2.1.1 e, utilizando o Teorema de Parseval (ver Malliavin, 1982, pg. 118) e as propriedades da transformada de Fourier, obtemos

$$\begin{aligned} \text{IVAR}(f_n) &= \frac{1}{nh_n} \left( \frac{1}{2\pi} \int |\varphi_K(t)|^2 dt \right) - \frac{1}{n} \left( \frac{1}{2\pi} \int |\varphi_{K_{h_n} * f}(t)|^2 dt \right) \\ &= \frac{1}{2\pi} \frac{1}{nh_n} \int |\varphi_K(t)|^2 dt - \frac{1}{2\pi} \frac{1}{n} \int |\varphi_{K_{h_n}}(t) \varphi_f(t)|^2 dt \\ &= \frac{1}{2\pi} \left( \frac{1}{nh_n} \int |\varphi_K(t)|^2 dt - \frac{1}{n} \int |\varphi_K(th_n)|^2 |\varphi_f(t)|^2 dt \right). \end{aligned}$$

Tendo em conta a expressão (2.5), o Teorema de Parseval (ver Malliavin, 1982, pg. 118) e as propriedades da transformada de Fourier, obtemos

$$\begin{aligned} \text{IVIES}(f_n) &= \int \{K_{h_n} * f(x) - f(x)\}^2 dx \\ &= \frac{1}{2\pi} \int |\varphi_{K_{h_n} * f - f}(t)|^2 dt \\ &= \frac{1}{2\pi} \int |\varphi_{K_{h_n}}(t) \varphi_f(t) - \varphi_f(t)|^2 dt \\ &= \frac{1}{2\pi} \int |\varphi_f(t)|^2 |\varphi_K(th_n) - 1|^2 dt. \end{aligned}$$

■

Tendo em conta o teorema anterior, no que se segue, apresentamos um majorante para o  $\text{EQMI}(f_n)$ , a partir do qual obtemos a expressão da janela assintoticamente óptima, no sentido da minimização desse majorante.

**Teorema 2.2.7.** *Sejam  $K$  um supernúcleo e  $f$  uma densidade, ambos de quadrado integrável. Se  $f$  admite derivada de ordem  $k$ , integrável e de quadrado integrável então*

$$\text{EQMI}(f_n) \leq \frac{1}{nh_n} R(K) + \frac{h_n^{2k}}{S_K^{2k}} R(f^{(k)}),$$

onde  $S_K = \inf \{t \geq 0 : |\varphi_K(t) - 1| \neq 0\}$ .

**Demonstração:** Relativamente ao termo de viés, temos

$$\begin{aligned} \text{IVIES}(f_n) &= \frac{1}{2\pi} \int |\varphi_f(t)|^2 |\varphi_K(th_n) - 1|^2 dt \\ &= \frac{1}{2\pi} \int_{|t| > \frac{S_K}{h_n}} |\varphi_f(t)|^2 |\varphi_K(th_n) - 1|^2 dt \\ &\leq \frac{1}{2\pi} \int_{|t| > \frac{S_K}{h_n}} |\varphi_f(t)|^2 dt \\ &\leq \frac{h_n^{2k}}{S_K^{2k}} \frac{1}{2\pi} \int |t|^{2k} |\varphi_f(t)|^2 dt. \end{aligned}$$

Atendendo a que,

$$\int |t|^{2k} |\varphi_f(t)|^2 dt = 2\pi R(f^{(k)}) < \infty,$$

(ver Donoghue, 1969, pg. 138) obtemos,

$$\text{IVIES}(f_n) \leq \frac{h_n^{2k}}{S_K^{2k}} R(f^{(k)}).$$

O resultado é agora uma consequência imediata da desigualdade

$$\text{IVAR}(f_n) \leq \frac{1}{nh_n} R(K)$$

que obtemos do teorema anterior. ■

Notemos que a majoração apresentada para o termo de variância, pode ser igualmente obtida a partir do Teorema 2.1.1.

Tendo em conta o teorema anterior, estamos agora em condições de obter a expressão para a janela assintoticamente óptima, no sentido da minimização desse majorante.

**Teorema 2.2.8.** *Se  $f$  é tal que  $R(f^{(k)}) > 0$ , o valor  $h_n$  que minimiza o majorante de EQMI, apresentado no Teorema 2.2.7 é dado por*

$$h_n = c_K R(f^{(k)})^{-1/(2k+1)} n^{-1/(2k+1)},$$

onde

$$c_K = \left( \frac{R(K) S_K^{2k}}{2k} \right)^{1/(2k+1)}.$$





# Capítulo 3

## A correcção de Glad-Hjort-Ushakov

No caso em que o núcleo  $K$  é de ordem superior à segunda ou é um supernúcleo, o estimador  $f_n$  baseado nesses núcleos não é mais uma densidade de probabilidade, visto poder tomar valores negativos, embora satisfaça a condição

$$\int f_n(x) dx = 1.$$

Neste capítulo, iremos apresentar um método proposto por Glad, Hjort e Ushakov (2003), que permite tornar próprio um estimador  $\hat{f}_n$  da densidade, que satisfaça a condição

$$\int \max\{0, \hat{f}_n(x)\} dx \geq 1, \text{ quase certamente,} \quad (3.1)$$

com erro quadrático integrado,

$$\text{EQI}(\hat{f}_n) = \int \{\hat{f}_n(x) - f(x)\}^2 dx, \quad (3.2)$$

inferior ao erro quadrático integrado do estimador inicial  $f_n$ . Os estimadores baseados nos núcleos de ordem superior à segunda ou nos supernúcleos satisfazem a condição (3.1).

O método proposto por Glad, Hjort e Ushakov (2003) consiste em alterar verticalmente o estimador inicial de forma a que o integral da parte positiva do estimador verticalmente modificado seja igual à unidade, tomando-se para estimador corrigido essa parte positiva. Ou seja, o estimador corrigido será definido por

$$\tilde{f}_n(x) = \max\{0, \hat{f}_n(x) - \gamma\}, \quad (3.3)$$

onde  $\gamma$  é escolhido de tal forma que

$$\int \tilde{f}_n(x) dx = 1. \quad (3.4)$$

### 3.1. O estimador $\tilde{f}_n$ está bem definido

Nesta secção, iremos mostrar que o estimador corrigido  $\tilde{f}_n$ , definido por (3.3), está bem definido, ou seja, que existe um único valor de  $\gamma$  que satisfaz a condição (3.4).

Para tal, começamos por apresentar alguns resultados auxiliares. No que se segue, denotaremos por  $m$  a medida de Lebesgue em  $\mathbb{R}$ .

**Lema 3.1.1** (Desigualdade de Chebychev). *Sejam  $f : \mathbb{R} \rightarrow \mathbb{R}^+$  mensurável,  $p > 0$  e  $\varepsilon > 0$ . Então,*

$$m(\{x \in \mathbb{R} : f(x) \geq \varepsilon\}) \leq \varepsilon^{-p} \int f^p(x) dx.$$

Dada uma função  $q$  real de variável real e um número  $\varepsilon > 0$ , consideremos a função  $q_\varepsilon$  definida por

$$q_\varepsilon(x) = \max\{0, q(x) - \varepsilon\},$$

e o conjunto

$$Q_\varepsilon = \{x \in \mathbb{R} : q_\varepsilon(x) \geq 0\} = \{x \in \mathbb{R} : q(x) \geq \varepsilon\}.$$

**Lema 3.1.2.** *Se  $q$  é limitada de quadrado integrável então,  $m(Q_\varepsilon) < \infty$  e a função  $q_\varepsilon$  é integrável para qualquer  $\varepsilon > 0$ .*

**Demonstração:** Para  $\varepsilon > 0$ , fixo, pela desigualdade de Chebychev temos, para  $p = 2$ ,

$$\int q^2(x) dx \geq \int_{Q_\varepsilon} q^2(x) dx \geq \varepsilon^2 m(Q_\varepsilon).$$

Atendendo a que  $q$  é de quadrado integrável, concluímos que  $Q_\varepsilon$  tem medida de Lebesgue finita.

Sabendo que  $q_\varepsilon(x) = 0$ , sempre que  $x \notin Q_\varepsilon$ , temos

$$\int q_\varepsilon(x) dx = \int_{Q_\varepsilon} q_\varepsilon(x) dx \leq \int_{Q_\varepsilon} \sup_{x \in \mathbb{R}} q_\varepsilon(x) dx \leq \sup_{x \in \mathbb{R}} |q(x)| m(Q_\varepsilon),$$

o que permite concluir que  $q_\varepsilon$  é integrável. ■

**Lema 3.1.3.** *Se  $q$  é limitada de quadrado integrável então, a função  $\lambda$  definida por  $\lambda(\varepsilon) = \int q_\varepsilon(x) dx$ , com  $\varepsilon > 0$ , é contínua, decrescente e*

$$\lim_{\varepsilon \rightarrow 0} \lambda(\varepsilon) = \int \max\{0, q(x)\} dx.$$

**Demonstração:** Para  $0 < \varepsilon_2 < \varepsilon_1$ , temos que

$$(\varepsilon_1 - \varepsilon_2)m(Q_{\varepsilon_1}) \leq \lambda(\varepsilon_2) - \lambda(\varepsilon_1) \leq (\varepsilon_1 - \varepsilon_2)m(Q_{\varepsilon_2}) \quad (3.5)$$

pois, como  $Q_{\varepsilon_1} \subset Q_{\varepsilon_2}$ ,

$$\begin{aligned} \lambda(\varepsilon_2) - \lambda(\varepsilon_1) &= \int_{Q_{\varepsilon_2}} \max\{0, q(x) - \varepsilon_2\} dx - \int_{Q_{\varepsilon_1}} \max\{0, q(x) - \varepsilon_1\} dx \\ &\leq \int_{Q_{\varepsilon_2}} \left( \max\{0, q(x) - \varepsilon_2\} - \max\{0, q(x) - \varepsilon_1\} \right) dx \\ &\leq \int_{Q_{\varepsilon_2}} \left( (q(x) - \varepsilon_2) - (q(x) - \varepsilon_1) \right) dx \\ &= \int_{Q_{\varepsilon_2}} (\varepsilon_1 - \varepsilon_2) dx \\ &= (\varepsilon_1 - \varepsilon_2) m(Q_{\varepsilon_2}), \end{aligned}$$

e também

$$\begin{aligned} \lambda(\varepsilon_2) - \lambda(\varepsilon_1) &\geq \int_{A_{\varepsilon_1}} \left( (q(x) - \varepsilon_2) - (q(x) - \varepsilon_1) \right) dx \\ &= \int_{A_{\varepsilon_1}} (\varepsilon_1 - \varepsilon_2) dx \\ &= (\varepsilon_1 - \varepsilon_2) m(A_{\varepsilon_1}). \end{aligned}$$

Tendo em conta (3.5) e a continuidade da medida  $m$ , concluímos que a função  $\lambda$  é contínua para  $\varepsilon > 0$ .

Para  $\varepsilon > 0$ , temos  $q_\varepsilon(x) \geq 0$  e, quando  $\varepsilon \rightarrow 0$ , temos

$$q_\varepsilon(x) \uparrow \max\{0, q(x)\},$$

o que, pelo Teorema da Convergência Monótona (ver Fernandez, 1976, pg. 69), permite concluir que

$$\lambda(\varepsilon) = \int q_\varepsilon(x) dx \uparrow \int \max\{0, q(x)\} dx.$$

■

**Observação 3.1.1.** Reparemos que, sempre que

$$m(Q_\varepsilon) > 0, \forall \varepsilon \in ]0, \sup_{x \in \mathbb{R}} q(x)[$$

a função  $\lambda$  é estritamente decrescente no intervalo anterior. Esta condição é, em particular, satisfeita se  $0 < q(x_0) = \sup_{x \in \mathbb{R}} q(x)$  para algum  $x_0 \in \mathbb{R}$  e  $q$  é contínua em  $x_0$ .

Estamos agora em condições de mostrar que existe um único valor de  $\gamma$  que satisfaz a condição (3.4).

**Teorema 3.1.1.** *Se  $\hat{f}_n(\cdot) = \hat{f}_n(X_1, X_2, \dots, X_n; \cdot)$  é, para quase toda a amostra, um estimador de  $f$  de quadrado integrável e limitado satisfazendo (3.1) e tal que*

$$m(\{x \in \mathbb{R} : \hat{f}_n(x) \geq \varepsilon\}) > 0, \text{ para todo } \varepsilon \in ]0, \sup_{x \in \mathbb{R}} \hat{f}_n(x)[, \quad (3.6)$$

*então, para quase toda a amostra, existe um e um só valor de  $\gamma \geq 0$  tal que o estimador  $\tilde{f}_n$ , definido por (3.3), satisfaz (3.4).*

**Demonstração:** Consideremos uma realização arbitrária de  $\hat{f}_n$ . Atendendo às hipóteses do teorema vamos admitir que esta realização é de quadrado integrável e que satisfaz  $\int \max\{0, \hat{f}_n(x)\} dx \geq 1$ .

No caso em que

$$\int \max\{0, \hat{f}_n(x)\} dx = 1,$$

para  $\gamma = 0$  temos (3.4). Este é o único valor que satisfaz (3.4) uma vez que, pelo Lema 3.1.3 e por (3.6), a função  $\lambda(\varepsilon) = \int \max\{0, \hat{f}_n(x) - \varepsilon\} dx$  é estritamente decrescente e satisfaz  $\lim_{\varepsilon \rightarrow 0} \lambda(\varepsilon) = 1$ .

No caso em que

$$\int \max\{0, \hat{f}_n(x)\} dx > 1,$$

usando a continuidade da função  $\lambda$ , definida por  $\lambda(\varepsilon) = \int \max\{0, \hat{f}_n(x) - \varepsilon\} dx$ , o facto desta função ser estritamente decrescente satisfazendo  $\lim_{\varepsilon \rightarrow 0} \lambda(\varepsilon) = \int \max\{0, \hat{f}_n(x)\} dx > 1$  e  $\lambda(b) = 0$ , quando  $b = \sup_{x \in \mathbb{R}} \hat{f}_n(x)$ , concluímos o pretendido. ■

### 3.2. Erro quadrático integrado de $\tilde{f}_n$

Vamos mostrar agora que, no sentido do erro quadrático integrado, definido por (3.2), o estimador  $\tilde{f}_n$ , definido por (3.3), é pelo menos tão bom quanto o estimador inicial  $\hat{f}_n$ .

Notemos que se  $\hat{f}_n$  é um estimador baseado num núcleo de ordem superior à segunda ou num supernúcleo tal que  $\lim_{|x| \rightarrow +\infty} K(x) = 0$  e  $K$  contínuo, as condições do teorema seguinte são válidas para um tal estimador.

**Teorema 3.2.1.** *Seja  $\hat{f}_n$  um estimador de  $f$  nas condições do Teorema 3.1.1 tal que*

*$\lim_{|x| \rightarrow +\infty} \hat{f}_n(x) = 0$ . Se  $f$  é de quadrado integrável então, para todo o  $n$ , temos que*

$$\text{EQI}(\tilde{f}_n) \leq \text{EQI}(\hat{f}_n).$$

**Demonstração:** No caso em que  $\gamma = 0$  vimos já que  $\tilde{f}_n(x) = \max\{0, \hat{f}_n(x)\}$  e, o resultado enunciado decorre da desigualdade

$$|\tilde{f}_n(x) - f(x)| \leq |\hat{f}_n(x) - f(x)|$$

que é válida para todo o  $x \in \mathbb{R}$ .

No caso em que  $\gamma > 0$ , isto significa, em particular, que

$$\int \max\{0, \hat{f}_n(x)\} dx > 1.$$

Fixemos  $0 < \varepsilon < 1$  arbitrariamente. Seja  $A_M = [-\frac{M}{2}, \frac{M}{2}]$  um conjunto, dependente de  $\varepsilon$ , com medida de Lebesgue  $M > 0$ , suficientemente grande, tal que

$$\int_{\mathbb{R} \setminus A_M} f(x)^2 dx < \varepsilon, \tag{3.7}$$

$$\int_{A_M} f(x) dx > 0, \tag{3.8}$$

$$\int_{A_M} \max\{0, \hat{f}_n(x)\} dx > 1, \tag{3.9}$$

$$\sup_{x \in \mathbb{R} \setminus A_M} \hat{f}_n(x) < \gamma. \tag{3.10}$$

De seguida, consideramos uma sucessão de funções  $q_1, q_2, \dots$  tal que: seja  $q_0 = \hat{f}_n$  e

$$q_k(x) = \max\{0, q_{k-1}(x)\} - c_k, \text{ para } k = 1, 2, \dots, \tag{3.11}$$

onde

$$c_k = \frac{1}{M} \left[ \int_{A_M} \max\{0, q_{k-1}(x)\} dx - \int_{A_M} f(x) dx \right],$$

ou seja,  $c_k$  é tal que

$$\int_{A_M} q_k(x) dx = \int_{A_M} f(x) dx. \tag{3.12}$$

Note-se que, atendendo a (3.9),  $c_k \geq 0$ , para todo o  $k$ .

Antes de continuarmos com a demonstração do teorema, vejamos dois resultados auxiliares, relativamente à sucessão de funções  $(q_k)$  definida por (3.11).

**Lema 3.2.1.** *Para todo  $k = 1, 2, \dots$ , temos*

$$\int_{A_M} (q_k(x) - f(x))^2 dx \leq \int_{A_M} (q_{k-1}(x) - f(x))^2 dx. \quad (3.13)$$

**Demonstração:** Sabendo que, para qualquer  $a \in \mathbb{R}$  e  $b \geq 0$ , temos  $|\max(0, a) - b| \leq |a - b|$ , então

$$\int_{A_M} (\max\{0, q_{k-1}(x)\} - f(x))^2 dx \leq \int_{A_M} (q_{k-1}(x) - f(x))^2 dx. \quad (3.14)$$

Consideremos a seguinte função quadrática definida, para  $z \in \mathbb{R}$ , por

$$\lambda_k(z) = \int_{A_M} (\max\{0, q_{k-1}(x)\} - z - f(x))^2 dx.$$

Pelo Teorema da derivação sob o sinal de integral, temos que

$$\lambda'_k(z) = -2 \int_{A_M} (\max\{0, q_{k-1}(x)\} - z - f(x)) dx$$

e, é igual a zero, para  $z = c_k$ . Assim,  $\lambda_k$  atinge o mínimo absoluto no ponto  $z = c_k$  e, portanto,

$$\int_{A_M} (q_k(x) - f(x))^2 dx \leq \int_{A_M} (\max\{0, q_{k-1}(x)\} - f(x))^2 dx. \quad (3.15)$$

Portanto, combinando (3.14) e (3.15) concluímos (3.13). ■

**Lema 3.2.2.** *Para  $k = 1, 2, \dots$ , temos*

$$q_k(x) + c_k = \max \left\{ 0, q_0(x) - \sum_{j=1}^{k-1} c_j \right\}. \quad (3.16)$$

**Demonstração:** Este resultado pode ser provado por indução. Para  $k = 1$ , a igualdade é verdadeira por definição de  $q_1(x)$ . Supondo agora que a igualdade é válida para  $k$ , provemos que é também válida para o seu sucessor. Para tal, vamos usar a igualdade  $\max(0, \max(0, a) - b) = \max(0, a - b)$  que é válida para qualquer  $a \in \mathbb{R}$  e  $b \geq 0$ .

Assim, temos

$$\begin{aligned}
 q_{k+1}(x) + c_{k+1} &= \max\{0, q_k(x)\} \\
 &= \max\{c_k, q_k(x) + c_k\} - c_k \\
 &= \max\left\{c_k, \max\left\{0, q_0(x) - \sum_{j=1}^{k-1} c_j\right\}\right\} - c_k \\
 &= \max\left\{0, \max\left\{0, q_0(x) - \sum_{j=1}^{k-1} c_j\right\} - c_k\right\} \\
 &= \max\left\{0, q_0(x) - \sum_{j=1}^k c_j\right\}.
 \end{aligned}$$

■

Retomando agora a demonstração do teorema, do Lema 3.2.2 concluímos que  $\sum_{j=1}^{\infty} c_j < \infty$ , pois, caso contrário, teríamos

$$\limsup_{k \rightarrow \infty} \int_{A_M} q_k(x) dx \leq \limsup_{k \rightarrow \infty} \int_{A_M} \max\left\{0, q_0(x) - \sum_{j=1}^{k-1} c_j\right\} dx = 0,$$

o que contraria (3.8) e (3.12).

Tomando agora limites em (3.16), quando  $k \rightarrow \infty$ , obtemos

$$\lim_{k \rightarrow \infty} q_k(x) = q^M(x),$$

com

$$q^M(x) = \max\left\{0, q_0(x) - \sum_{j=1}^{\infty} c_j\right\},$$

onde este limite satisfaz

$$\limsup_{k \rightarrow \infty} \sup_{x \in \mathbb{R}} |q_k(x) - q^M(x)| = 0 \tag{3.17}$$

e,

$$\lim_{k \rightarrow \infty} \int \{q_k(x) - q^M(x)\}^2 dx = 0. \tag{3.18}$$

Considerando agora o estimador  $\tilde{f}_n$  definido por (3.3), de (3.10) concluímos que

$$\tilde{f}_n(x) = 0, \quad \text{para } x \in \mathbb{R} \setminus A_M, \tag{3.19}$$

e, portanto, pela definição de  $\gamma$ , temos que

$$\int_{A_M} \tilde{f}_n(x) dx = \int \tilde{f}_n(x) dx = 1. \tag{3.20}$$

Tendo em conta (3.12) e o facto de

$$\begin{aligned} \left| \int_{A_M} q_k(x) dx - \int_{A_M} q^M(x) dx \right| &\leq \int_{A_M} |q_k(x) - q^M(x)| dx \\ &\leq m(A_M) \sup_{x \in A_M} |q_k(x) - q^M(x)| \rightarrow 0, \end{aligned}$$

concluimos que

$$\int_{A_M} q^M(x) dx = \int_{A_M} f(x) dx \leq 1. \quad (3.21)$$

Por (3.20) e (3.21), concluimos que

$$\int_{A_M} q^M(x) dx \leq \int_{A_M} \tilde{f}(x) dx$$

e, portanto,  $\gamma \leq \sum_{j=1}^{\infty} c_j$ . Tendo em conta esta desigualdade, concluimos que

$$q^M(x) \leq \tilde{f}_n(x), \quad \text{para todo o } x \in \mathbb{R}. \quad (3.22)$$

Atendendo a (3.19) e (3.22), temos

$$q^M(x) = 0, \quad \text{para } x \in \mathbb{R} \setminus A_M. \quad (3.23)$$

Por (3.13) e, como

$$\int_{A_M} (q_k(x) - f(x))^2 dx \longrightarrow \int_{A_M} (q^M(x) - f(x))^2 dx$$

concluimos que

$$\begin{aligned} \int_{A_M} (q^M(x) - f(x))^2 dx &\leq \int_{A_M} (q_0(x) - f(x))^2 dx \\ &= \int_{A_M} (\hat{f}_n(x) - f(x))^2 dx. \end{aligned} \quad (3.24)$$

Tendo em consideração (3.7), (3.23) e (3.24), temos que

$$\begin{aligned} \int (q^M(x) - f(x))^2 dx &= \int_{A_M} (q^M(x) - f(x))^2 dx + \int_{\mathbb{R} \setminus A_M} f(x)^2 dx \\ &\leq \int_{A_M} (\hat{f}_n(x) - f(x))^2 dx + \varepsilon \\ &\leq \int (\hat{f}_n(x) - f(x))^2 dx + \varepsilon. \end{aligned} \quad (3.25)$$

Por (3.21), concluimos que

$$\lim_{M \rightarrow \infty} \int_{A_M} q^M(x) dx = \lim_{M \rightarrow \infty} \int q^M(x) dx = 1. \quad (3.26)$$



Atendendo às condições (3.19), (3.22), (3.23), obtemos

$$\lim_{M \rightarrow \infty} \sup \int \left( \tilde{f}_n(x) - q^M(x) \right)^2 dx \leq \lim_{M \rightarrow \infty} \sup \int \left( \tilde{f}_n(x) - q^M(x) \right) \left( \tilde{f}_n(x) + q^M(x) \right) dx.$$

Tomando  $c = \sup_{x \in \mathbb{R}} \hat{f}_n(x)$  (que é limitado pelas condições do teorema) e, por (3.22), temos

$$\tilde{f}_n(x) + q^M(x) \leq 2\tilde{f}_n(x) \leq 2 \sup_{x \in \mathbb{R}} \tilde{f}_n(x) \leq 2 \sup_{x \in \mathbb{R}} \hat{f}_n(x)$$

e, assim, ficamos com

$$\lim_{M \rightarrow \infty} \sup \int \left( \tilde{f}_n(x) - q^M(x) \right)^2 dx \leq 2c \lim_{M \rightarrow \infty} \sup \left( \int \tilde{f}_n(x) dx - \int q^M(x) dx \right).$$

Portanto, por (3.20) e (3.26) concluimos que

$$\lim_{M \rightarrow \infty} \sup \int \left( \tilde{f}_n(x) - q^M(x) \right)^2 dx = 0.$$

Atendendo a (3.25), concluimos que

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} \int \left( q^M(x) - f(x) \right)^2 dx &\leq \lim_{\varepsilon \rightarrow 0} \left\{ \int \left( \hat{f}_n(x) - f(x) \right)^2 dx + \varepsilon \right\} \\ \iff \lim_{\varepsilon \rightarrow 0} \int \left( q^M(x) - f(x) \right)^2 dx &\leq \int \left( \hat{f}_n(x) - f(x) \right)^2 dx. \end{aligned}$$

Tendo em conta que

$$\left[ \int \left( \tilde{f}_n(x) - f(x) \right)^2 dx \right]^{1/2} \leq \left[ \int \left( \tilde{f}_n(x) - q^M(x) \right)^2 dx \right]^{1/2} + \left[ \int \left( q^M(x) - f(x) \right)^2 dx \right]^{1/2}$$

tomando agora limites, quando  $\varepsilon \rightarrow 0$  e  $M \rightarrow \infty$ , ficamos

$$\left[ \int \left( \tilde{f}_n(x) - f(x) \right)^2 dx \right]^{1/2} \leq 0 + \left[ \int \left( \hat{f}_n(x) - f(x) \right)^2 dx \right]^{1/2},$$

ou seja,  $\text{EQI}(\tilde{f}_n) \leq \text{EQI}(\hat{f}_n)$ . ■

### 3.3. Outra correcção

Além da correcção apresentada, uma alternativa possível seria considerar o novo estimador

$$\tilde{f}_n(x) = \gamma \max\{0, \hat{f}_n(x)\},$$

onde  $\gamma$  é escolhido de tal forma que

$$\int \tilde{f}_n(x) dx = 1.$$

Apesar de não conhecermos qualquer resultado teórico que suporte a utilização deste estimador corrigido, no estudo de simulação que apresentamos no capítulo seguinte, este estimador será aí considerado.



## Capítulo 4

# Estudo comparativo: Núcleo de 2.<sup>a</sup> ordem vs. Supernúcleo

Neste capítulo apresentamos um estudo de simulação que tem por objectivo comparar o estimador baseado num núcleo de 2.<sup>a</sup> ordem com o estimador baseado num supernúcleo. Para tal, iremos considerar os núcleos mais utilizados na prática, que são o núcleo normal standard, como núcleo de 2.<sup>a</sup> ordem, e o supernúcleo trapezoidal, como supernúcleo.

Este estudo foi desenvolvido recorrendo ao *Software R* (R Core Team, 2012), tendo sido utilizadas as *packages* KernSmooth (Matt Wand, 2012) e nor1mix (Martin Mächler, 2013). Os gráficos apresentados ao longo do texto foram também obtidos recorrendo a este *Software*.

### 4.1. Comparação baseada no EQMI

Nesta secção, iremos calcular numericamente os valores das janelas óptimas e assintoticamente óptimas e comparar os erros quadráticos médios integrados, associados a estas janelas, dos estimadores baseados no núcleo normal standard e no supernúcleo trapezoidal.

#### 4.1.1. Desenvolvimento exacto do EQMI

Apresentamos nesta secção expressões exactas para os termos de viés e variância,  $\text{IVAR}(f_n)$  e  $\text{IVIES}(f_n)$ , respectivamente, para o estimador baseado no núcleo normal standard e para o estimador baseado no supernúcleo trapezoidal.

##### Núcleo Normal

Considerando  $K$  o núcleo normal standard, Fryer (1976) e Deheuvels (1977) observaram que sendo  $f$  uma mistura de densidades normais, podem ser obtidas expressões simples para as quantidades  $\text{IVAR}(f_n)$  e  $\text{IVIES}(f_n)$ .

De facto, se  $f$  tomar a forma

$$f(x) = \sum_{i=1}^k w_i \phi_{\sigma_i}(x - \mu_i)$$

onde  $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ ,  $\phi_{\sigma}(x) = \frac{1}{\sigma} \phi\left(\frac{x}{\sigma}\right)$ ,  $w_1, \dots, w_k$  números reais positivos cuja soma é unitária,  $\mu_i \in \mathbb{R}$  e  $\sigma_i > 0$ , dizemos que  $f$  é uma mistura de densidades normais com parâmetros  $(w_i, \mu_i, \sigma_i^2)$ , para  $i = 1, \dots, k$ , e podemos então escrever

$$\text{IVAR}(f_n) = \frac{1}{2\sqrt{\pi}} \frac{1}{nh_n} - \frac{1}{n} U(h_n, 2)$$

e

$$\text{IVIES}(f_n) = U(0, 0) - 2U(h_n, 1) + U(h_n, 2),$$

onde

$$U(h, q) = \sum_{i=1}^k \sum_{i'=1}^k w_i w_{i'} \phi_{\sigma_{ii'q}}(\mu_i - \mu_{i'})$$

com  $\sigma_{ii'q} = (\sigma_i^2 + \sigma_{i'}^2 + qh^2)^{1/2}$ .

### Supernúcleo Trapezoidal

Considerando  $K$  o supernúcleo trapezoidal, os desenvolvimentos apresentados de seguida são obtidos tendo em conta as expressões apresentadas na Secção 2.2.3 (ver Chacón et al., 2007a).

Assim, temos

$$\begin{aligned} \text{IVAR}(f_n) &= \frac{1}{2\pi} \left( \frac{2}{nh} \int_0^2 |\varphi_K(t)|^2 - \frac{2}{n} \int_0^{\frac{2}{h}} |\varphi_f(t)|^2 |\varphi_K(th)|^2 dt \right) \\ &= \frac{1}{2\pi} \left( \frac{2}{nh} \frac{4}{3} - \frac{2}{n} \int_0^{\frac{1}{h}} \int |\varphi_f(t)|^2 dt - \frac{2}{n} \int_{\frac{1}{h}}^{\frac{2}{h}} |\varphi_f(t)|^2 |2 - ht|^2 dt \right) \end{aligned}$$

e

$$\begin{aligned} \text{IVIES}(f_n) &= \frac{1}{2\pi} \left( 2 \int_{\frac{1}{h}}^{+\infty} |\varphi_f(t)|^2 |\varphi_K(th) - 1|^2 dt \right) \\ &= \frac{1}{2\pi} \left( 2 \int_{\frac{1}{h}}^{\frac{2}{h}} |\varphi_f(t)|^2 |1 - th|^2 dt + 2 \int_{\frac{2}{h}}^{+\infty} |\varphi_f(t)|^2 dt \right), \end{aligned}$$

onde

$$|\varphi_f(t)|^2 = \sum_{i,j=1}^k \omega_i \omega_j \cos((\mu_i - \mu_j)t) e^{-\frac{1}{2}(\sigma_i^2 + \sigma_j^2)t^2}.$$

## 4.1.2. Densidades de teste

Neste estudo, iremos considerar para densidades de teste,  $f$ , um conjunto de 9 misturas de densidades normais, representadas na Figura 4.1. Estas densidades que apresentam diferentes características distribucionais, foram escolhidas de entre o conjunto de 15 densidades consideradas em Marron e Wand (1992). A numeração que utilizamos para as diferentes densidades é a considerada por Marron e Wand (1992).

A primeira densidade escolhida é a densidade Gaussiana. De seguida, temos as densidades #2 e #3 que apresentam, respectivamente, uma ligeira assimetria e uma forte assimetria. As densidades #6, #7 e #8 são bimodais e as densidades #9, #12 e #14 são multimodais, tanto simétricas como assimétricas.

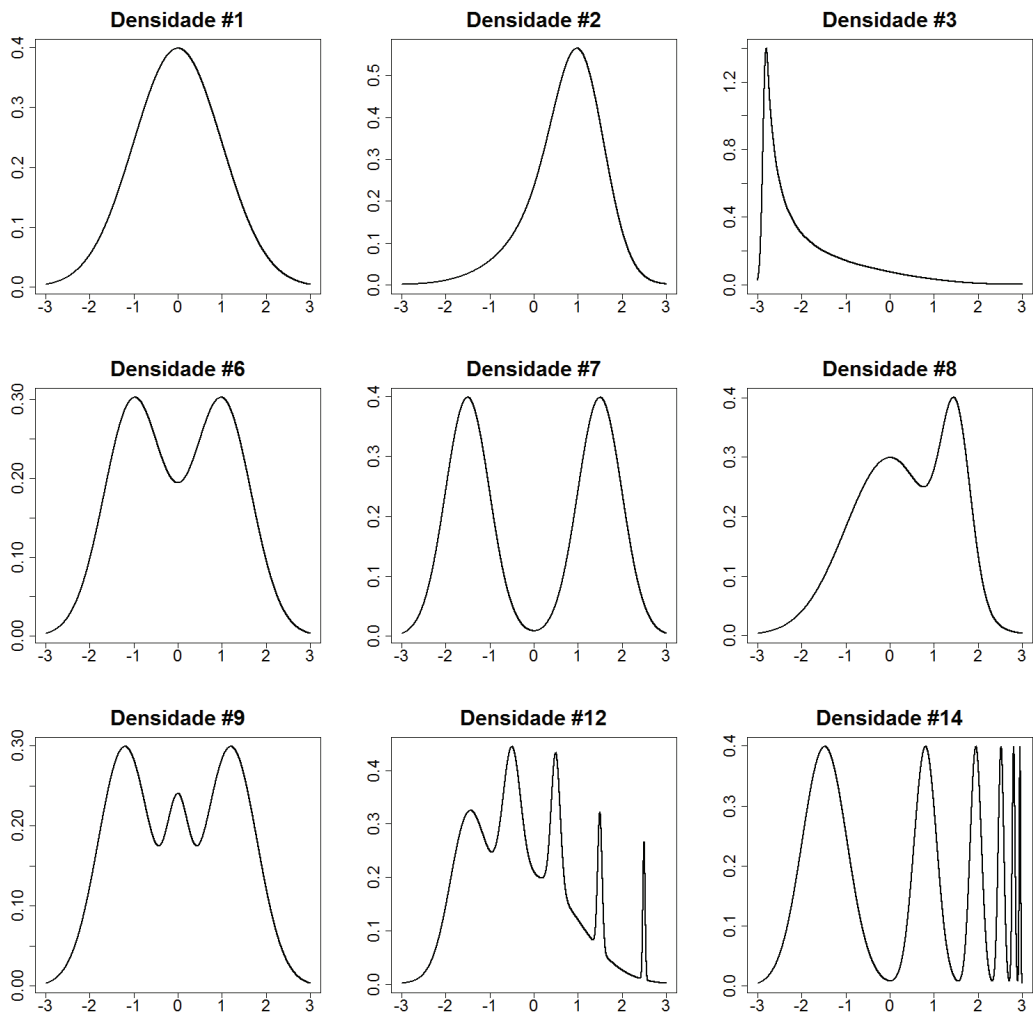


Figura 4.1: Exemplos de mistura de densidades normais

### 4.1.3. Resultados numéricos

Neste parágrafo, vamos comparar os EQMI associados às janelas óptimas, no sentido da minimização do EQMI, isto é,  $h_n = \arg \min_{h>0} \text{EQMI}$ , e associados às janelas assintoticamente óptimas, dos estimadores baseados no núcleo normal e no supernúcleo trapezoidal. O cálculo da janela óptima, tendo em conta as expressões exactas apresentadas anteriormente, é feito numericamente e, as janelas assintoticamente óptimas são calculadas tendo em conta as expressões apresentadas na Secção 2.2. No caso do estimador baseado no supernúcleo trapezoidal a janela assintoticamente óptima, no sentido da minimização dum majorante do EQMI, é obtida tomando  $k = 2$  no Teorema 2.2.8.

Representamos as janelas óptimas do estimador baseado no núcleo normal e no supernúcleo trapezoidal, respectivamente, por  $h_N$  e  $h_T$  e, as janelas assintoticamente óptimas são representadas, respectivamente, por  $h_{0N}$  e  $h_{0T}$ .

Para cada densidade, os resultados são obtidos para amostras de tamanhos,  $n$ , iguais a 50, 100, 200, 300, 400 e 500 e estão representados na Figura 4.2.

Analisando os resultados apresentados, verificamos que no caso do estimador baseado no núcleo normal, para as densidades com características de multimodalidade, densidades #12 e #14, a janela assintoticamente óptima,  $h_{0N}$ , não é uma boa aproximação da janela óptima, enquanto que nas outras densidades verificamos que é uma boa aproximação. No caso do estimador baseado no supernúcleo trapezoidal verificamos que, para todas as densidades, a janela que minimiza um majorante do EQMI,  $h_{0T}$ , não é uma boa aproximação da janela óptima.

É também possível verificar que o EQMI do estimador baseado no supernúcleo trapezoidal com janela óptima é algumas vezes inferior ao EQMI do estimador baseado no núcleo normal com janela óptima. Assim, pelo menos quando a janela que utilizamos é a janela óptima, embora desconhecida para o utilizador, existe uma vantagem do estimador baseado no supernúcleo trapezoidal relativamente ao estimador baseado no núcleo normal. No entanto, sendo a janela assintoticamente óptima uma má aproximação da janela óptima, poderão existir problemas na escolha automática da janela quando esta é baseada na janela assintoticamente óptima.

## 4.2. Comparação baseada no EQI

A comparação efectuada no parágrafo anterior é aqui extendida ao caso em que a escolha da janela é feita de forma automática, ou seja, a janela  $h_n$  depende exclusi-

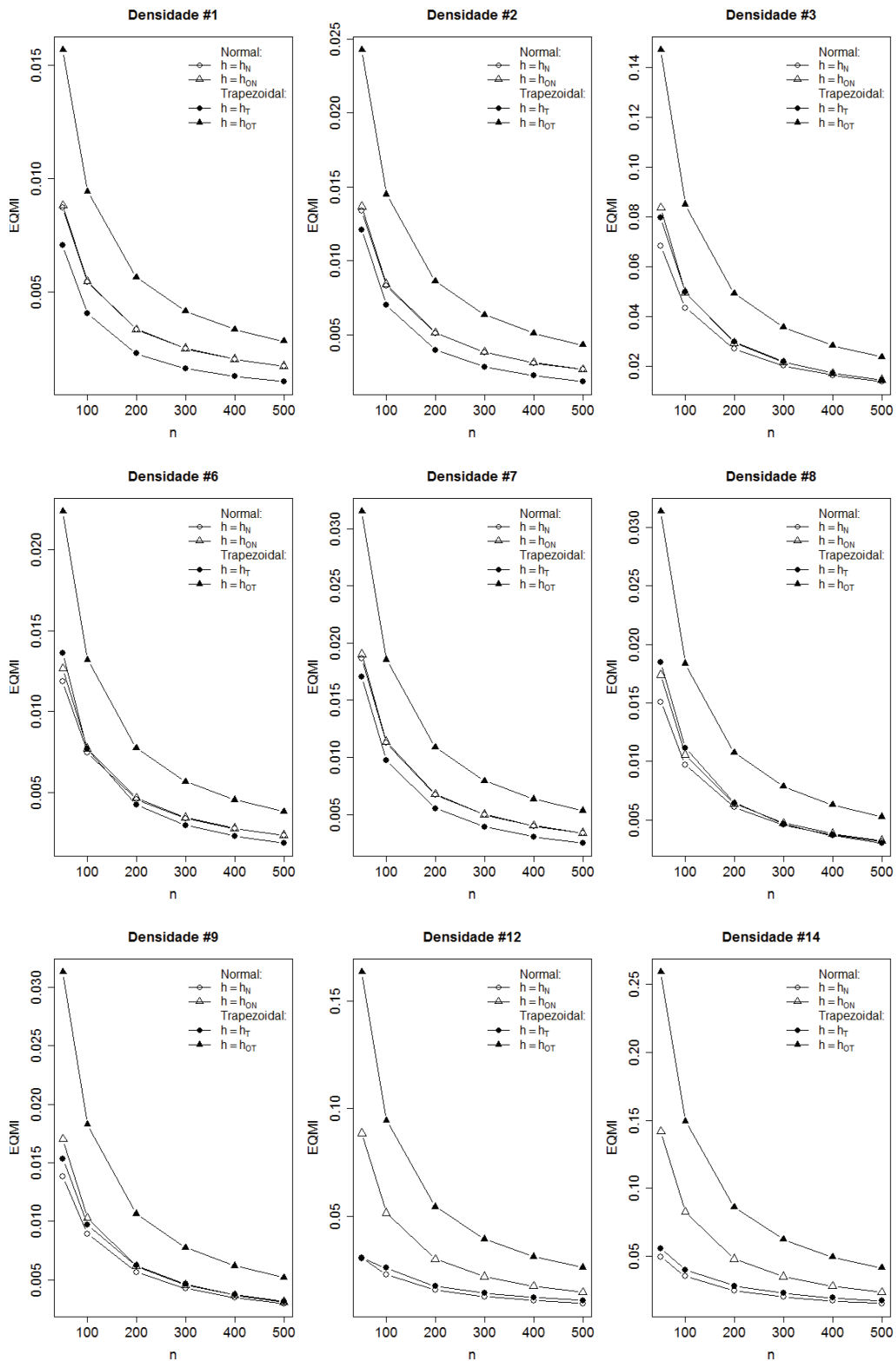


Figura 4.2: Comparação dos EQMI associados às janelas óptimas e às janelas assintoticamente óptimas.

vamente da amostra observada e não de características populacionais desconhecidas.

Para tal, iremos apresentar dois métodos automáticos para a escolha da janela: o método *plug-in* e o método de validação cruzada.

Neste caso, os estimadores serão comparados a partir dos erros quadráticos integrados, associados às janelas aqui obtidas de forma automática. Iremos comparar os estimadores baseados no núcleo normal, no supernúcleo trapezoidal e nas correcções deste último, proposta por Glad-Hjort-Ushakov e a apresentada na Secção 3.3.

#### 4.2.1. Escolha da janela pelo método *plug-in*

Como primeiro método de escolha da janela, consideramos o método *plug-in* a duas etapas que consiste na utilização das expressões das janelas assintoticamente óptimas presentes nos Teoremas 2.2.3 e 2.2.8 substituindo a quantidade desconhecida  $R(f'')$  por um estimador.

Assim, sendo  $\hat{R}(f'')$  um estimador de  $R(f'')$ , as janelas *plug-in* são dadas por

$$\hat{h}_N = \left( \frac{R(K)}{m_2^2(K)} \right)^{1/5} n^{-1/5} \hat{R}(f'')^{-1/5} \quad (4.1)$$

e

$$\hat{h}_T = \left( \frac{R(K)}{4} \right)^{1/5} n^{-1/5} \hat{R}(f'')^{-1/5}. \quad (4.2)$$

Para estimar a quantidade desconhecida  $R(f'')$  consideramos o estimador dado por

$$\hat{R}(f'') = \frac{1}{n^2} \sum_{i,j=1}^n \phi_g^{(4)}(X_i - X_j),$$

(ver Wand e Jones, 1995, pg. 67) onde  $\phi$  é o núcleo normal standard,  $g = g_n > 0$  é a janela e  $\phi_g^{(s)}$  representa a derivada de ordem  $s$  da função  $\phi_g(x) = \phi(x/g)/g$ , isto é,  $\phi_g^{(s)}(x) = \phi^{(s)}(x/g)/g^{s+1}$ .

Atendendo à definição de  $\hat{R}(f'')$  verificamos que este estimador não é completamente automático pois depende da escolha de uma nova janela,  $g$ . Pode demonstrar-se que a janela assintoticamente óptima  $g$ , no sentido da minimização do erro quadrático médio, é dada por

$$g = \left( \frac{-2\phi^{(6)}(0)}{R(f^{(4)})n} \right)^{1/9},$$

(ver Wand e Jones, 1995, pg. 67–72) dependendo assim da funcional desconhecida  $R(f^{(4)})$ , que será estimada pelo método das distribuições de referência. Tomando como distribuição de referência a distribuição normal, de média 0 e desvio-padrão



$\sigma$ ,  $R(f^{(4)})$  será então substituído por  $R(\phi_{\hat{\sigma}}^{(4)})$  onde  $\hat{\sigma}$  é o desvio-padrão empírico corrigido definido por

$$\hat{\sigma} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}.$$

Assim, a janela  $g$  que utilizaremos na prática é dada por

$$\hat{g} = \left( \frac{-2\phi^{(6)}(0)}{R(\phi_{\hat{\sigma}}^{(4)})n} \right)^{1/9} = \left( \frac{32\hat{\sigma}^9}{7n\sqrt{\pi}} \right)^{1/9}.$$

Na realização desta simulação, geramos 100 amostras de tamanho,  $n$ , igual a 50, 100 e 200 e calculamos o erro quadrático integrado associado às janelas definidas por (4.1) e (4.2).

Nas Figuras 4.3, 4.4 e 4.5, os resultados obtidos para o núcleo normal, supernúcleo trapezoidal, correcção de Glad-Hjort-Ushakov e a correcção apresentada na Secção 3.3 serão representados, respectivamente, por  $N_{PI}$ ,  $T_{PI}$ ,  $C_{PI}$  e  $\bar{C}_{PI}$ . Analisando os resultados apresentados, podemos concluir que, para todas as densidades, o estimador baseado no núcleo normal apresenta sempre melhores resultados que os estimadores baseados no supernúcleo trapezoidal não corrigido e corrigido (Glad-Hjort-Ushakov e a correcção apresentada na Secção 3.3). Estes resultados não são surpreendentes atendendo aos resultados apresentados na Secção 4.1.3 onde vimos que  $h_{0T}$  era, para todas as densidades consideradas, uma má aproximação da janela óptima.

Quanto à correcção de Glad-Hjort-Ushakov, verificamos que à medida que o tamanho da amostra aumenta, o EQI associado à correcção tende a igualar o EQI do estimador baseado no supernúcleo trapezoidal. Tal não surpreende pois com o aumento do tamanho da amostra o estimador corrigido tende a aproximar-se do estimador baseado no supernúcleo. Quando a amostra é de tamanho 50 e 100, as densidades #3, #12 e #14 são as que apresentam um EQI da correcção de Glad-Hjort-Ushakov inferior ao EQI do estimador baseado no supernúcleo trapezoidal enquanto que para as outras densidades os EQI são muito próximos. No caso em que a amostra é de tamanho 200 só para as densidades #12 e #14 é que verificamos que o EQI da correcção de Glad-Hjort-Ushakov continua a ser inferior ao EQI do estimador baseado no supernúcleo trapezoidal.

Comparando agora as duas correcções apresentadas, verificamos que, para todas as densidades estudadas, a correcção apresentada na Secção 3.3 apresenta sempre um EQI melhor ou igual que o EQI da correcção de Glad-Hjort-Ushakov.

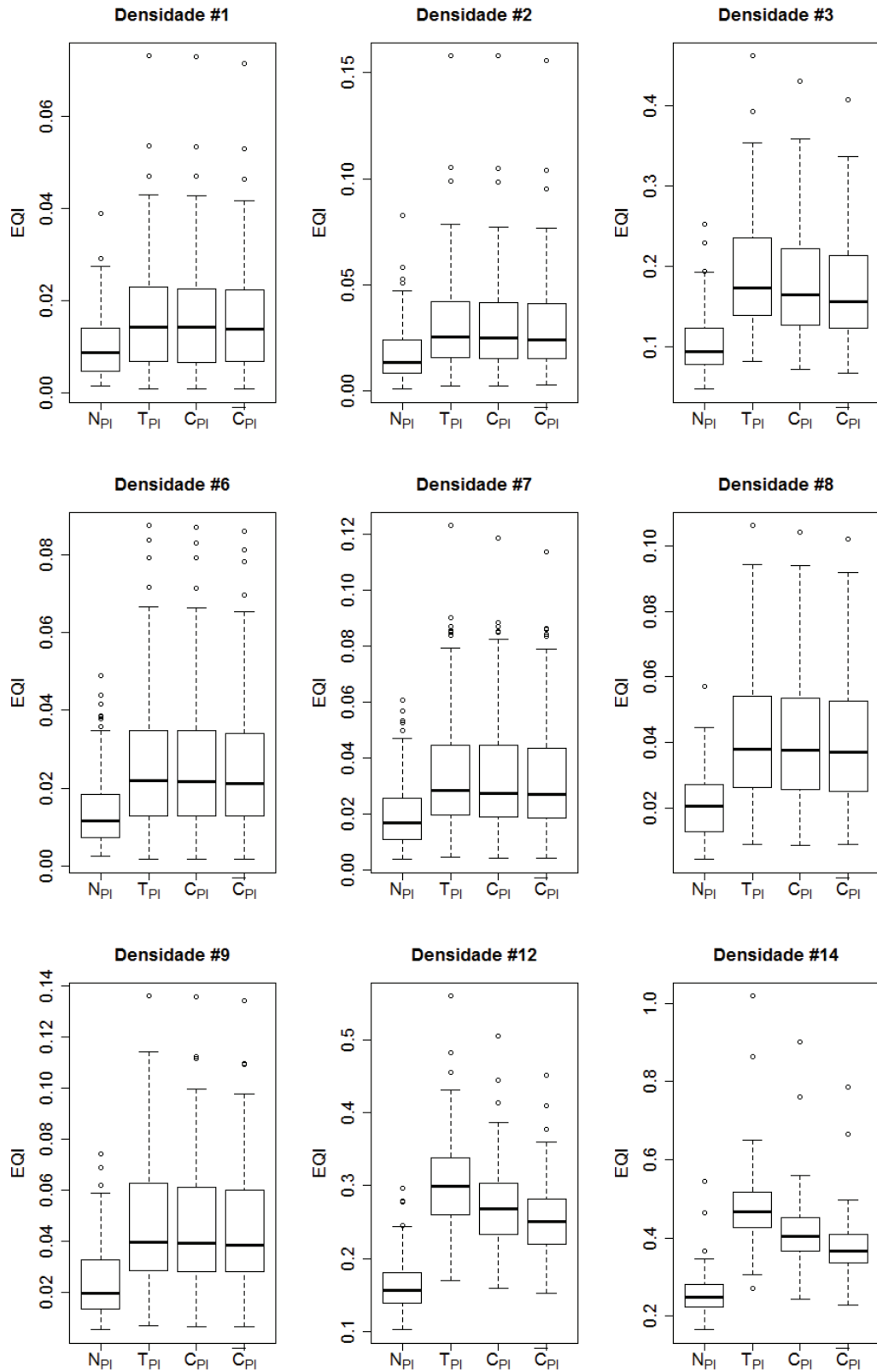


Figura 4.3: Erro quadrático integrado, com o método *plug-in*, com o tamanho da amostra  $n = 50$  e o número de repetições para cada densidade é de 100.

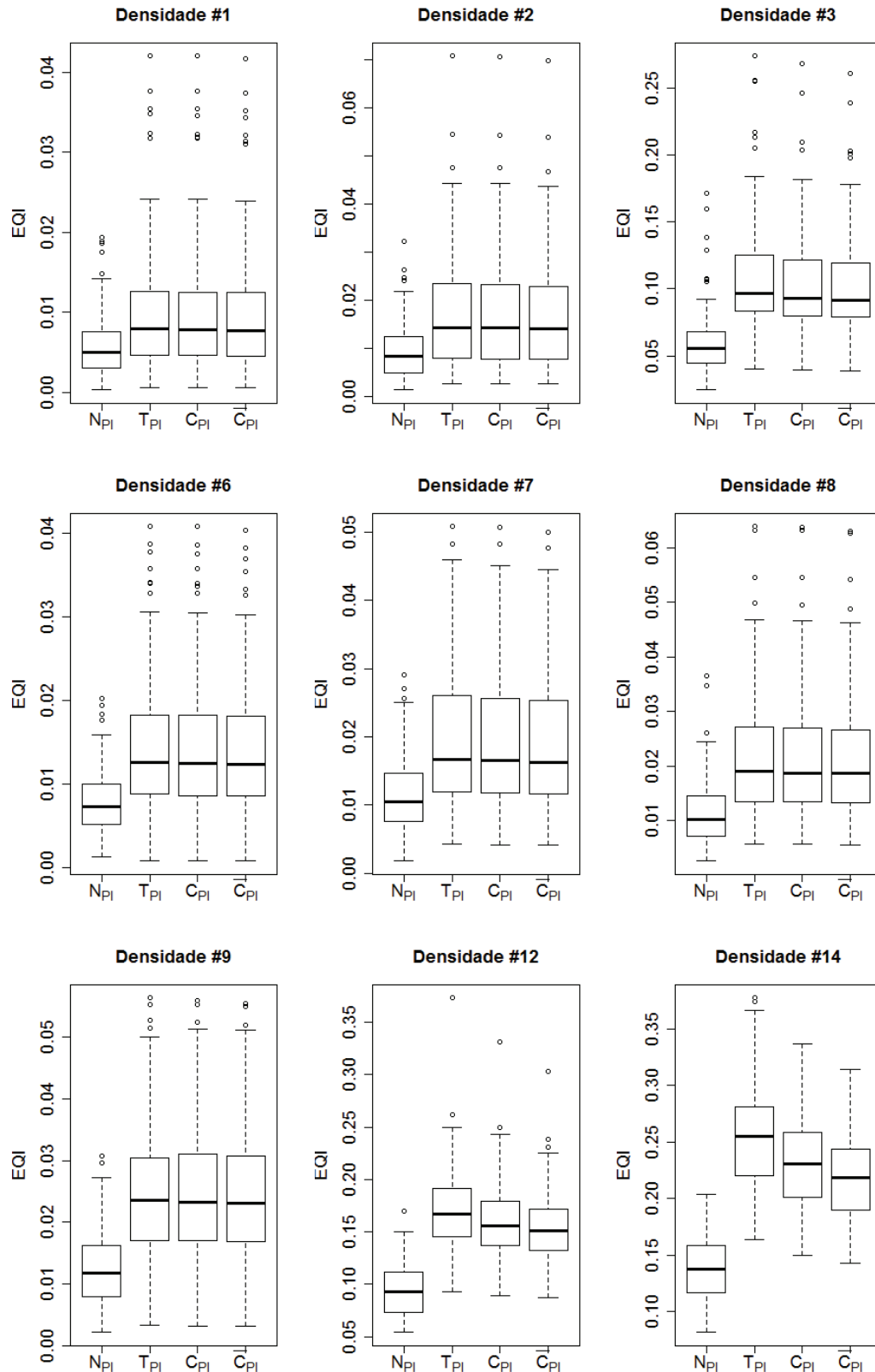


Figura 4.4: Erro quadrático integrado, com o método *plug-in*, com o tamanho da amostra  $n = 100$  e o número de repetições para cada densidade é de 100.

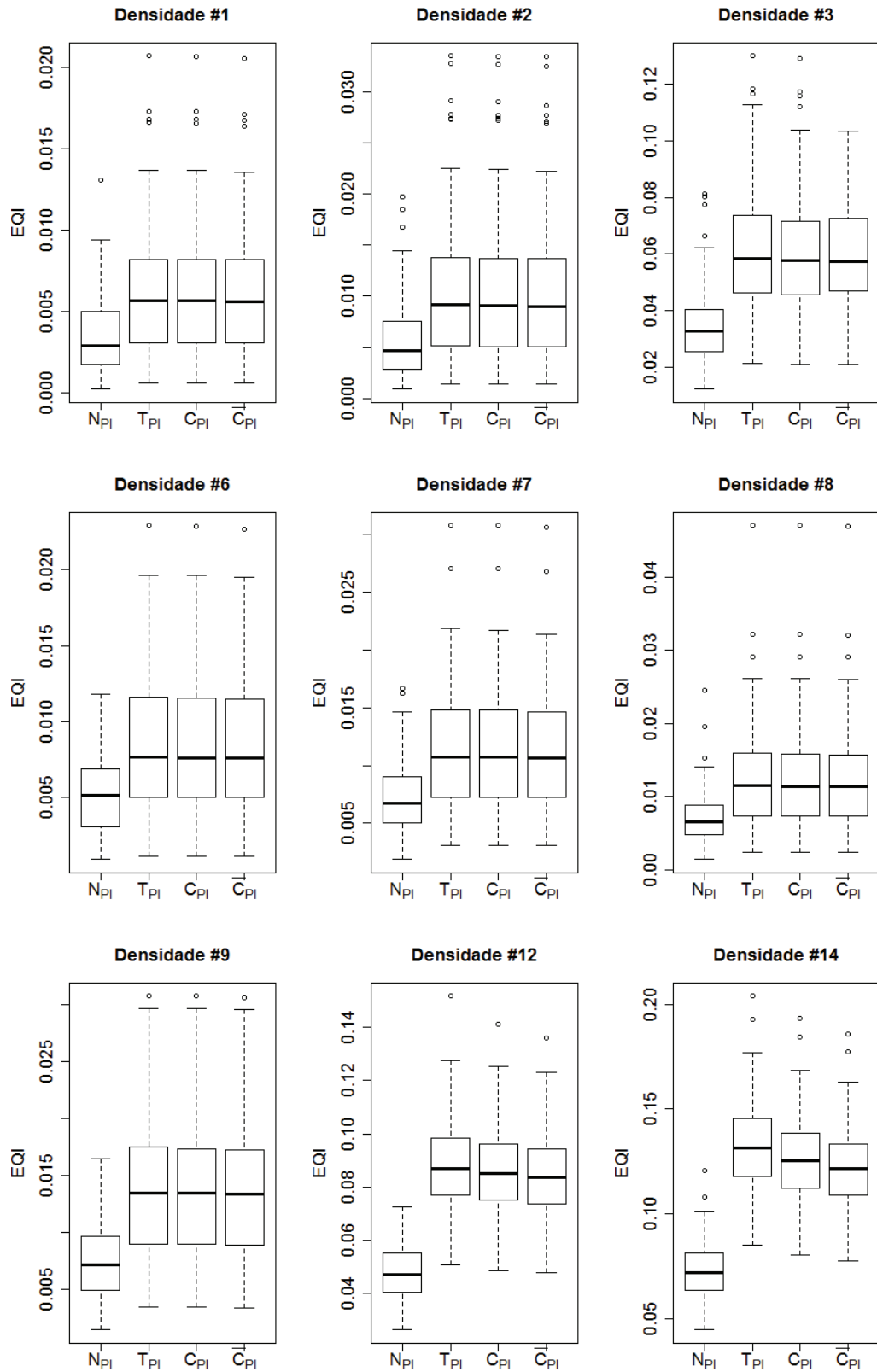


Figura 4.5: Erro quadrático integrado, com o método *plug-in*, com o tamanho da amostra  $n = 200$  e o número de repetições para cada densidade é de 100.

### 4.2.2. Escolha da janela pelo método de validação cruzada

Como no caso do estimador baseado no supernúcleo trapezoidal a janela assintoticamente óptima, no sentido da minimização dum majorante do EQMI, é uma má aproximação da janela óptima, facto visto na Secção 4.1.3, vamos agora abordar um método de escolha da janela alternativo ao método do *plug-in* que não se baseia nas expressões das janelas assintoticamente óptimas.

O método aqui abordado é o chamado método de validação cruzada, baseado no erro quadrático integrado, definido no Capítulo 3, e tem por objectivo escolher uma janela que minimiza o estimador cêntrico de  $\text{EQI}(h) - R(f)$  definido por

$$\text{CV}(h) = \frac{R(K)}{nh} + \frac{1}{n(n-1)} \sum_{i \neq j} \left( \frac{n-1}{n} \bar{K}_h * K_h - 2K_h \right) (X_i - X_j).$$

Assim, a janela é dada por

$$\hat{h}_{\text{CV}} = \arg \min_{h>0} \text{CV}(h).$$

As primeiras propriedades teóricas sobre a janela anterior, no caso do núcleo normal, foram estabelecidas por Hall (1983) e por Stone (1984). Quando o núcleo  $K$  é um supernúcleo não são conhecidos resultados teóricos sobre a janela  $\hat{h}_{\text{CV}}$ .

Para o nosso estudo é necessário ter em conta que no caso do supernúcleo trapezoidal a expressão do produto de convolução é dada por

$$\bar{K} * K(x) = \begin{cases} \frac{2}{\pi} \left( \frac{\cos x}{x^2} + \frac{\sin x}{x^3} - \frac{\sin(2x)}{x^3} \right) & , x \neq 0 \\ \frac{4}{3\pi} & , x = 0 \end{cases}$$

e no caso do núcleo normal o produto de convolução é dado pela densidade normal de média 0 e variância 2.

Nestas simulações, analogamente às apresentadas, geramos 100 amostras de tamanho,  $n$ , igual a 50, 100 e 200 e, os resultados estão representados nas Figuras 4.6, 4.7 e 4.8.

Analisando os resultados obtidos verificamos que, para as densidades #1, #2, #7 e #12, o estimador baseado no supernúcleo trapezoidal (não corrigido) apresenta um EQI inferior ao do estimador baseado no núcleo normal e, para as outras densidades é o estimador baseado no núcleo normal que apresenta melhor EQI. No caso em que o tamanho da amostra é 200 verificamos que, também para a densidade #6, o estimador baseado no supernúcleo trapezoidal apresenta um EQI inferior ao do

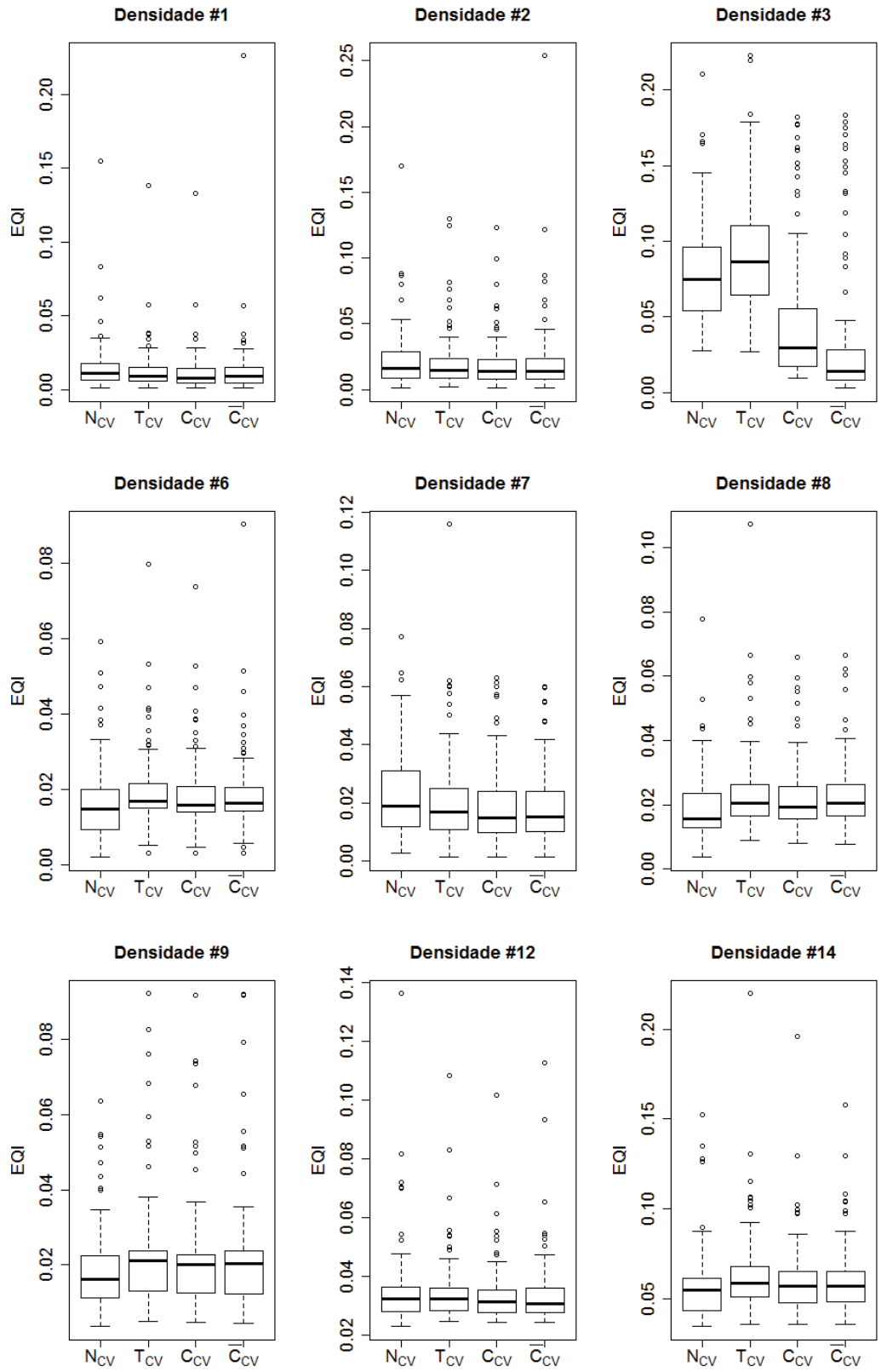


Figura 4.6: Erro quadrático integrado, com o método de validação cruzada, com o tamanho da amostra  $n = 50$  e o número de repetições para cada densidade é de 100.

estimador baseado no núcleo normal. Comparando estes resultados com os resultados teóricos apresentados no parágrafo 4.1.3 verificamos que os resultados aqui obtidos ilustram os apresentados no parágrafo 4.1.3.

Comparando agora o estimador baseado no núcleo normal com o estimador baseado no supernúcleo trapezoidal corrigido (Glad-Hjort-Ushakov), quando o tamanho da amostra é 50, para as densidade #1, #2, #3, #7 e #12, verificamos que o EQI do estimador corrigido é inferior ao do estimador baseado no núcleo normal. Quando o tamanho da amostra é 100, o estimador corrigido apresenta um EQI inferior para as densidades anteriores e também para a densidade #14. No caso em que a amostra é de tamanho 200 tal acontece para as densidades #1, #2, #3, #6 e #7. Verificamos também que o estimador corrigido apresenta grande vantagem para a densidade #3. Tal facto pode ser explicado devido à sua característica distribucional, pois é caracterizada por uma forte assimetria.

Relativamente às correcções apresentadas verificamos que, quando o tamanho da amostra é 50, para as densidades #3 e #12, a correcção apresentada na Secção 3.3 apresenta EQI inferior ao da correcção de Glad-Hjort-Ushakov, enquanto que para as outras densidades é a correcção de Glad-Hjort-Ushakov que apresenta EQI inferior. Quando a amostra é de tamanho 100, para as densidades #3 e #6 a correcção apresentada na Secção 3.3 apresenta EQI inferior ao da correcção de Glad-Hjort-Ushakov e, no caso em que o tamanho da amostra é 200 tal acontece apenas para a densidade #3.

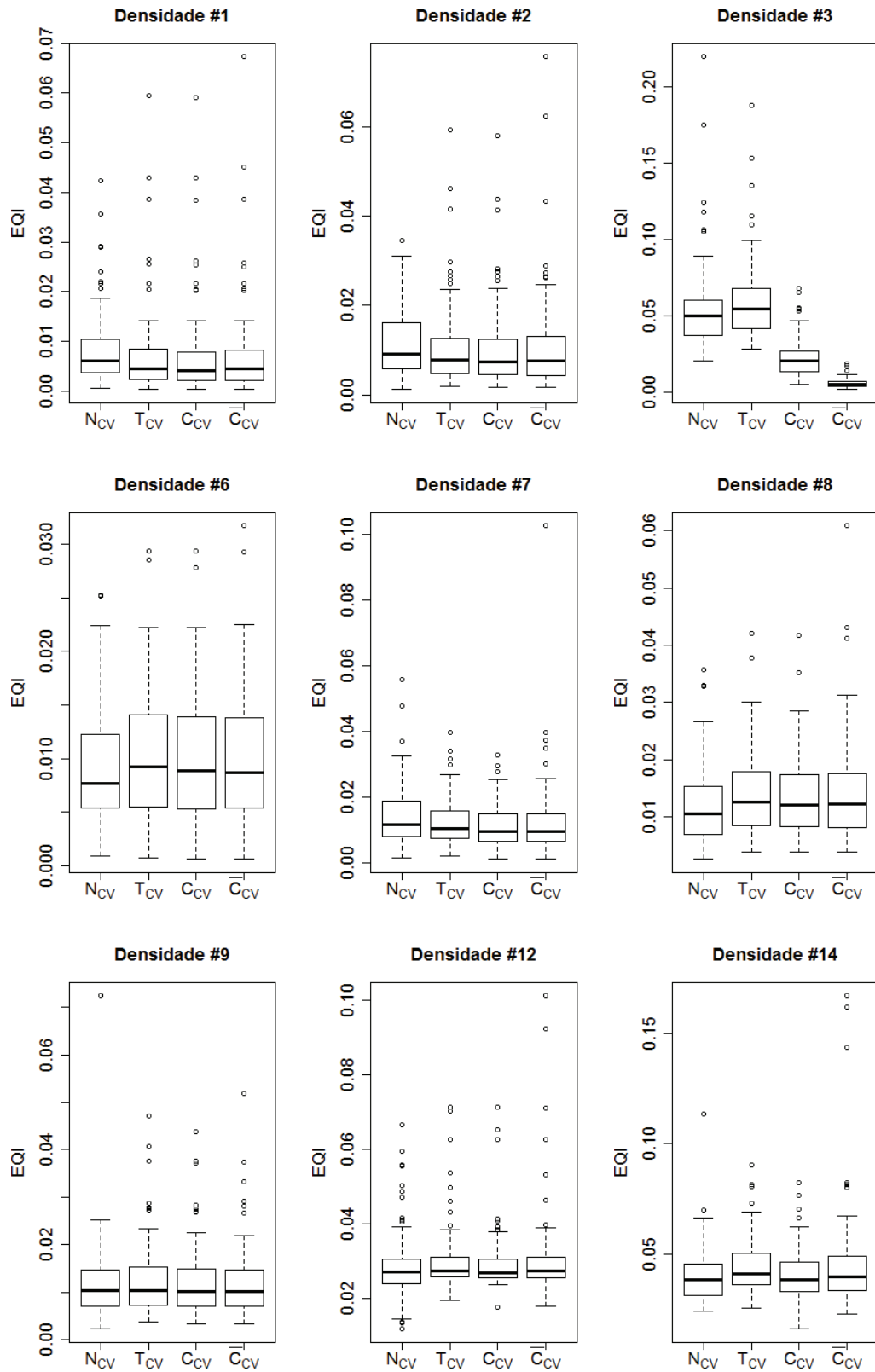


Figura 4.7: Erro quadrático integrado, com o método de validação cruzada, com o tamanho da amostra  $n = 100$  e o número de repetições para cada densidade é de 100.



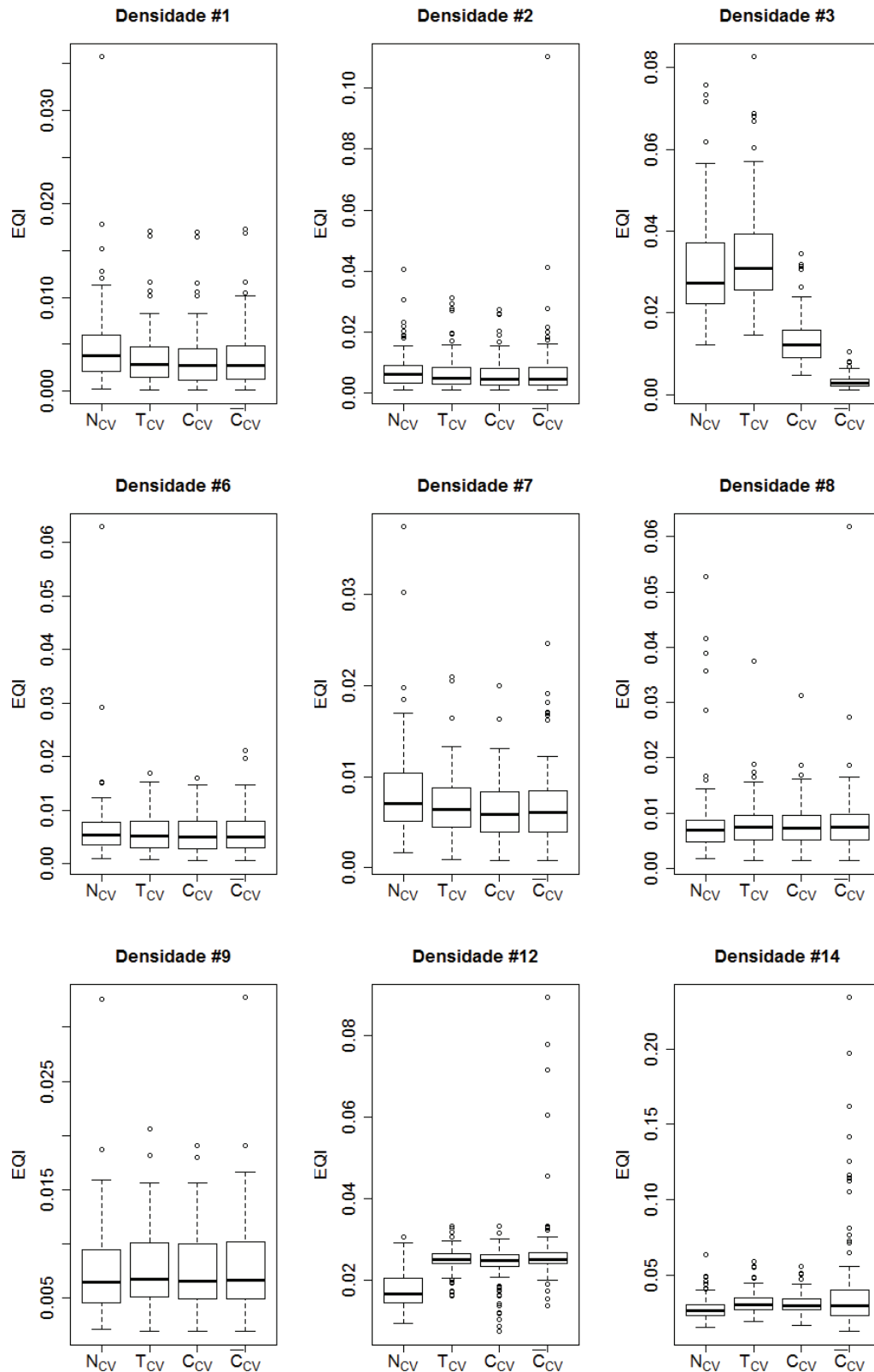


Figura 4.8: Erro quadrático integrado, com o método de validação cruzada, com o tamanho da amostra  $n = 200$  e o número de repetições para cada densidade é de 100.



# Bibliografia

- Chacón, J.E., Montanero, J., Nogales, A.G. (2007a). A note on kernel density estimation at a parametric rate. *Journal of Nonparametric Statistics* 19, 13–21.
- Chacón, J.E., Montanero, J., Nogales, A.G., Pérez, P. (2007b). On the existence and limit behaviour of the optimal bandwidth in kernel density estimation. *Statistica Sinica* 17, 289–300.
- Deheuvels, P. (1977). Estimation non paramétrique de la densité par histogrammes généralisés. *Revue de Statistique Appliquée* 25, 5–42.
- Devroye, L. (1987). *A course in density estimation*. Birkhäuser, Boston.
- Donoghue, W.F. (1969). *Distributions and Fourier Transforms*. Academic Press, New York and London.
- Epanechnikov, V.A. (1969). Nonparametric estimation of a multivariate probability density. *Theory of Probability and its applications* 14, 153–158.
- Fernandez, P.J. (1976). *Medida e Integração*. Projecto Euclides, Brasil.
- Fryer, M.J. (1976). Some errors associated with the nonparametric estimation of density functions *IMA Journal of Applied Mathematics* 18, 371–380.
- Glad, I.K., Hjort, N.L., Ushakov, N.G. (2003). Correction of density estimators that are not densities. *Scandinavian Journal of Statistics* 30, 415–427.
- Hall, P. (1983). Large sample optimality of least squares cross-validation in density estimation. *The Annals of Statistics* 11, 1156–1174.
- Jones, M.C., Foster, P.J. (1993). Generalized jackknifing and higher order kernels. *Journal of Nonparametric Statistics* 3, 81–94.
- Lima, E.L. (1992). *Curso de análise, Vol. 1*. Projecto Euclides, Brasil.

Mächler, M. (2013). *nor1mix: Normal (1-d) Mixture Models (S3 Classes and Methods)*. R package version 1.1-4. <http://CRAN.R-project.org/package=nor1mix>.

Malliavin, P. (1982). *Intégration et probabilités, analyse de Fourier et analyse spectrale*. Masson, Paris.

Marron, J.S., Wand, M.P. (1992). Exact mean integrated square error. *The Annals of Statistics* 20, 712–736.

Parzen, E. (1962). On estimation of a probability density function and mode. *The Annals of Mathematical Statistics* 33, 1065–1076.

R Core Team (2012). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.

Rosenblatt, M. (1956). Remarks on some non-parametric estimates of a density function. *The Annals of Mathematical Statistics* 27, 832–837.

Ruppert, D., Wand, M.P. (1992). Multivariate locally weighted least squares regression. *The Annals of Statistics* 22, 1346–1370.

Stone, C.J., (1984). An asymptotically optimal window selection rule for kernel density estimates. *The Annals of Statistics* 12, 1285–1297.

Tenreiro, C. (2010). Uma Introdução à Estimaco No-Paramétrica da Densidade. *Minicurso, XVIII Congresso Anual da Sociedade Portuguesa de Estatística*. Coimbra.

Wand, M. (2012). *KernSmooth: Functions for kernel smoothing for Wand & Jones (1995)*. R package version 2.23-8. <http://CRAN.R-project.org/package=KernSmooth>.

Wand, M.P., Jones, M.C. (1995). *Kernel Smoothing*. Chapman & Hall, London.