

Dissertação de Mestrado

Escolha do Parâmetro de Suavidade na  
Estimativa da Função de Distribuição

por

Lupércio França Bessegato

Orientador:

Gregorio Saravia Atuncar

Co-orientador:

Luiz Duczmal

Dezembro de 2001

Lupércio França Bessegato

# Escolha do Parâmetro de Suavidade na Estimativa de Função de Distribuição

Dissertação apresentada ao Departamen-  
to de Estatística do Instituto de Ciências  
Exatas da Universidade Federal de Minas  
Gerais, como requisito parcial à obtenção  
do título de Mestre em Estatística.

Universidade Federal de Minas Gerais

Belo Horizonte, dezembro de 2001

# Agradecimentos

Agradeço ao professor Gregório Saravia Atuncar pelo incentivo, orientação, paciência e dedicação dispensados no decorrer de todo o trabalho; ao professor Luiz Duczmal pelo apoio e pela atenção; aos professores e funcionários do Departamento de Estatística pelo incentivo e orientação, em especial à Marcinha, incentivadora desde a primeira hora; ao Vagner Facin, por disponibilizar parte fundamental do tempo necessário aos meus estudos; ao Fabrício, pelo apoio, pelo auxílio e pela presença sempre constante.

A minha mãe por seu exemplo e amor, dedico este trabalho,

# Sumário

<b>Resumo</b>	<b>5</b>
<b>1 Introdução</b>	<b>6</b>
<b>2 Revisão da Literatura</b>	<b>8</b>
<b>3 Metodologia</b>	<b>13</b>
3.1 Definições Básicas e Alguns Resultados . . . . .	13
3.2 Estimativa pelo núcleo estimador da função de distribuição . . . . .	15
3.2.1 Propriedades de $\hat{\mathbf{F}}_n$ . . . . .	16
3.2.2 Escolha da janela ótima . . . . .	19
3.2.3 Método “plug-in” . . . . .	22
3.2.4 Convergência Quase Certa de $\hat{\mathbf{H}}$ . . . . .	23
3.3 Estimativa pela integral do núcleo estimador da função de densidade . . . . .	29
<b>4 Implementação Computacional</b>	<b>30</b>
4.1 Considerações Gerais . . . . .	30

4.2	Geração de Números Aleatórios . . . . .	32
4.3	Limite de Integração $\Lambda$ . . . . .	32
4.4	Integração Numérica . . . . .	33
<b>5</b>	<b>Simulações e Aplicações</b>	<b>34</b>
5.1	Simulação . . . . .	34
5.2	Estimação de Capacidade de Processo . . . . .	41
5.3	Estimação de Taxa de Falha . . . . .	43
5.4	Aplicação . . . . .	46
<b>6</b>	<b>Conclusões e Pesquisas Futuras</b>	<b>49</b>
	<b>Apêndice</b>	<b>51</b>
	<b>Referências Bibliográficas</b>	<b>67</b>

# Lista de Tabelas

4.1	Resumo de Algumas Funções Disponíveis em S-plus . . . . .	31
5.1	Valores de H, Janela Ótima e MISE Assintótico . . . . .	35
5.2	Sumário das Simulações - Estimacão de H . . . . .	36
5.3	Sumário das Simulações - Estimacão de $h_{opt}$ . . . . .	37
5.4	Sumário das Simulações - Estimacão do MISE . . . . .	38
5.5	Resumo de Simulacão - Estimacão de $p$ . . . . .	43
5.6	Resultado de Simulacão - Estimativa de Taxa de Falha . . . . .	44

# Lista de Figuras

5.1	Histogramas de Acompanhamento de Simulação . . . . .	40
5.2	Box-plot de Amostras Normais . . . . .	40
5.3	Box-plot de Amostras Gama . . . . .	41
5.4	Box-plot de Amostras de Distribuição Bimodal . . . . .	41
5.5	Estimação de Taxa de Falha - Amostra Weibull (2,1) . . . . .	45
5.6	Tensão de Escoamento de Aço - Histograma . . . . .	46
5.7	Tensão de Escoamento de Aço - Estimação da Densidade . . . . .	47
5.8	Tensão de Escoamento de Aço - Estimação da Distribuição . . . . .	47

# Resumo

O método do núcleo estimador tem sido largamente utilizado na estimação da função de distribuição de probabilidades, consolidando-se como uma alternativa às abordagens paramétricas.

No Capítulo 1, apresentamos a motivação e os objetivos deste trabalho, enquanto que no Capítulo 2, uma revisão da literatura, dando uma visão geral do núcleo estimador e de suas propriedades, assim como sobre os métodos para a escolha da janela ótima neste contexto.

No Capítulo 3, desenvolvemos a metodologia para escolha de janela ótima e propomos o estimador tipo “plug-in”, baseado naquele desenvolvido por Chiu [8] em estimação de função de densidade, e provamos a consistência forte deste estimador.

No Capítulo 4, resumimos a implementação computacional necessária para suportar as simulações efetuadas e para a utilização do estimador proposto em aplicações reais. No Capítulo 5, apresentamos os resultados das simulações e aplicações deste método, comparando seu desempenho com alguns outros estabelecidos na literatura.

Finalmente, concluímos a presente dissertação apresentando os próximos passos da pesquisa no sentido de obter-se o aprimoramento do estimador e para adequá-lo à maior gama de situações possíveis.

# Capítulo 1

## Introdução

Dada uma amostra aleatória, podem-se fazer várias abordagens para se estimar a distribuição de probabilidades da variável de interesse, ou através de sua função de densidade, ou através de sua função de distribuição.

Dentre os processos de estimação, o método do núcleo estimador, aproveitando-se de suas propriedades assintóticas, tem sido largamente usado. A taxa de convergência e a suavidade do estimador dependem da escolha de uma largura de janela. Neste sentido, torna-se de extrema relevância estudar os estimadores da janela ótima para a obtenção da estimação mais apropriada da função que exprima a probabilidade de ocorrência dos dados.

O objetivo principal deste trabalho é estimar a função de distribuição de uma variável aleatória. Abordaremos este problema de duas maneiras:

1. Núcleo estimador da função de distribuição:

Seja uma amostra aleatória  $X_1, X_2, \dots, X_n$  de uma variável aleatória  $X$  com função de distribuição  $F$ . O estimador de  $F$ , avaliada no ponto  $x$ , será dado por:

$$\hat{F}_{1n}(x) = \frac{1}{n} \sum_{i=1}^n W\left(\frac{x - X_i}{h_n}\right) \quad (1.1)$$

onde  $W$  é uma função de distribuição e será chamada núcleo.

O trabalho, nesta abordagem, estará centrado na escolha da janela ótima  $h_{opt}$ . Sali-

tamos que a janela ótima depende também de  $n$ .

2. Estimador de  $F$  como integral do núcleo estimador da função de densidade:

Seja uma amostra aleatória  $X_1, \dots, X_n$  de uma variável  $X$  com função de densidade  $f = F'$ . O estimador de  $F$ , avaliada no ponto  $x$ , será dado por:

$$\hat{F}_{2n}(x) = \int_{-\infty}^x \hat{f}_n(x) dx \quad (1.2)$$

onde,

$$\hat{f}_n(x) = \frac{1}{nh_n} \sum_{i=1}^n w\left(\frac{x - X_i}{h_n}\right) \quad (1.3)$$

sendo  $w$  uma função de densidade e será chamado núcleo.

Usaremos os resultados obtidos por Damasceno [10] e compararemos o estimador obtido desta maneira com aquele obtido através do método descrito no item anterior.

Salienta-se que a largura da janela ótima do núcleo estimador ( $h_{opt}$ ) da função de densidade é da ordem de  $n^{-1/5}$  e no caso da estimação da função de distribuição, é da ordem de  $n^{-1/3}$ .

Esta dissertação pretende estudar as várias abordagens que podem ser feitas para se escolher a janela a ser usada na definição do núcleo estimador  $\hat{F}_{1n}$  de funções de distribuições. Os métodos a serem estudados contemplam o caso i.i.d., embora possam ser estendidos a situações gerais, como no caso Markoviano, para estimar a função de distribuição invariante de um processo. Temos a intenção de fazer essa extensão em trabalhos posteriores.

# Capítulo 2

## Revisão da Literatura

O uso do método do núcleo para estimação de funções de distribuição tem sido largamente estudado na literatura, embora as técnicas usadas na estimação da função de densidade não sejam diretamente aplicáveis no caso da estimação da função de distribuição.

Diversos autores investigaram as propriedades do estimador  $\hat{F}_n(x)$ , obtido a partir da integral do núcleo estimador da função de densidade. Nadaraya [27], Winter [47] e Yamato [44] provaram a convergência de  $\hat{F}_n(x)$  para  $F(x)$ , com probabilidade 1. Reiss [33] provou que a deficiência relativa da função de distribuição empírica  $\tilde{F}_n(x)$ , com respeito a uma escolha apropriada de  $\hat{F}_n(x)$ , tende rapidamente ao infinito quando o tamanho da amostra cresce.

Jones [23] estudou o desempenho de diversos núcleos  $W$  para a estimação de funções de distribuição e verificou que a escolha da função núcleo não apresenta um grande impacto no desempenho do núcleo estimador  $\hat{F}_n(x)$ . Falk [14] estudou se um dado núcleo estimador da função de distribuição, em um ponto, tem melhor desempenho assintótico que o da função de distribuição empírica e conjecturou que não há núcleo ótimo, considerada a deficiência ou eficiência relativa à função de distribuição empírica.

Hjort e Walker [21] estudaram o comportamento da função de distribuição acumulada, correspondendo ao núcleo estimador da função de densidade, verificando a necessidade de uma janela de tamanho  $O(n^{-1/4})$ , para que a estimativa da função de distribuição assim

calculada ficasse dentro de qualquer intervalo de confiança razoável da função de distribuição empírica, com probabilidade tendendo a 1, quando o tamanho amostral tende ao infinito. O fato de o parâmetro de suavidade ótimo para o núcleo estimador da função de distribuição ser da forma  $an^{-1/3}$  assegura que as estimativas assim obtidas pertencerão a esta banda de confiança em torno da função de distribuição empírica.

Azzalini [4] derivou uma expressão assintótica para o erro quadrático médio de  $\hat{F}_n(x)$  e determinou o parâmetro de suavidade ótimo, estudado posteriormente por Hill [20] e Swanepoel [43].

Para medir o desempenho global das estimativas, é comum o uso do erro quadrático médio integrado, MISE (Mean Integrated Squared Error), definido por

$$MISE(h) = E \int \left\{ \hat{F}_n(x) - F(x) \right\}^2 dx \quad (2.1)$$

Nas condições de o núcleo  $w$  ser função de densidade e  $F$  ser duas vezes continuamente diferenciável, Swanepoel [43] derivou a expressão do erro quadrático médio integrado (MISE), estudando um parâmetro de suavidade ótimo de maneira a minimizar o MISE e indicou que, assim, o núcleo estimador  $\hat{F}_n(x)$  é assintoticamente mais eficiente que a função de distribuição empírica  $\tilde{F}_n(x)$ . A questão referente à adequabilidade do ISE ou do MISE como as expressões de minimização apropriadas à escolha da janela ótima foram discutidas por Jones (1991a), embora Shirahata [38] tenha indicado que a superioridade do núcleo estimador  $\hat{F}_n(x)$  sobre a função de distribuição empírica  $\tilde{F}_n(x)$  não é necessariamente verdadeira em termos do erro quadrático integrado (ISE), definido como  $ISE(h) = \int \left\{ \hat{F}_n(x) - F(x) \right\}^2 dx$ .

O parâmetro  $h$  é chamado parâmetro de suavidade, ou janela, e controla a suavidade da estimativa de  $\hat{F}_n(x)$  e sua escolha é uma questão crítica na estimação de função de distribuição, sendo discutida extensamente na literatura. Em geral,  $h$  é escolhido de maneira que  $\hat{F}_n(x)$  seja um ótimo estimador de  $F$ , de acordo com alguma medida de desempenho.

O MISE é difícil de ser calculado para amostras finitas, tendo Jones [23] derivado uma expressão para uma aproximação de seu valor. Está disponível há algum tempo uma expressão para a janela que minimiza seu valor, verificando-se que este valor ótimo,  $h_{opt}$

infelizmente depende da função desconhecida  $F$ , significando que  $h_{opt}$  deve ser estimado baseando-se nos dados observados.

A literatura aborda de várias maneiras a escolha da janela ótima  $h_{opt}$ . Embora, na prática, seja possível escolher o parâmetro de suavização de maneira subjetiva, há uma grande demanda por procedimentos automáticos para seleção da janela. Algumas razões para se usar este tipo de procedimento estão dadas em Silverman [39].

O seletor automático mais estudado em estimação função de densidade é o da função escore de validação cruzada de mínimos quadrados, proposto por Rudemo [34] e Bowman [5]. Foi mostrado em Hall [17] e Stone [41] que o minimizador de função escore de validação cruzada é uma estimativa consistente da janela ótima. A normalidade assintótica da estimativa da janela foi estabelecida em Hall e Marron [18] e Scott e Terrell [37]. Entretanto, dos resultados assintóticos dos procedimentos de validação cruzada, verifica-se que a estimativa da janela está sujeita a uma grande variação amostral, sendo que em estudos de simulação observou-se que o seletor tende a escolher valores de janela menores, com mais frequência que o predito pelos teoremas assintóticos (veja Chiu [8]). Alguns estudos acerca do aprimoramento da validação cruzada podem ser encontrados em Scott e Terrell [37], Park e Marron [29] e Hall, Marron e Park [19]. Sabe-se que janelas menores tendem a mostrar falsas estruturas dos dados, estando esta dificuldade claramente demonstrada em Chiu [8]. Isto limita, na prática, a utilização do método de validação cruzada.

De maneira similar àquela utilizada em estimação da função de densidade, Bowman e outros [6] propuseram um método de validação cruzada para a suavização de funções de distribuição. Neste método, a seleção do parâmetro da janela é baseado em estimação não-viciada do MISE (Mean Integrated Squared Error), levando seu procedimento de minimização a uma escolha assintótica da janela ótima.

Outra abordagem possível na escolha da janela ótima é através da utilização de método "plug-in", que estima o valor da única quantidade desconhecida na expressão do erro quadrático médio integrado assintótico, ou seja, a parcela dependente da função que se quer estimar ( $\int [f'']^2$ , no caso da estimação de função de densidade ou  $\int [F''']^2$ , no caso da

estimação da função de distribuição). Salienta-se que o método “plug-in” tem a aparente vantagem de, em seu cálculo, não necessitar de uma rotina de otimização.

No caso de estimação de densidade, este tipo de procedimento foi estudado em Woodroffe [48] e Scott e Factor [36]. Chiu [8] prestou importante colaboração ao propor estimadores “plug-in” e “plug-in” ajustados, baseados em funções características, além de estudar suas distribuições assintóticas e de comparar o desempenho desses estimadores com aqueles obtidos a partir de métodos de validação cruzada. Um dos objetivos da presente pesquisa é estudar o comportamento de um estimador da função de distribuição análogo ao estimador proposto por Chiu.

Altman e Leger [1] propuseram métodos “plug-in” para estimação da função de distribuição análogos àqueles usados largamente em estimação de função de densidade. Polanski [31] propôs também um estimador tipo “plug-in” para a escolha da janela ótima do núcleo estimador da função de distribuição, derivando suas propriedades assintóticas e verificando seu desempenho em um conjunto de 15 misturas de normais, conforme considerado por Marron e Wand [25].

Embora em algumas aplicações seja desejável obter-se uma estimativa contínua e suavizada de  $F$ , a função de distribuição empírica é uma estimativa muito boa e não se espera que, assintoticamente, as estimativas através do núcleo forneçam uma melhoria substancial no erro quadrático médio integrado. De fato, verifica-se que a suavização terá apenas um efeito de segunda ordem no erro médio quadrático integrado (MISE), esperando-se que um grande ganho no desempenho, como resultado de suavização, seja evidente apenas em amostras pequenas ou moderadamente grandes.

Uma aplicação possível no uso dos núcleos estimadores da função de distribuição é na avaliação de capacidade de um processo, quando é necessário obter-se estimativas dos quantis situados nas caudas da distribuição. Polansky [30] propôs um estimador desta capacidade, baseado em núcleo, de maneira a evitar o uso de estimador paramétrico em situações de desvios da normalidade dos dados e, em estudo de simulação, compara o desempenho deste estimador em face de outros estimadores comumente utilizados.

Outro problema bastante especial é a estimação de função de densidade ou de distribuição na presença de dados censurados, pois em situações de estudos de tempo de falha é importante se obter estimativas não-paramétricas de várias das características da função de sobrevivência. O estimador de Kaplan-Meier é baseado em dados censurados à direita, sendo um estimador de máxima verossimilhança da função de sobrevivência  $S$ . Uma das características usuais envolvendo esta função é o interesse em se estimar quantis. Padgett [28] propôs um núcleo estimador da função quantil, baseado na suavização do estimador de Kaplan-Meier. Entretanto, como não há resultados do erro quadrático médio integrado do estimador proposto e, conseqüentemente, não se pode obter os valores da janela que minimizam o MISE neste caso, o autor sugere o uso de método bootstrap para escolha da janela em algumas situações.

A estimação da função taxa de falha é também uma questão recorrente na literatura, tendo sido estudada por Watson e Leadbetter [46], que derivaram a fórmula assintótica da variância do estimador  $\hat{r}(t) = \hat{f}(t)/[1 - \hat{F}(t)]$ . Esta abordagem foi estudada também por Silverman [39], que discorre sobre sua utilização em problemas envolvendo dados de sobrevivência e confiabilidade. Klein [24] apresenta um estimador da taxa de falha baseado na suavização da estimativa de Nelson-Allen, apresentando método de validação cruzada para escolha da janela. Similarmente, Sun [42] propôs núcleo estimador da função taxa de falha, em presença de dados censurados à direita, apresentando também método de validação cruzada para escolha da janela ótima.

Uma alternativa para os últimos métodos mencionados é usar o método estabilizado de validação cruzada. Precisamos para isso encontrar resultados referentes ao erro quadrático médio integrado. É um problema interessante para ser abordado no futuro.

# Capítulo 3

## Metodologia

Apresenta-se abaixo a metodologia da pesquisa, assim como alguns requisitos para sua compreensão.

### 3.1 Definições Básicas e Alguns Resultados

No decorrer deste trabalho, utilizaremos diversos resultados e definições da literatura, os quais transcrevemos abaixo:

#### Definição 3.1.1 Função Característica

Seja  $X$  uma variável aleatória, sua função característica é a função  $\varphi : \mathfrak{R} \rightarrow \mathbb{C}$  definida por  $\varphi(\lambda) = \varphi_X(\lambda) = E[e^{i\lambda X}]$ , onde se define  $E[e^{i\lambda X}] = E[\cos(\lambda X)] + iE[\text{sen}(\lambda X)]$ ,  $\lambda \in \mathfrak{R}$ .

#### Resultado 3.1.1 Teorema da Inversão

Se  $\varphi \in L^1(-\infty, \infty)$ , então  $F$  é continuamente diferenciável e tem-se que:

$$F'(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ixt} \varphi(\lambda) d\lambda$$

Este é o Teorema 6.2.3, em Chung [9], pag. 155.

### Resultado 3.1.2 Identidade de Parseval

Seja  $\varphi$  a função característica de  $f$ . Então:

$$\int (f(x))^2 dx = \frac{1}{2\pi} \int |\varphi(\lambda)|^2 d\lambda$$

Este é o Teorema 9.13, em Rudin [35], pag. 187.

### Definição 3.1.2 Função Característica Empírica

Seja  $X_1, \dots, X_n$  uma amostra aleatória com função de distribuição  $F$  e seja  $\varphi$  a função característica de  $F$ . Define-se a função característica empírica,  $\hat{\varphi}$ , avaliada em  $\lambda$ , por  $\hat{\varphi}(\lambda) = \frac{1}{n} \sum_{i=1}^n e^{i\lambda X_j}$ . A partir dessa definição temos que:

$$|\hat{\varphi}(\lambda)|^2 = \left[ \frac{\sum_j \cos(\lambda X_j)}{n} \right]^2 + \left[ \frac{\sum_j \text{sen}(\lambda X_j)}{n} \right]^2$$

### Definição 3.1.3 Convergência Quase Certa

Seja  $\{Y_n : n \geq 1\}$  uma seqüência de variáveis aleatórias. Dizemos que  $Y_n$  converge para  $Y$  quase certamente, se  $P(Y_n \rightarrow Y \text{ quando } n \rightarrow \infty) = 1$ .

### Resultado 3.1.3 Lema de Borel-Cantelli

Sejam  $A_1, A_2, \dots$  eventos em um espaço de probabilidade  $(\Omega, \mathcal{A}, P)$ , isto é,  $A_n \in \mathcal{A} \forall n$ .

(a) Se  $\sum_{n=1}^{\infty} P(A_n) < \infty$ , então  $P(A_n \text{ infinitas vezes}) = 0$ .

(b) Se  $\sum_{n=1}^{\infty} P(A_n) = \infty$  e os  $A_n$  são independentes, então  $P(A_n \text{ infinitas vezes}) = 1$ .

Esta é a proposição 5.2, em Barry James [22], pag.200.

### Resultado 3.1.4

Seja  $X_1, X_2, \dots$  uma seqüência de variáveis aleatórias tais que para  $\forall \epsilon > 0$ ,  $\sum_{n=1}^{\infty} P(X_n > \epsilon) < \infty$ , então  $X_n \rightarrow 0$  quase certamente.

Este é o Teorema 4.2.2, em Chung [9], pag. 73.

### Resultado 3.1.5 Teorema de Glick

Se para quase todo  $x$  real,  $f_n(x)$  é uma seqüência de estimativas de densidade de  $f(x)$  convergindo quase certamente para  $f(x)$  em probabilidade (ou quase certamente), então  $\int |f_n - f| \rightarrow 0$ , em probabilidade (ou quase certamente), quando  $n \rightarrow \infty$ .

Este é o Teorema 2.6, em Devroye [11], pag. 25.

O Teorema de Glick implica que, para todo  $x$ ,  $\int_{-\infty}^x f_n(t) dt \rightarrow \int_{-\infty}^x f(t) dt$ , como pode ser visto da seguinte forma:

Para todo subconjunto  $A$  de  $\mathfrak{R}$ , tem-se que:

$$\begin{aligned} \int_A |f_n(t) - f(t)| dt &\leq \int_{-\infty}^{\infty} |f_n(t) - f(t)| dt \\ \int_A |f_n(t)| dt - \int_A |f(t)| dt &\leq \int_A |f_n(t) - f(t)| dt \rightarrow 0 \end{aligned}$$

Então:  $\int_A |f_n(t)| dt \rightarrow \int_A |f(t)| dt$ , e, para  $A = (-\infty, x]$ , tem-se que:

$$\int_{-\infty}^x f_n(t) dt \rightarrow \int_{-\infty}^x f(t) dt$$

## 3.2 Estimativa pelo núcleo estimador da função de distribuição

Dada a amostra  $X_1, X_2, \dots, X_n$  de uma variável aleatória contínua, o núcleo estimador da função de distribuição  $F$ , duas vezes continuamente diferenciável, avaliada no ponto  $x$ , é definida por:

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^n W\left(\frac{x - X_i}{h_n}\right) \quad (3.1)$$

Assumiremos que:

1. A função de densidade  $w = W'$  é limitada, simétrica e tem suporte compacto, sendo que:  $\int w(t)dt = 1$ ,  $\int tw(t)dt = 0$  e  $0 < \int t^2w(t)dt = k_2 < \infty$ .
2. Tem-se que  $h_n \rightarrow 0$  e  $nh_n \rightarrow \infty$ , quando  $n \rightarrow \infty$ .

A partir daqui, para simplificação, escreveremos  $h$  em lugar de  $h_n$  e, quando não houver indicação dos limites de integração, assume-se que a integral é sobre toda a reta.

### 3.2.1 Propriedades de $\hat{F}_n$

As propriedades teóricas de  $\hat{F}_n$  como um estimador da função de distribuição desconhecida  $F$  têm sido investigadas por vários autores, sendo que no decorrer deste trabalho foram provadas as proposições 3.1 e 3.2 a seguir:

**Proposição 3.1** *A esperança de  $\hat{F}_n$  é dada por:  $E[\hat{F}_n] = F(x) + h^2C_2 + o(h^2)$*

*Prova*

$$\begin{aligned} E[\hat{F}_n] &= E\left[\frac{1}{n}\sum_{i=1}^n W\left(\frac{x-X_i}{h}\right)\right] \\ &= E\left[W\left(\frac{x-X_i}{h}\right)\right] \\ &= \int W\left(\frac{x-u}{h}\right)f(u)du \\ &= \frac{1}{h}\int w\left(\frac{x-u}{h}\right)F(u)du \end{aligned}$$

Mudando-se para a variável  $z = \frac{x-u}{h}$ , tem-se:

$$E[\hat{F}_n] = \int w(z)F(x-zh)dz \tag{3.2}$$

Aplicando a expansão de Taylor, temos:

$$F(x-zh) = F(x) - zhF'(x) + \frac{z^2h^2}{2}F''(x) + o(h^2)$$

A expressão (3.2) pode ser transformada em:

$$\begin{aligned} E[\hat{F}_n] &= \int w(z)F(x)dz - \int w(z)zhF'(x)dz + \int w(z)\frac{z^2h^2}{2}F''(x)dz + o(h^2) \\ &= F(x)\int w(z)dz - hF'(x)\int zw(z)dz + \frac{h^2}{2}F''(x)\int z^2w(z)dz + o(h^2) \end{aligned}$$

Desde que  $w$  é simétrica e tem suporte compacto, tem-se que:

$$\begin{aligned} E[\hat{F}_n] &= F(x) + \frac{h^2}{2}F''(x)k_2 + o(h^2) \\ &= F(x) + h^2C_2 + o(h^2) \quad \blacksquare \end{aligned} \quad (3.3)$$

Nota-se que o vício na estimação de  $F$  não depende diretamente do tamanho da amostra, mas sim do valor da janela  $h$ .

**Proposição 3.2** *A variância de  $\hat{F}_n$  é dada por:  $Var(\hat{F}_n) = \frac{1}{n}F(x)[1-F(x)] - \frac{h}{n}C_1 + o(h/n)$*

*Prova*

$$\begin{aligned} Var(\hat{F}_n) &= \frac{1}{n^2} \sum_{i=1}^n Var\left[W\left(\frac{x-X_i}{h}\right)\right] \\ &= \frac{1}{n} Var\left[W\left(\frac{x-X_1}{h}\right)\right] \\ &= \frac{1}{n} \left\{ E\left[W^2\left(\frac{x-X_1}{h}\right)\right] - E\left[W\left(\frac{x-X_1}{h}\right)\right]^2 \right\} \\ &= \frac{1}{n} \left\{ I - E\left[W\left(\frac{x-X_1}{h}\right)\right]^2 \right\} \end{aligned} \quad (3.4)$$

A primeira igualdade por serem  $X_1, X_2, \dots$  independentes e a segunda por serem identicamente distribuídas.

Sendo:

$$I = E\left[W^2\left(\frac{x-X_1}{h}\right)\right],$$

integrando por partes, temos que:

$$\begin{aligned} I &= \int W^2\left(\frac{x-u}{h}\right) f(u) du \\ &= \frac{2}{h} \int W\left(\frac{x-u}{h}\right) w\left(\frac{x-u}{h}\right) F(u) du \end{aligned}$$

Mudando-se para a variável  $z = \frac{x-u}{h}$ , tem-se:

$$I = 2 \int W(z) w(z) F(x-zh) dz$$

Aplicando novamente a expansão de Taylor, temos  $F(x - zh) = F(x) - zhF'(x) + o(h)$  e então:

$$\begin{aligned} I &= 2 \left\{ \int W(z) w(z) F(x) dz - \int W(z) w(z) z h F'(x) dz + o(h) \right\} \\ &= 2 \left\{ F(x) \int W(z) W'(z) dz - h F'(x) \int z W(z) W'(z) dz + o(h) \right\} \end{aligned}$$

Como  $\int W(z) W'(z) dz = 1/2$ , tem-se que:

$$I = F(x) - 2h F'(x) \int z W(z) w(z) dz + o(h)$$

Assim, das expressões (3.3) e (3.4), tem-se que:

$$\begin{aligned} Var [\hat{F}_n] &= \frac{1}{n} \left\{ F(x) - 2h F'(x) \int z W(z) w(z) dz + o(h) - \left[ F(x) + \frac{h^2}{2} F''(x) k_2 + o(h^2) \right]^2 \right\} \\ &= \frac{1}{n} \{ F(x)[1 - F(x)] \} - 2 \frac{h}{n} F'(x) \int z W(z) w(z) dz - \frac{h^2}{n} F(x) F''(x) \int z^2 w(z) dz + \\ &\quad - \frac{h^4}{4n} \left[ F''(x) \int z^2 w(z) dz \right]^2 + o(h/n) \\ &= \frac{1}{n} F(x)[1 - F(x)] - \frac{h}{n} C_1 + o(h/n) \end{aligned} \quad (3.5)$$

onde  $C_1 = 2 F'(x) \int z W(z) w(z) dz > 0$  ■

A variância da função de distribuição empírica,  $\tilde{F}_n = \frac{1}{n} \# \{i : X_i \leq x\}$  é dada por:

$$\frac{F(x)\{1 - F(x)\}}{n},$$

onde  $n$  é o tamanho da amostra. Verifica-se então que o núcleo estimador reduz esta variância no termo da forma  $-C_1 h/n$ , com  $C_1 > 0$ , salientando que  $Var(\hat{F}_n) - Var(\tilde{F}_n) \rightarrow 0$ , pois  $C_1 \frac{h}{n} \rightarrow 0$  quando  $n \rightarrow \infty$ . Prova-se que  $\int z W(z) w(z) dz > 0$ . Dessa maneira,  $Var [\hat{F}_n] \leq Var [\tilde{F}_n]$ .

Entretanto,  $\hat{F}_n$  está sujeito a um vício cujo termo principal é da forma  $h^2 C_2$ , conforme é verificado na expressão (3.3). As constantes  $C_1$  e  $C_2$  dependem do núcleo  $W$  e da função de distribuição desconhecida  $F$ . Salientamos assim que a escolha da janela ótima visa um balanceamento entre o vício e a variância de  $\hat{F}_n$ .

Verifica-se pelas expressões (3.3) e (3.5) que o ajuste no valor do parâmetro de suavidade pode reduzir o vício do estimador, impondo entretanto um acréscimo em sua variância. Se é utilizado um valor pequeno de  $h$  para eliminar o vício, então a variância torna-se grande. De outra maneira, a escolha de um valor grande de  $h$  reduzirá a variação aleatória, quantificada pela variância, em troca da introdução de um erro sistemático, ou vício na estimativa. Assim, a escolha do parâmetro de suavidade implica um balanceamento entre erro aleatório e erro sistemático, sendo este um dos problemas fundamentais em estimação.

### 3.2.2 Escolha da janela ótima

Como apontado por diversos autores, a escolha do núcleo  $W$  não é muito crucial, mas a escolha do parâmetro de suavidade é um sério problema que tem sido tratado exaustivamente na literatura.

A medida mais usada do desempenho global do estimador é o erro quadrático integrado  $ISE$ , que é definido como :  $ISE(h) = \int \{ \hat{F}_n(x) - F(x) \}^2 dx$ .

Um valor ótimo do parâmetro de suavidade ( $h_{opt}$ ) pode ser definido como o valor que minimiza o  $MISE(h)$ . Veja expressão (2.1).

Esse valor é desconhecido na prática, mas demonstraremos que  $h_{opt}$  tem a forma assintótica de  $C_3 n^{-1/3}$  e encontraremos  $C_3$ .

Percebe-se aqui uma diferença importante com o caso da estimação de densidade, onde o parâmetro de suavidade tem a forma assintótica de  $C_4 n^{-1/5}$ .

Retomando a expressão (2.1) e assumindo as hipóteses do modelo, expressas anteriormente, tem-se que:

$$\begin{aligned} MISE(h) &= \int E \{ \hat{F}_n(x) - F(x) \}^2 dx \\ &= \int \{ var [\hat{F}_n(x)] + Vicio^2 [\hat{F}_n(x)] \} dx \end{aligned}$$

Dadas as expressões (3.3) e (3.5) e dados que:  $C_1 = 2F'(x) \int z W(z) w(z) dz$ ,

$C_2 = \frac{F''(x)k_2}{2}$  e  $k_2 = \int z^2 w(z) dz$ , tem-se que:

$$\begin{aligned}
MISE(h) &= \int \left\{ \frac{1}{n} F(x)[1 - F(x)] - \frac{h}{n} C_1 + o(h/n) + [F(x) + h^2 C_2 + o(h^2) - F(x)]^2 \right\} dx \\
&= \frac{1}{n} \int F(x)[1 - F(x)] dx - \frac{h}{n} \int 2F'(x) \left[ \int zW(z)w(z) dz \right] dx + \\
&\quad + \int \left[ \frac{h^2 F''(x) k_2}{2} \right]^2 dx + \int \left\{ [o(h^2)]^2 + o(h/n) \right\} dx \\
&= n^{-1} \int F(x)[1 - F(x)] dx - h n^{-1} \int 2zW(z)w(z) dz + \frac{h^4 k_2^2}{2} \int [F''(x)]^2 dx + \\
&\quad + o(h n^{-1} + h^4) \tag{3.6}
\end{aligned}$$

**Proposição 3.3** *Dado que, por hipótese,  $W$  tem suporte compacto, tem-se que:*

$$2 \int_{-\alpha}^{\alpha} z W(z) w(z) dz = \int_{-\alpha}^{\alpha} W(z)[1 - W(z)] dz$$

*Prova*

Dado que, por hipótese,  $W$  tem suporte compacto e é simétrica, então existe  $\alpha$ ,  $0 < \alpha < \infty$ , tal que  $W(-\alpha) = 0$  e  $W(\alpha) = 1$ , sendo que  $W(z) = 0$ ,  $z < -\alpha$  e  $W(z) = 1$ ,  $z > \alpha$ .

Seja  $I = \int_{-\alpha}^{\alpha} z W(z) w(z) dz$ , então:

$$I = \int_{-\alpha}^0 z W(z) w(z) dz + \int_0^{\alpha} z W(z) w(z) dz = I_1 + I_2 \quad , \text{ onde :}$$

$$I_1 = \int_0^{\alpha} z W(z) w(z) dz = \frac{\alpha}{2} - \int_0^{\alpha} \frac{1}{2} W^2(z) dz \quad , \text{ e } \quad I_2 = \int_0^{\alpha} z W(z) w(z) dz$$

Fazendo uma mudança para a variável  $t = -z$ , obtém-se:

$$I_2 = \int_{\alpha}^0 (-t) W(-t) w(-t) (-dt),$$

como  $W$  é simétrica, tem-se que  $w(t) = w(-t)$  e  $W(-t) = 1 - W(t)$ .

Assim:

$$I_2 = - \int_0^{\alpha} t w(t) [1 - W(t)] dt$$

$$\begin{aligned}
&= -\alpha + \int_0^\alpha W(t) dt + \int_0^\alpha t w(t) W(t) dt, \text{ e} \\
I &= \int_0^\alpha W(t) [1 - W(t)] dt
\end{aligned}$$

$W(t) [1 - W(t)]$  é simétrica, então:

$$I = \frac{1}{2} \int_{-\alpha}^\alpha W(t) [1 - W(t)] dt \blacksquare$$

Assim, uma forma alternativa para o erro quadrático médio integrado será dado por:

$$\begin{aligned}
MISE(h) &= n^{-1} \int F(x) [1 - F(x)] dx - h n^{-1} \int W(z) [1 - W(z)] dz + \frac{h^4 k_2^2}{4} \int [F''(x)]^2 dx \\
&\quad + o(h n^{-1} + h^4)
\end{aligned}$$

Sendo da forma :

$$MISE(h) = n^{-1} \int F(x) [1 - F(x)] dx - C_1 h n^{-1} + C_2 h^4 + o(h n^{-1} + h^4). \quad (3.7)$$

A janela ótima  $h_{opt}$  será dada por  $\frac{dMISE(h)}{dh} = 0$ , para  $h = h_{opt}$ . Assim derivando (3.7) com respeito a  $h$ , obtemos:

$$\frac{dMISE(h)}{dh} = -C_1 n^{-1} + 4 C_2 h^3$$

Os pontos críticos serão dados por  $h_{crit}$ , tal que:

$$h_{crit} = \left[ \frac{C_1}{4 C_2} \right]^{1/3} n^{-1/3}$$

Entretanto,  $\frac{d^2 MISE(h)}{dh^2} = 12 C_2 h^2 > 0$ , para todo  $h$ , já que  $C_2 > 0$ . Logo  $h_{crit}$  é ponto de mínimo global do erro quadrático médio integrado e  $h_{opt} = C_3 n^{-1/3}$ , sendo que  $C_3 = \left[ \frac{C_1}{4 C_2} \right]^{1/3}$ , e a expressão da janela ótima será dada por:

$$h_{opt} = \left\{ \frac{\int W(x) [1 - W(x)] dx}{\left[ \int z^2 dW(z) \right]^2 \int [F''(x)]^2 dx} \right\}^{1/3} n^{-1/3} \quad (3.8)$$

### 3.2.3 Método “plug-in”

O núcleo estimador da função de distribuição  $F(x)$  foi definido como:

$$\hat{F}_n = \frac{1}{n} \sum_{i=1}^n W\left(\frac{x - X_i}{h}\right),$$

onde  $W$  é uma função de distribuição. Por este método, o valor ótimo de  $h$  é dado pela expressão (3.8), sendo da forma  $h_{opt} = C_3 n^{-1/3}$ , verificando-se que a parcela  $H = \int [F''(x)]^2 dx$  da expressão de  $h_{opt}$  depende da função  $F$  desconhecida que se quer estimar. O método “plug-in” baseia-se em estimar  $H$ .

Seguindo as idéias de Chiu [8], para estimar o valor de  $H$  usamos a função característica empírica da amostra. Para isso, apresentaremos algumas definições e resultados básicos da literatura :

- Assumimos que a função característica (Definição 3.1.1) satisfaz a condição de que existem constantes positivas  $M_1, M_2, K_1$  e  $K_2$ , tais que  $M_1|\lambda|^{-K_1} \geq |\varphi(\lambda)|^2 \geq M_2|\lambda|^{-K_2}$  quando  $|\lambda| \rightarrow \infty$  e que  $|\varphi(\lambda)|^2 > 0$ , para todo  $\lambda$ .
- Pela Fórmula da Inversão (Resultado 3.1.1), tem-se que:

$$f(x) = F'(x) = \frac{1}{2\pi} \int e^{-i\lambda x} \varphi(\lambda) d\lambda$$

logo:

$$\begin{aligned} f'(x) = F''(x) &= \frac{1}{2\pi} \int (-i\lambda) e^{-i\lambda x} \varphi(\lambda) d\lambda \\ &= \frac{1}{2\pi} \int e^{-i\lambda x} [-i\lambda \varphi(\lambda)] d\lambda \end{aligned}$$

- Da Identidade de Parseval (Resultado 3.1.2), tem-se que:

$$\int (f(x))^2 dx = \frac{1}{2\pi} \int |\varphi(\lambda)|^2 d\lambda$$

logo:

$$H = \int (F''(x))^2 dx$$

$$\begin{aligned}
&= \int (f'(x))^2 dx \\
&= \frac{1}{2\pi} \int \lambda^2 |\varphi(\lambda)|^2 d\lambda
\end{aligned}$$

Dada a função característica empírica,  $\hat{\varphi}(\lambda)$ , (Definição 3.1.1), então:

$$|\hat{\varphi}(\lambda)|^2 = \left[ \frac{\sum_j \cos(\lambda X_j)}{n} \right]^2 + \left[ \frac{\sum_j \text{sen}(\lambda X_j)}{n} \right]^2$$

Podendo-se provar que:

$$E \left[ |\hat{\varphi}(\lambda)|^2 \right] \cong \frac{1}{n} \left[ (n-1) |\varphi(\lambda)|^2 + 1 \right]$$

Usando este resultado, pode-se aproximar  $H$  por:

$$\hat{H} = \frac{1}{\pi} \int_0^\Lambda \lambda^2 \left[ |\hat{\varphi}(\lambda)|^2 - \frac{1}{n} \right] d\lambda \quad (3.9)$$

onde  $\Lambda = \min \left\{ \lambda : |\hat{\varphi}(\lambda)|^2 \leq \frac{C}{n} \right\}$ , para algum  $C > 1$ . O termo  $1/n$  é utilizado para reduzir a variação de  $|\hat{\varphi}(\lambda)|^2$  para  $\lambda > \Lambda$ , identificada a parte que contém a maioria da informação de  $f(x)$ .

Substituímos  $\hat{H}$  em lugar de  $H$  na expressão (3.8) e obtemos um estimador  $\hat{h}_{opt}$ , para  $h_{opt}$ .

### 3.2.4 Convergência Quase Certa de $\hat{H}$

Conforme Chiu (1991) [8], as condições de  $f(x)$  e  $w(x)$  estão resumidas nas hipóteses de 1 a 3. Já que o procedimento de escolha da janela está baseado em termos de função característica, estas são expressas em termos de  $\varphi(\lambda)$  e  $W(\lambda)$ :

1. Existem constantes positivas  $M_1$ ,  $M_2$ ,  $K_1$  e  $K_2$ , tais que  $M_1 |\lambda|^{-K_1} \geq |\varphi(\lambda)|^2 \geq M_2 |\lambda|^{-K_2}$  quando  $|\lambda| \rightarrow \infty$ . Assume-se também que  $|\varphi(\lambda)|^2 > 0$ , para todo  $\lambda$

2. A densidade  $f(x)$  tem derivada uniformemente limitada e satisfaz a relação  $\int_{|x|>M} f(x) dx \leq O(M^{-1})$ , quando  $M \rightarrow \infty$ .
3. O núcleo  $w(x)$  é uma função de densidade simétrica e satisfaz a relação  $\int |x|^3 w(x) dx < \infty$ . A função característica de  $w(x)$ , dada por  $\varphi_W(\lambda) = \int e^{i\lambda x} w(x) dx$ , satisfaz as seguintes condições:  $\varphi_W(\lambda) = O(\lambda^{-3})$  e  $\varphi'_W(\lambda) = O(\lambda^{-3})$ , quando  $\lambda \rightarrow \infty$ .

Considerando as hipóteses listadas acima, transcrevemos os seguintes resultados de Chiu ([8]):

**Resultado 3.2.1** *Considerando a hipótese 1 acima, então, para qualquer  $\delta > 0$ ,  $\Lambda \leq n^{1/K_1+\delta}$ , com probabilidade tendendo a 1 .*

Este é o Lema 2, em Chiu [8], pag. 1900.

**Resultado 3.2.2** *Considerando as hipóteses 1 e 3 acima, então, para qualquer  $\delta > 0$ ,  $n^{1/K_2-\delta} \leq \Lambda$ , com probabilidade tendendo a 1 .*

*Este resultado implica que  $\max |\hat{\varphi}_d(\lambda)| \leq Mn^{-1/2+\delta}$  , para  $0 \leq \lambda \leq n$*

Este é o Lema 3, em Chiu [8], pag. 1901.

O seguinte resultado foi provado no desenvolvimento de nossa pesquisa:

### **Proposição 3.4** *Convergência Quase Certa de $\hat{H}$*

*Sob as condições das hipóteses 1 e 2, com  $K_1 > 5$ ,  $\hat{H}$  converge quase certamente para  $H$ , quando  $n \rightarrow \infty$ .*

*Prova:*

Dado que:

$$H = \int [F''(x)]^2 dx = \frac{1}{2\pi} \int \lambda^2 |\varphi(\lambda)|^2 d\lambda$$

e

$$\hat{H} = \frac{1}{\pi} \int_0^\Lambda \lambda^2 \left[ |\hat{\varphi}(\lambda)|^2 - \frac{1}{n} \right] d\lambda$$

então:

$$\begin{aligned} \pi \hat{H} - \pi H &= \int_0^\Lambda \lambda^2 \left[ |\hat{\varphi}(\lambda)|^2 - \frac{1}{n} \right] d\lambda - \frac{1}{2} \int \lambda^2 |\varphi(\lambda)|^2 d\lambda \\ &= \int_0^\Lambda \lambda^2 \left[ |\hat{\varphi}(\lambda)|^2 - \frac{1}{n} \right] d\lambda - \int_0^\infty \lambda^2 |\varphi(\lambda)|^2 d\lambda \\ &= \int_0^\Lambda \lambda^2 \left[ |\hat{\varphi}(\lambda)|^2 - \frac{1}{n} \right] d\lambda - \int_0^\Lambda \lambda^2 |\varphi(\lambda)|^2 d\lambda + B_1 \end{aligned} \quad (3.10)$$

com

$$B_1 = - \int_\Lambda^\infty \lambda^2 |\varphi(\lambda)|^2 d\lambda \quad (3.11)$$

Definindo-se  $\hat{\varphi}_d(\lambda) = \hat{\varphi}(\lambda) - \varphi(\lambda)$ , então :

$$\begin{aligned} |\hat{\varphi}(\lambda)|^2 &= |\hat{\varphi}(\lambda) - \varphi(\lambda) + \varphi(\lambda)|^2 \\ &= |\hat{\varphi}_d(\lambda) + \varphi(\lambda)|^2 \end{aligned}$$

Usando propriedades de números complexos, temos que:

$$|\hat{\varphi}(\lambda)|^2 = |\varphi(\lambda)|^2 + |\hat{\varphi}_d(\lambda)|^2 + 2 \operatorname{Re} [\varphi(\lambda) \hat{\varphi}_d(-\lambda)]$$

Assim a expressão (3.10) torna-se:

$$\begin{aligned} \pi \hat{H} - \pi H &= \int_0^\Lambda \lambda^2 \left[ |\hat{\varphi}(\lambda)|^2 - |\varphi(\lambda)|^2 - \frac{1}{n} \right] d\lambda + B_1 \\ &= \int_0^\Lambda \lambda^2 \left[ |\hat{\varphi}_d(\lambda)|^2 + 2 \operatorname{Re} \varphi(\lambda) \hat{\varphi}_d(-\lambda) - \frac{1}{n} \right] d\lambda + B_1 \\ &= \int_0^\Lambda \lambda^2 \left[ |\hat{\varphi}_d(\lambda)|^2 - \frac{1}{n} \right] d\lambda + 2 \operatorname{Re} \int_0^\Lambda \lambda^2 \varphi(\lambda) \hat{\varphi}_d(-\lambda) d\lambda + B_1 \end{aligned} \quad (3.12)$$

Logo a expressão (3.12) é da forma:

$$\pi \hat{H} - \pi H = B_1 + B_2 + B_3 \quad (3.13)$$

com:

$$B_1 = - \int_{\Lambda}^{\infty} \lambda^2 |\varphi(\lambda)|^2 d\lambda \quad (3.14)$$

$$B_2 = 2 \operatorname{Re} \int_0^{\Lambda} \lambda^2 \varphi(\lambda) \hat{\varphi}_d(-\lambda) d\lambda \quad (3.15)$$

$$B_3 = \int_0^{\Lambda} \lambda^2 \left[ |\hat{\varphi}_d(\lambda)|^2 - \frac{1}{n} \right] d\lambda \quad (3.16)$$

### Convergência de $B_1$

Pelo Resultado 3.2.2 tem-se que  $P\left(\Lambda \geq n^{1/K_2-\delta}\right) \rightarrow 1$ , quando  $n \rightarrow \infty$ .

Dado que  $\int_{\Lambda}^{\infty} \lambda^2 |\varphi(\lambda)|^2 d\lambda = g_1(\Lambda)$  é uma função contínua e decrescente em  $\Lambda$ , tem-se que  $P\left[g_1(\Lambda) \leq g_1(n^{1/K_2-\delta})\right] \rightarrow 1$  quase certamente, implicando:

$$\int_{\Lambda}^{\infty} \lambda^2 |\varphi(\lambda)|^2 d\lambda \leq \int_{n^{1/K_2-\delta}}^{\infty} \lambda^2 |\varphi(\lambda)|^2 d\lambda \quad (3.17)$$

Pela hipótese 1, tem-se que:

$$\begin{aligned} M_1 |\lambda|^{-K_1} &\geq |\varphi(\lambda)|^2 \geq M_2 |\lambda|^{-K_2} \\ |\varphi(\lambda)|^2 &\leq M_1 |\lambda|^{-K_1} \\ \lambda^2 |\varphi(\lambda)|^2 &\leq M_1 |\lambda|^{-K_1+2}, \text{ para } \lambda \rightarrow \infty \end{aligned}$$

Assim, da expressão (3.17), tem-se:

$$\begin{aligned} \int_{n^{1/K_2-\delta}}^{\infty} \lambda^2 |\varphi(\lambda)|^2 d\lambda &\leq \int_{n^{1/K_2-\delta}}^{\infty} M_1 \lambda^{-K_1+2} d\lambda \\ &\leq \frac{M_1}{-3+K_1} n^{(1/K_2-\delta)(-K_1+3)} \end{aligned}$$

Considerando  $K_1 > 3$ , escolhe-se  $\delta$  tal que  $(1/K_2 - \delta)(-K_1 + 3) < 0$ .

Sendo  $A_n = \left| \int_{\Lambda}^{\infty} \lambda^2 |\varphi(\lambda)|^2 d\lambda \right|$  quer-se provar que  $A_n \rightarrow 0$  quase certamente, tendo-se que  $|A_n| \leq \frac{M_1}{k_1-3} n^{\alpha}$ , com  $\alpha < 0$ . Para  $\forall \epsilon > 0$ , existe  $n_0$  tal que  $n^{\alpha} < \epsilon, \forall n > n_0$ .

Teremos então:

$$\sum_{n=1}^{\infty} P(A_n > \epsilon) = \sum_{n=1}^{n_0} P(A_n > \epsilon) + \sum_{n=n_0+1}^{\infty} P(A_n > \epsilon)$$

Mas  $\sum_{n=n_0+1}^{\infty} P(A_n > \epsilon) = 0$ , então:

$$\sum_{n=1}^{\infty} P(A_n > \epsilon) = \sum_{n=1}^{n_0} P(A_n > \epsilon) \leq n_0 < \infty$$

Assim, usando o Lema de Borel-Cantelli 3.1.3, conclui-se que  $B_1 \rightarrow 0$  quase certamente.

### *Convergência de $B_2$*

Dado  $B_2$  (3.15), tem-se que:

$$\begin{aligned} |B_2| &= 2 \operatorname{Re} \int_0^{\Lambda} \lambda^2 \varphi(\lambda) \hat{\varphi}_d(-\lambda) d\lambda \leq 2 \left| \int_0^{\Lambda} \lambda^2 \varphi(\lambda) \hat{\varphi}_d(-\lambda) d\lambda \right| \leq \int_0^{\Lambda} \lambda^2 |\varphi(\lambda) \hat{\varphi}_d(-\lambda)| d\lambda \leq \\ &\int_0^{\Lambda} \lambda^2 |\varphi(\lambda)| |\hat{\varphi}_d(-\lambda)| d\lambda \end{aligned}$$

Sabe-se que  $|\varphi(\lambda)| \leq 1$  e do Resultado 3.2.2 tem-se que  $|\hat{\varphi}_d(-\lambda)| \leq \max |\hat{\varphi}_d(\lambda)| \leq Mn^{-1/2+\delta}$ , para  $0 \leq \lambda \leq n$ , assim:

$$2 \operatorname{Re} \int_0^{\Lambda} \lambda^2 \varphi(\lambda) \hat{\varphi}_d(-\lambda) d\lambda \leq 2 \int_0^{\Lambda} \lambda^2 Mn^{-1/2+\delta} d\lambda$$

Pelo Resultado 3.2.1, para qualquer  $\delta > 0$ , tem-se que  $P(\Lambda \leq n^{1/K_1+\delta}) \rightarrow 1$ , quando  $n \rightarrow \infty$ .

Dado que  $2 \int_0^{\Lambda} \lambda^2 Mn^{-1/2+\delta} d\lambda = g_2(\Lambda)$  é uma função contínua e crescente em  $\Lambda$ , então  $P[g_2(\Lambda) \leq g_2(n^{1/K_1+\delta})] \rightarrow 1$  quando  $n \rightarrow \infty$ , implicando:

$$\begin{aligned} \int_0^{\Lambda} \lambda^2 Mn^{-1/2+\delta} d\lambda &\leq \int_0^{n^{1/K_1+\delta}} \lambda^2 Mn^{-1/2+\delta} d\lambda = \frac{1}{3} Mn^{3/K_1-1/2+4\delta} = \\ &= O(n^{3/K_1-1/2+4\delta}) \end{aligned}$$

Sendo  $A_n = 2 \operatorname{Re} \int_0^{\Lambda} \lambda^2 \varphi(\lambda) \hat{\varphi}_d(-\lambda) d\lambda$ , considerando  $K_1 > 6$  e escolhendo-se  $\delta$  de maneira que  $(3/K_1 - 1/2 + 4\delta) < 0$ , tem-se que  $A_n \leq n^\alpha$ , com  $\alpha < 0$ . Assim, aplicando

o Lema de Borel-Cantelli de maneira similar ao efetuado no caso da convergência de  $B_1$ , prova-se que  $B_2 \rightarrow 0$  quase certamente.

### Convergência de $B_3$

Dado  $B_3$  (3.16), tem-se que:

$$\begin{aligned} |B_3| &= \left| \int_0^\Lambda \lambda^2 \left[ |\hat{\varphi}_d(\lambda)|^2 - \frac{1}{n} \right] d\lambda \right| \leq \int_0^\Lambda \lambda^2 \left| \left[ |\hat{\varphi}_d(\lambda)|^2 - \frac{1}{n} \right] \right| d\lambda \leq \\ &\int_0^\Lambda \lambda^2 \left| \left[ |\hat{\varphi}_d(\lambda)|^2 + \frac{1}{n} \right] \right| d\lambda \leq \int_0^\Lambda \lambda^2 |\hat{\varphi}_d(\lambda)|^2 d\lambda + \int_0^\Lambda \lambda^2 \frac{1}{n} d\lambda \end{aligned}$$

Pelo Resultado 3.2.1, para qualquer  $\delta > 0$ , tem-se que  $P(\Lambda \leq n^{1/K_1+\delta}) \rightarrow 1$ , quando  $n \rightarrow \infty$ .

Dado que  $\int_0^\Lambda \lambda^2 |\hat{\varphi}_d(\lambda)|^2 d\lambda = g_3(\lambda)$  e  $\int_0^\Lambda \lambda^2 \frac{1}{n} d\lambda = g_4(\lambda)$  são funções contínuas e crescentes em  $\Lambda$ , então  $P[g_3(\Lambda) \leq g_3(n^{1/K_1+\delta})] \rightarrow 1$  e  $P[g_4(\Lambda) \leq g_4(n^{1/K_1+\delta})] \rightarrow 1$  quando  $n \rightarrow \infty$ , implicando:

(i)

$$\int_0^\Lambda \lambda^2 \frac{1}{n} d\lambda \leq \int_0^{n^{1/K_1+\delta}} \frac{\lambda^2}{n} d\lambda = \frac{n^{3/K_1+3\delta-1}}{3}$$

(ii)

$$\int_0^\Lambda \lambda^2 |\hat{\varphi}_d(\lambda)|^2 d\lambda \leq \int_0^{n^{1/K_1+\delta}} \lambda^2 |\hat{\varphi}_d(\lambda)|^2 d\lambda$$

Entretanto, do Resultado 3.2.2 tem-se que  $\max |\hat{\varphi}_d(\lambda)| \leq Mn^{-1/2+\delta} \Rightarrow \max |\hat{\varphi}_d(\lambda)|^2 \leq M^2 n^{-1+2\delta}$ , para  $0 \leq \lambda \leq n$ , assim:

$$\begin{aligned} \int_0^{n^{1/K_1+\delta}} \lambda^2 |\hat{\varphi}_d(\lambda)|^2 d\lambda &\leq \int_0^{n^{1/K_1+\delta}} \lambda^2 \max |\hat{\varphi}_d(\lambda)|^2 d\lambda \leq \int_0^{n^{1/K_1+\delta}} \lambda^2 M^2 n^{-1+2\delta} d\lambda = \\ &= \frac{M^2}{3} n^{-1+3/K_1+5\delta} \end{aligned}$$

Assim:

$$\left| \int_0^\Lambda \lambda^2 \left[ |\hat{\varphi}_d(\lambda)|^2 - \frac{1}{n} \right] d\lambda \right| \leq \frac{M^2}{3} n^{3/K_1-1+5\delta} + \frac{n^{3/K_1-1+3\delta}}{3} = O(n^{3/K_1-1+5\delta})$$

Sendo  $A_n = \left| \int_0^\Lambda \lambda^2 \left[ |\hat{\varphi}_d(\lambda)|^2 - \frac{1}{n} \right] d\lambda \right|$ , considerando  $K_1 > 3$  e escolhendo-se  $\delta$  de maneira que  $(3/K_1 - 1 + 5\delta) < 0$ , tem-se que  $A_n \leq n^\alpha$ , com  $\alpha < 0$ . Assim, aplicando o Lema de Borel-Cantelli de maneira similar ao efetuado no caso da convergência de  $B_1$ , prova-se que  $B_3 \rightarrow 0$  quase certamente.

Verificou-se que  $B_1$ ,  $B_2$  e  $B_3$  convergem quase certamente para zero. Então, dadas a definição 3.1.3 e a expressão 3.12, conclui-se que  $\hat{H}$  converge para  $H$  quase certamente.

Desde que  $\hat{h}$  é uma função contínua de  $\hat{H}$ , uma consequência imediata é o seguinte:

**Corolário 3.5 Convergência Quase Certa de  $\hat{h}_{opt}$**

*Sob as condições das hipóteses 1 e 2,  $\hat{h}_{opt}$  converge quase certamente para  $h_{opt}$ , quando  $n \rightarrow \infty$*

### 3.3 Estimativa pela integral do núcleo estimador da função de densidade

A função  $F(x)$  pode ser estimada através da estimação da função de densidade  $f$ , ou seja,  $\hat{F}_{2n}(x) = \int_{-\infty}^x \hat{f}_n(t) dt$ , onde  $f_n(t)$  é o núcleo estimador de  $f$ , a partir da amostra  $X_1, X_2, \dots, X_n$ , ou seja:

$$\hat{f}_n(x) = \frac{1}{n} \sum_{i=1}^n w \left( \frac{x - X_i}{h} \right) \tag{3.18}$$

onde  $w$  é uma função de densidade e  $h$  é o parâmetro de suavidade da função de densidade. O Teorema de Glick garante que  $\hat{F}_{2n}(x)$  assim definida converge quase certamente para  $F(x)$ .

Na literatura, prova-se que o valor ótimo de  $h$ , na definição 3.18, é dado por:

$$h_{opt} = k_2^{-2/5} \left\{ \int [w(t)]^2 dt \right\}^{1/5} \left\{ \int [f''(x)]^2 dx \right\}^{-1/5} n^{-1/5}$$

sendo  $k_2 = \int t^2 w(t) dt$ .

Veja, por exemplo, Damasceno [10] ou Silverman [39]. Damasceno, em particular, apresenta alguns métodos de estimação da janela ótima.

# Capítulo 4

## Implementação Computacional

### 4.1 Considerações Gerais

Para atender aos objetivos principais desta dissertação foi necessária a elaboração de um algoritmo em uma linguagem de programação que fosse rápida o bastante para processar simulações com uma grande quantidade de amostras. A flexibilidade e facilidade de linguagem também foi um requisito importante, pois buscávamos o controle das etapas mais importantes de todo o processo de estimação visando um estudo detalhado do comportamento do método. Outro objetivo foi facilitar a estimação funcional a partir de um único conjunto de dados. Esta característica é conveniente às aplicações práticas do método.

Nas primeiras versões do algoritmo, optou-se pela utilização do S-plus, versão 4.5, devido à sua disponibilidade e à facilidade de suas funções estatísticas e gráficas. Posteriormente, para a execução das simulações mais longas, desenvolvemos uma rotina em linguagem C, em ambiente Linux, de maneira a garantirmos um tempo de operação reduzido, assegurando-nos de uma grande capacidade de memória, haja visto a quantidade de amostras e de cálculos necessários ao desenvolvimento do trabalho. O programa em C não fez uso de nenhuma biblioteca estabelecida, tendo sido desenvolvidas todas as funções necessárias ao funcionamento do algoritmo de simulação e implementadas rotinas gráficas para

construção dos histogramas das distribuições envolvidas.

O funcionamento do algoritmo pode ser resumido nas seguintes etapas: geração das amostras da simulação; determinação do limite de integração  $\Lambda$ ; cálculo de  $\hat{H}$ ; cálculo de  $h_{opt}$ ; cálculo e gráfico de  $\hat{F}_n$ ; cálculo do  $ISE(h)$ ; cálculo da estimativa da função de distribuição a partir da estimativa da função de densidade. Para cada uma delas demos um tratamento específico, principalmente em relação aos cuidados técnicos necessários à precisão das aproximações envolvidas.

Tabela 4.1: Resumo de Algumas Funções Disponíveis em S-plus

Comando	Argumentos	Descrição
func.phi2	(vetor.dados,x)	Calcula $ \hat{\varphi}(\lambda) ^2$
alg.lambda	(vetor.dados, cota.sup=4, cota.inf=0)	Calcula $\Lambda$
alg.H	(vetor.dados, lambda)	Calcula $\hat{H}$
alg.hop	(vetor.dados, $\hat{H}$ , flag) flag(0: N. Gaussiano, 1: N.Epanechnikov)	Calcula $h_{opt}$ (para núcleo estimador de $\hat{F}_n$ )
func.F.norm	(x, $h_{opt}$ , vetor.dados)	Calcula $\hat{F}_n(x)$ (núcleo gaussiano)
func.F.ise.norm	(hop, vet, inf=-5, delta=0.01)	Calcula ISE de amostra $N(0,1)$
alg.G	(vetor.dados, lambda)	Calcula $\hat{G}$
alg.hop.f	(vetor.dados, $\hat{G}$ , flag) flag(0: N. Gaussiano, 1: N.Epanechnikov)	Calcula $h_{opt}$ (para núcleo estimador de $\hat{f}_n$ )
func.f.dens.norm	(x, $h_{opt}$ , vetor.dados)	Calcula $\hat{f}_n(x)$ - núcleo gaussiano
func.F.dens	(x, hop, vetor.dados, inf=-5, sup=-5)	Calcula $\hat{f}_n$ e $\int \hat{f}_n$
func.f.dens.ise	(hop, vet, inf=-5, delta=0.01)	Calcula ISE de amostra $N(0,1)$

O programa em C, correspondendo a todo o processo de simulação e de estimação, está apresentado no Apêndice. A biblioteca em S desenvolvida possibilita o cálculo de cada uma das etapas descritas anteriormente, podendo ser utilizada tanto para a estimação da função de distribuição quanto da função de densidade. Apresentamos o detalhe de algumas destas funções no mesmo apêndice. A tabela 4.1 apresenta um resumo de algumas das principais funções disponíveis em S.

## 4.2 Geração de Números Aleatórios

As simulações envolveram amostras com distribuição normal padrão, distribuição gama (2,1) e mistura de normais com médias -1,5 e 1,5, com o mesmo desvio padrão 1. Na implementação, em C, da função para geração de números aleatórios de distribuição normal, utilizamos o algoritmo de Box-Muller devido à sua eficiência. No caso da geração de números aleatórios com distribuição gama, utilizamos de algoritmo de método da rejeição. Uma boa referência para ambos é Press et all [32]. Na formação das amostras de mistura de normais, a distribuição de origem de cada um de seus elementos foi decidida por sorteio, através de número aleatório gerado a partir de distribuição uniforme, garantidas iguais probabilidades de escolha. Além disso, utilizamos as funções internas do S-plus, para a geração de alguns conjuntos reduzidos de amostras necessárias à pesquisa e ao desenvolvimento e validação das funções nesta linguagem.

## 4.3 Limite de Integração $\Lambda$

Preocupou-nos inicialmente a precisão na escolha do limite superior de integração  $\Lambda$ . Através de simulações, verificou-se que valores da ordem de até  $10^{-3}$  são suficientes para assegurar a precisão da integração numérica necessária à obtenção do valor de  $\hat{H}$ .

Ressalta-se que, segundo a literatura, para o caso da estimativa de  $f$ , a escolha do valor da constante  $c$  não é importante, quando a função é suficientemente suave e  $n$  grande. Adotamos o valor sugerido por Chiu para o seletor estabilizado ( $c = 3$ ) e verificamos que a exatidão dos resultados não era significativamente afetada quando se escolhiam os valores de  $c$  na faixa sugerida por Chiu, ou seja, entre  $-\log(0,15)$  e  $-\log(0,05)$ .

## 4.4 Integração Numérica

Por estar presente na maioria das etapas do processo de estimação, a integração numérica foi uma parte importante do desenvolvimento computacional deste trabalho. O método de Simpson, conforme Mathews [26], mostrou-se adequado às integrações numéricas, dadas suas características de robustez, fornecendo uma precisão razoável aos resultados. Utilizamos este processo na estimação de  $H$  e na obtenção dos valores de  $\hat{F}_n$ , a partir da estimativa de densidade  $f_n$ . Para obter-se precisão na obtenção de  $\hat{H}$ , verificamos ser suficiente a utilização de um passo de tamanho 0,1; não havendo uma alteração significativa quando utilizado um passo de 0,01. Entretanto, no processo de validação do algoritmo, verificamos que em algumas situações o passo de tamanho 0,1 poderia influenciar os resultados das estimativas de  $F$  obtidas através da  $\int f_n$ . Assim, optamos pela utilização do passo de tamanho 0,01 em ambas as situações, de maneira a termos garantida a precisão dos valores obtidos. Para o cálculo do ISE, por tratar-se da integração de uma diferença ao quadrado, utilizamos a integral de Riemman, com um passo de tamanho 0,01, procedimento que se mostrou estável nas validações efetuadas, obtendo-se uma precisão razoável.

Salientamos, entretanto, que o principal cuidado a ser tomado refere-se ao intervalo de truncamento da integral, já que grande parte dos cálculos e simulações envolvem integrais que possuem pelo menos um de seus limites tendendo ao infinito. Utilizamos assim intervalos de integração grandes o suficiente para garantir a precisão dos resultados e a correção de nossas conclusões.

Aprimoramos as funções desenvolvidas em S-plus, no sentido de disponibilizarmos uma biblioteca de funções para estimação funcional que seja flexível e que possibilite uma boa interatividade com o usuário, aumentando sua acessibilidade pela facilidade de manipulação das funções. No caso do programa em C, vamos implantar rotinas que possibilitem seu uso em simulações e em estudos mais aprofundados do comportamento do método na estimativa de outros tipos de funções.

# Capítulo 5

## Simulações e Aplicações

Em todas as utilizações de técnicas de suavização por núcleo, tanto para estimação da função de distribuição quanto para função de densidade, utilizamos o núcleo gaussiano, exceto onde indicado.

### 5.1 Simulação

Realizamos um conjunto de simulações com o objetivo de observar o comportamento do método na estimação de dados provenientes de algumas distribuições conhecidas. Buscamos algum conjunto de simulações existente na literatura de maneira a podermos comparar o desempenho do método frente a outros já existentes. Optamos por utilizar aquele conjunto de simulações comentados por Bowman [6] porque parece-nos conveniente para uma primeira avaliação do comportamento do estimador proposto.

As simulações compõem-se de amostras aleatórias oriundas das seguintes distribuições: normal padrão, gama(2,1) e uma mistura de normais  $N(-1,5;1)$  e  $N(1,5;1)$ . Bowman realizou este experimento, através da geração de 100 amostras de tamanhos 25, 50, 100 e 200, para cada uma destas distribuições. Apesar de termos re-executado este experimento, resolvemos expandi-lo, repetindo os mesmos tamanhos amostrais só que aumentando para

1000 o número de amostras, pois observamos algumas outras características destas distribuições. Verificamos o comportamento do estimador  $H$ , da janela ótima e do  $ISE(h)$  de cada uma das amostras. Além disso, comparamos nossas estimativas com a estimativa efetuada pela função característica empírica e a estimativa obtida pela integral da função de densidade. A estimativa da função de densidade foi obtida através do método proposto por Chiu [8] e dos resultados obtidos por Damasceno [10]. Para permitir uma análise mais precisa, efetuamos um teste t-pareado, no sentido de verificarmos a hipótese de igualdade entre o ISE obtido a partir do núcleo estimador e o ISE apresentado pela função característica empírica. Realizamos o mesmo tipo de teste no sentido de compararmos o ISE apresentado pela estimativa a partir do núcleo estimador com aquele da integral do núcleo estimador da função de densidade. Em ambos os casos, a hipótese alternativa é de que o ISE do núcleo estimador é o menor deles.

Tabela 5.1: Valores de  $H$ , Janela Ótima e MISE Assintótico

Distribuição	n	$H_0$	$h_{opt_0}$	$MISE(h)$
Normal (0,1)	25		0,5436	0,016458
	50		0,4315	0,008859
	100	0,1411	0,3425	0,004680
	200		0,2718	0,002439
	500		0,2003	0,001016
Gama (2,1)	25		0,4486	0,024938
	50		0,3560	0,012991
	100	0,2500	0,2826	0,006703
	200		0,2243	0,003434
	500		0,1653	0,001407
Mistura de Normais	25		0,7974	0,032458
	50		0,6329	0,017157
	100	0,0445	0,5023	0,008947
	200		0,3987	0,004620
	500		0,2937	0,001907

Para fins de comparação, a Tabela (5.1) apresenta valores teóricos dos parâmetros sendo estimados.

Apresentamos na tabela 5.2 os resultados obtidos nas simulações com 1.000 amostras das distribuições de interesse, nos tamanhos amostrais definidos para as simulações. Para facilitar o entendimento, os resultados serão resumidos em três tabelas, referentes às estimativas de  $H$ , de  $h_{opt}$  e nos valores do  $ISE$  encontrados nas estimativas obtidas pelo núcleo estimador, pela integral do núcleo estimador da densidade e pela função de distribuição empírica.

Tabela 5.2: Sumário das Simulações - Estimação de H

Distribuição	n	$\hat{E}(\hat{H})$	$\hat{\sigma}(\hat{H})$	$\hat{E}(\hat{H} - H_0)^2$
Normal (0,1)	25	0,1436	0,1814	0,03289
	50	0,1395	0,0710	0,00504
	100	0,1426	0,0484	0,00234
	200	0,1412	0,0313	0,00098
	500	0,1413	0,0194	0,00038
Gama (2,1)	25	0,1202	0,1697	0,04561
	50	0,1277	0,0872	0,02255
	100	0,1432	0,0693	0,01621
	200	0,1570	0,0549	0,01166
	500	0,1731	0,0410	0,00759
Mistura de Normais	25	0,0169	0,01226	0,009103
	50	0,0174	0,00714	0,000786
	100	0,0178	0,00683	0,000760
	200	0,0176	0,00460	0,000745
	500	0,0180	0,00518	0,000728

Com relação aos resultados encontrados na estimação da distribuição normal, podemos observar que os valores de  $\hat{H}$  estão bastante próximos do valor teórico, observando-

se que sua variabilidade diminui à medida em que o tamanho amostral aumenta, percebendo-se também que à medida em que aumenta o valor de  $\Lambda$ , o estimador fica mais próximo do valor teórico. Quanto aos valores de  $\hat{h}_{opt}$ , verifica-se que a proximidade ao valor teórico aumenta à medida em que o tamanho amostral aumenta, enquanto diminui sua variabilidade. Observando-se os valores estimados do MISE, verifica-se que o núcleo estimador da função apresenta um bom desempenho, comparado ao apresentado pela função empírica e pela integral do núcleo estimador da função de densidade. Observando-se o p-valor do teste t- pareado executado, verifica-se que há evidências suficientes para se rejeitar a hipótese de igualdade entre o ISE do núcleo estimador e da função empírica, em favor da hipótese alternativa de que o ISE da função empírica é maior que o ISE do núcleo estimador. Chega-se à mesma conclusão quando se compara o ISE da integral do núcleo estimador da função de densidade ao do núcleo estimador.

Tabela 5.3: Sumário das Simulações - Estimação de  $h_{opt}$

Distribuição	n	$\hat{E}(\hat{h}_{opt})$	$\hat{\sigma}(\hat{h}_{opt})$	$\hat{E}(\hat{h}_{opt} - h_{opt_0})^2$
Normal (0,1)	25	0,5889	0,1052	0.01311
	50	0,4500	0,0593	0.00386
	100	0,3479	0,0339	0.00117
	200	0,2740	0,0187	0.00035
	500	0,2007	0,0089	0.00008
Gama (2,1)	25	0,6589	0,1549	0.06823
	50	0,4780	0,0851	0.02211
	100	0,3541	0,0474	0.00735
	200	0,2680	0,0276	0.00267
	500	0,1889	0,0137	0.00075
Mistura de Normais	25	1,1460	0,1408	0.14134
	50	0,8840	0,0804	0.06952
	100	0,6918	0,0493	0.03833
	200	0,5473	0,0288	0.02292
	500	0,4005	0,0209	0.01183

No caso da distribuição gama, observamos que os valores da estimativa de  $H$  estão bastante afastados de seu valor verdadeiro, induzindo um vício na estimação funcional. A variabilidade destas estimativas apresenta uma lenta diminuição à medida em que o tamanho amostral aumenta, o mesmo ocorrendo com a distância entre a estimativa e o valor verdadeiro. Com relação aos valores de  $\hat{h}_{opt}$ , verifica-se que os valores estimados estão sempre acima do valor teórico, principalmente nas amostras de menor tamanho, indicando que o núcleo estará “sobre-suavizando” a estimativa da função.

Tabela 5.4: Sumário das Simulações - Estimação do MISE

Distribuição	n	$\widehat{MISE}$			p-valor	
		$\hat{F}_n$	$\int \hat{f}_n$	$\tilde{F}$	Emp	Dens
Normal (0,1)	25	0,018543	0,019269	0,023131	0	0
	50	0,009711	0,010383	0,011854	0	0
	100	0,004724	0,005189	0,005698	0	0
	200	0,002454	0,002772	0,002861	0	0
	500	0,000966	0,001146	0,001099	0	0
Gama (2,1)	25	0,025181	0,093018	0,029735	0	0
	50	0,013025	0,044311	0,015122	0	0
	100	0,006593	0,020821	0,007471	0	0
	200	0,003309	0,010013	0,003698	0	0
	500	0,001386	0,003834	0,001510	0	0
Mistura de Normais	25	0,031275	0,033044	0,038496	0	0
	50	0,016883	0,018834	0,020007	0	0
	100	0,009108	0,010805	0,010395	0	0
	200	0,004766	0,006158	0,005257	0	0
	500	0,001939	0,002856	0,002081	0	0

$\hat{F}_n$ : N. estim. distribuição;  $\hat{f}_n$ : N. estim. densidade;  $\tilde{F}$ : Função dist. empírica

O ISE apresentado pelo núcleo estimador está mais próximo do ISE da função de distribuição empírica, embora o teste t-pareado indique que o desempenho do núcleo estimador é melhor que o da função empírica. Salientamos que, neste caso, há um problema

de fronteira bastante evidente, já que a função gama é definida a partir de zero, com o inconveniente de que a função gama(2,1) não possui derivada segunda em zero, comprometendo a aplicação do método. Julgamos que, numa primeira tentativa, uma transformação poderia resolver esta situação possibilitando melhores estimativas. Este tipo de situação motiva-nos a encaminhar a pesquisa no sentido de aprofundarmos o estudo das alternativas para a melhoria das estimativas funcionais nestes casos.

Das estimativas da mistura de normais, verifica-se que uma situação análoga àquela apresentada pela distribuição gama, com os valores de  $\hat{h}_{opt}$  consistentemente maiores que os valores assintóticos, para todos os tamanhos amostrais. A variabilidade não é tão elevada quanto na situação discutida anteriormente sobre a distribuição gama. Percebe-se também que as estimativas de  $H$  estão bastante distantes de seu valor verdadeiro, com a variabilidade diminuindo razoavelmente com o tamanho amostral. Embora o ISE do núcleo estimador esteja próximo daquele apresentado pela função de distribuição empírica, o teste t-pareado indica que o mesmo é menor que o ISE da mesma. Esta situação pode indicar uma sub-avaliação no valor de  $H$ . Salienta-se que a função característica desta distribuição apresenta um componente periódico que leva a função a apresentar um máximo local, situação que impede que as informações dos dados contidas nesta faixa de frequência sejam captadas da maneira como  $\Lambda$  está definido. Julgamos que deveríamos estudar um aprimoramento nesta variável no sentido de permitir sua atuação eficaz em situações similares.

Apresentamos a figura (5.1), ilustrando uma das simulações efetuadas, neste caso com distribuição normal. Na figura estão dispostos os histogramas de diversas estimativas de interesse, cujo comportamento acompanhamos. Este tipo de visualização permitiu um grande auxílio na realização de verificações empíricas necessárias ao andamento da pesquisa.

Com relação ao experimento original de Bowman, apresentamos as figuras (5.2), (5.3) e (5.4) que comparam os gráficos “box-plot” dos vários tamanhos amostrais estudados referenciados ao  $\log(ISE)$ . Utilizamos, neste caso, o núcleo de Epanechnikov, da mesma maneira que definido no artigo mencionado, para preservarmos as características originais do experimento. Comparando-se com a figura apresentada em Bowman [6], verifica-se que o

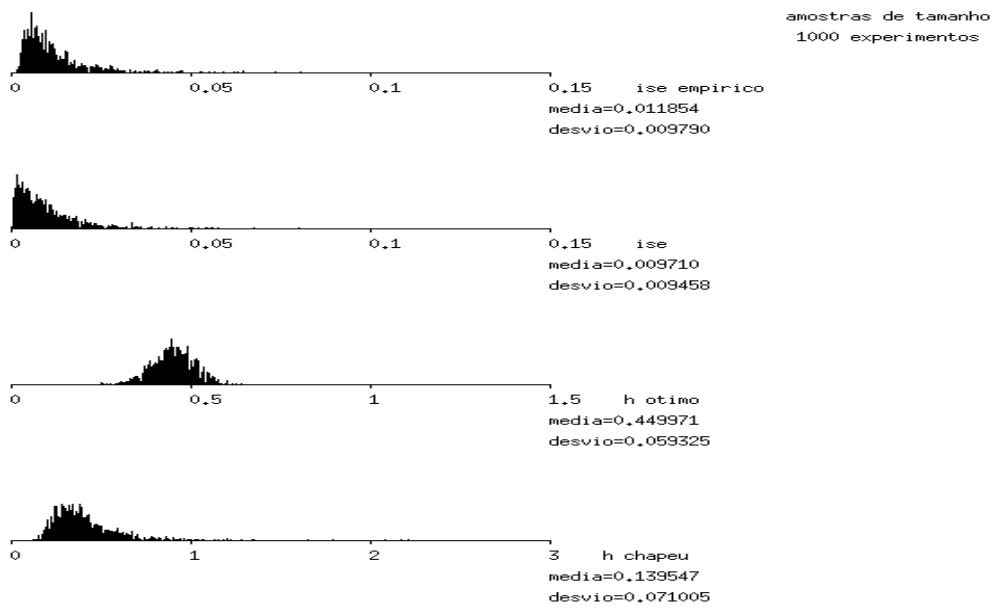


Figura 5.1: Histogramas de Acompanhamento de Simulação

desempenho das estimativas através do estimador proposto são melhores que aquelas então observadas .

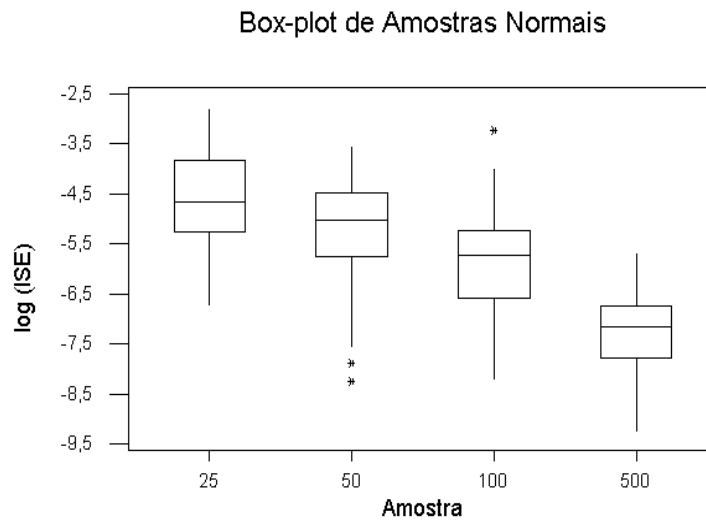


Figura 5.2: Box-plot de Amostras Normais

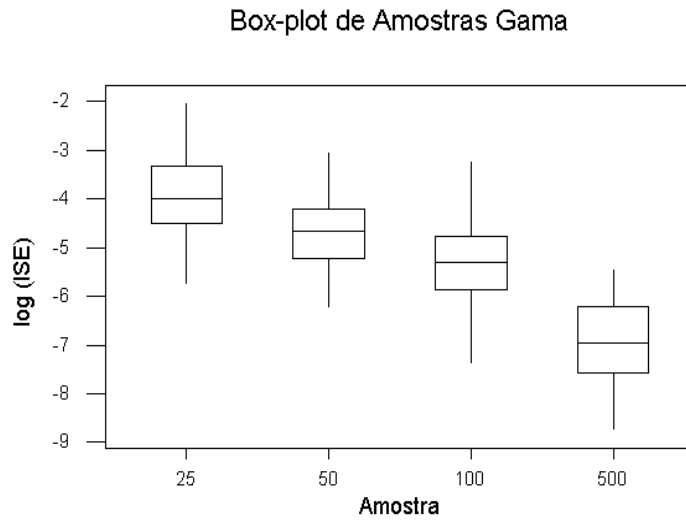


Figura 5.3: Box-plot de Amostras Gama

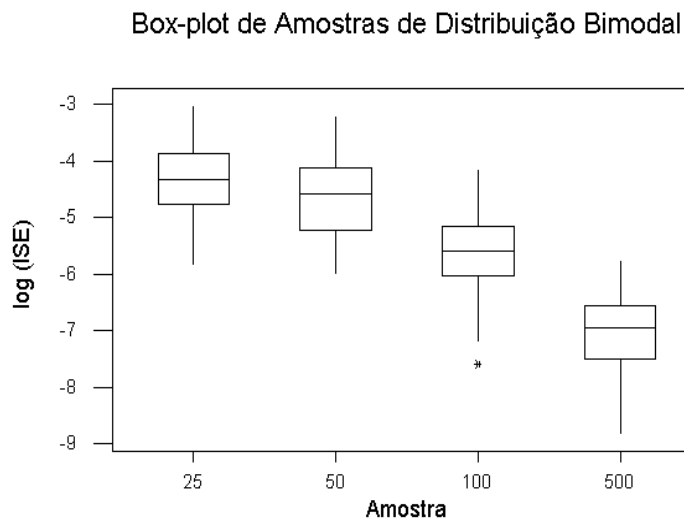


Figura 5.4: Box-plot de Amostras de Distribuição Bimodal

## 5.2 Estimação de Capacidade de Processo

A motivação para esta simulação foi dada por Polansky [30] em sua proposta para se estimar a capacidade de um processo através de núcleo estimador. Define-se capa-

cidade de um processo como sendo a habilidade em manter, dentro do especificado, uma característica de qualidade  $X$ . Sendo  $L$  e  $U$  respectivamente, os limites inferior e superior desta especificação e assumindo que  $X$  segue uma distribuição contínua  $F$ , a capacidade de um processo é função de:

$$p = 1 - F(U) + F(L)$$

em que  $p$  é a probabilidade de um item ser produzido fora da especificação.

Da literatura verifica-se que o valor de  $p$  para processos capazes é razoavelmente baixo, indicando uma dificuldade técnica para estimar seu valor, desde que, dado um conjunto de dados, o mesmo estará situado nas caudas da distribuição, situação que normalmente dificulta o desempenho das estimações.

Uma solução paramétrica ao problema é estimar  $F$  como sendo uma distribuição normal. Então se estima  $F$  por  $\tilde{F}(x) = \Phi((x - \hat{\mu})/\hat{\sigma})$ , onde  $\Phi(\cdot)$  é a distribuição normal padrão,  $\hat{\mu}$  é a média amostral e  $\hat{\sigma}$  é o desvio padrão amostral. Assim, a estimativa de  $p$  será dada por:

$$\tilde{p} = 1 - \Phi\left(\frac{U - \hat{\mu}}{\hat{\sigma}}\right) + \Phi\left(\frac{L - \hat{\mu}}{\hat{\sigma}}\right)$$

Nesta abordagem, a confiabilidade das estimativas de capacidade somente será garantida quando os dados não se desviarem da distribuição normal. Outra abordagem é a estimação de  $p$  através de núcleo estimador.

O experimento de Polansky consistiu na avaliação do desempenho de estimador, com amostras de tamanhos 50, 100, 200 e 400 de vários tipos de distribuição. Para fins de simulação, a especificação do processo era selecionar aqueles valores de maneira que  $p = 0,001$ , com metade de defeitos abaixo de  $L$  e a outra metade acima de  $U$ . Repetimos parte do experimento efetuado por Polansky de maneira a podermos comparar o núcleo estimador proposto em situações de estimação de valores com a característica de estarem situados na cauda da distribuição. Realizamos 1000 experimentos ao invés dos 10000 efetuados por Polansky, já que nossa comparação restringiu-se aos resultados referentes à distribuição normal padrão. Realizamos a estimativa também com o estimador paramétrico  $\tilde{p}$  descrito acima.

Os resultados estão apresentados na tabela (5.5), sendo que os valores encontrados para o estimador paramétrico parecem indicar um equívoco de ordem de grandeza nos resultados apresentados por Polansky. Ao se confirmar esta hipótese, as estimativas efetuadas pelo núcleo estimador proposto estão similares às aquelas observadas por Polansky. Saliente-se que Polansky recomenda o uso de núcleo estimador em situações em que a distribuição subjacente aos dados seja não normal, já que em situações de normalidade o estimador paramétrico comporta-se melhor.

Os próximos passos da pesquisa serão estudar o comportamento do núcleo estimador ao efetuar estimação funcional na região das caudas de distribuição. Parece-nos um bom caminho verificar nas distribuições estudadas por Polansky, que se baseiam no estudo de distribuição assimétrica ( $\chi^2$ , com 4 graus de liberdade), distribuição pesada na cauda ( $t$ , com 4 graus de liberdade) e distribuição com caudas pequenas (beta(2,4)).

### 5.3 Estimação de Taxa de Falha

Dada uma distribuição com função densidade de probabilidade  $f$  e sendo  $F'(t) = f(t)$ , a função taxa de falha é definida como :

$$r(t) = \frac{f(t)}{1 - F(t)}$$

Dada uma amostra  $X_1, \dots, X_n$  uma abordagem natural à estimação não paramétrica

Tabela 5.5: Resumo de Simulação - Estimação de  $p$

Distribuição	n	$\hat{E}(\hat{p})$	$\hat{\sigma}(\hat{p})$	$\hat{E}(\hat{p} - p_0)^2$	$\hat{E}(\tilde{p} - p_0)^2$
Normal (0,1)	50	0,00293	0,0044	$2,293 \times 10^{-5}$	$3,780 \times 10^{-6}$
	100	0,00200	0,0029	$9,183 \times 10^{-6}$	$1,138 \times 10^{-6}$
	200	0,00157	0,0020	$4,174 \times 10^{-6}$	$4,580 \times 10^{-7}$
	500	0,00126	0,0012	$1,511 \times 10^{-6}$	$1,548 \times 10^{-7}$

$$p_0 = 0,001$$

é através do estimador:

$$\hat{r}(t) = \frac{\hat{f}(t)}{1 - \hat{F}(t)} \quad (5.1)$$

Para estimar  $F$  utilizaremos o núcleo estimador da função de distribuição e a integral da função de densidade. Assim trabalharemos com dois estimadores de  $r$  de maneira a estudarmos o desempenho de ambos nesta situação. Os estimadores que definimos são:  $\hat{r}_{n_1} = \hat{f}_n / (1 - \hat{F}_n)$  e  $\hat{r}_{n_2} = \hat{f}_n / (1 - \int \hat{f}_n)$ .

Iniciamos a verificação da aplicação do uso do núcleo estimador na solução de problemas de estimação funcional em análise de sobrevivência, através do estudo do comportamento dos estimadores descritos acima, em simulações que envolvessem amostras de tamanho relativamente pequeno, embora ainda sem a presença de dados censurados.

Realizamos simulação com 100 amostras Weibull (2,1), com 50 elementos em cada uma e verificamos o desempenho da estimativa da taxa de falha através do cálculo do ISE de cada amostra desde a origem até os quantis  $p_{0,95} = 1,731$  e  $p_{0,99} = 2,146$ . Verificamos então os valores das média e variância de cada um destes estimadores, assim como o valor do  $\widehat{MSE} = \hat{E}(\hat{r} - r_0)^2$  em cada um dos quantis citados. Os resultados estão apresentados na tabela (5.6).

Tabela 5.6: Resultado de Simulação - Estimativa de Taxa de Falha

	$p_{0,95}$		$p_{0,99}$	
	$\hat{r}_{n_1}$	$\hat{r}_{n_2}$	$\hat{r}_{n_1}$	$\hat{r}_{n_2}$
Média	1.5547	0.7952	247.36	6.53
Variância	47.6086	4.3293	442151.32	239.33
$\widehat{MSE}$	49.5497	4.9183	4429218.21	279.63

Observa-se que os erros quadráticos médios das estimativas efetuadas por  $\hat{r}_{n_2}$  aparentam ser menores que aquelas apresentadas por  $\hat{r}_{n_1}$ . Entretanto um teste t-pareado não forneceu evidências para rejeitar a hipótese de igualdade entre os erros das duas estimativas. De qualquer maneira, os estimadores assim constituídos não se comportaram bem nesta região. Verifica-se, também, neste caso a existência de problema de fronteira nas estimativas

efetuadas por núcleo, embora a estimativa da função tenha se mantido estável nas regiões mais à esquerda da distribuição. Para uma visão mais completa da situação, apresentamos a figura 5.5, com a estimação da taxa de falha de uma das amostras, podendo observar-se que há uma situação de estabilidade, com uma explosão nos valores de taxas de falha nos quantis mais elevados. A situação das estimativas nas proximidades de zero apresenta-se bem evidenciada.

Baseados em Silverman [39], cremos que a principal contribuição para os erros envolvidos na estimação de  $r(t)$  deve-se ao numerador da expressão (5.1), podendo a estimativa da taxa de falha explodir na cauda direita da distribuição, em que flutuações relativamente pequenas em  $\hat{f}(t)$ , para  $t$  grande, tornar-se-ão muito maiores quando divididos pelo quociente de  $1 - \hat{F}(t)$ . Um primeiro passo na solução deste problema seria a tentativa de transformar os dados para diminuir sua variabilidade e poderemos obter estimativas mais estáveis de  $\hat{r}(t)$ . Esta ampliação de erros nas caudas pode também estar indicando o uso de janela local na estimação da taxa de falha, além de termos aqui presente a questão de fronteira que pode estar influenciando a estimativa. O estudo realizado contempla dados completos. Um dos próximos passos da pesquisas será o aprofundamento no uso de núcleos estimadores em presença de dados censurados.

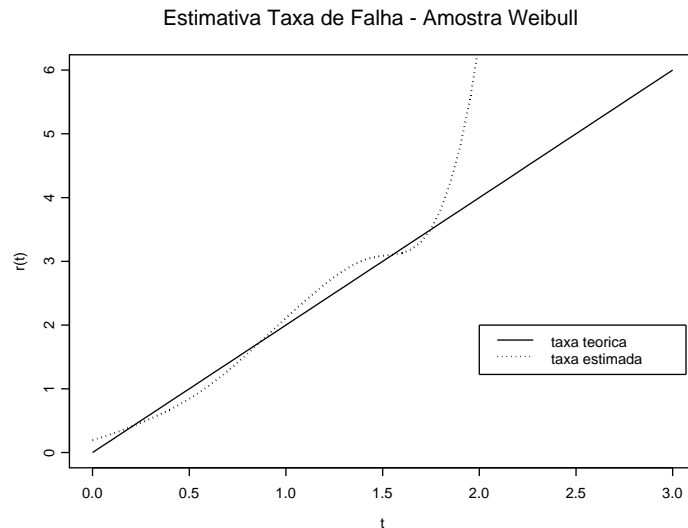


Figura 5.5: Estimação de Taxa de Falha - Amostra Weibull (2,1)

## 5.4 Aplicação

Uma das utilizações mais comuns da técnica de núcleo estimador é a análise da estrutura de um conjunto qualquer de dados. Para exemplificar este fato, analisamos os valores da tensão de escoamento de corpos de prova da matéria prima de cantoneiras de aço, utilizadas na fabricação de estruturas metálicas. Foi utilizada uma amostra de tamanho 628. A tensão de escoamento é uma característica inerente ao tipo de aço, independente de suas dimensões ou geometria.

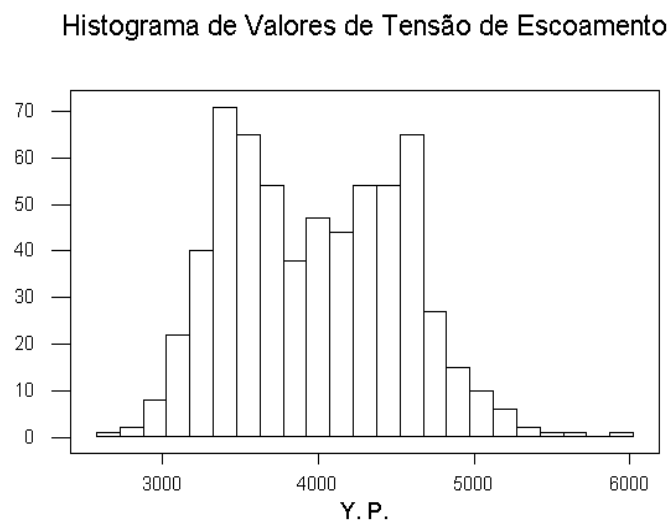


Figura 5.6: Tensão de Escoamento de Aço - Histograma

Em uma análise inicial, através do histograma da figura (5.6), verificou-se uma variação de valores elevados, com a aparência de bimodalidade, embora não pudéssemos localizar com precisão a posição destas modas. Numa primeira tentativa de suavização, com os dados originais, verificamos que o método não era eficiente pois a função característica empírica, por ter um decaimento muito rápido, não fornecia as informações contidas nas frequências relativamente altas. O valor de janela assim calculado fornecia uma estimativa da função de densidade que apresentava oscilações elevadas. Optamos, então, por uma transformação linear dos dados, dividindo-os pelo desvio-padrão amostral, no sentido de a

função característica empírica fornecer mais informações.

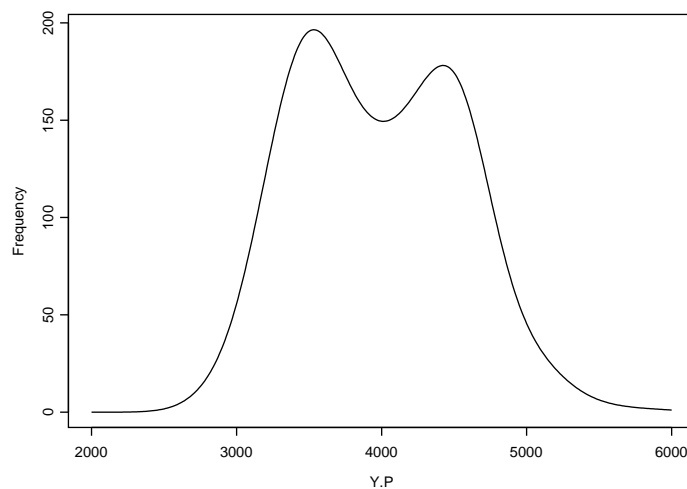


Figura 5.7: Tensão de Escoamento de Aço - Estimação da Densidade

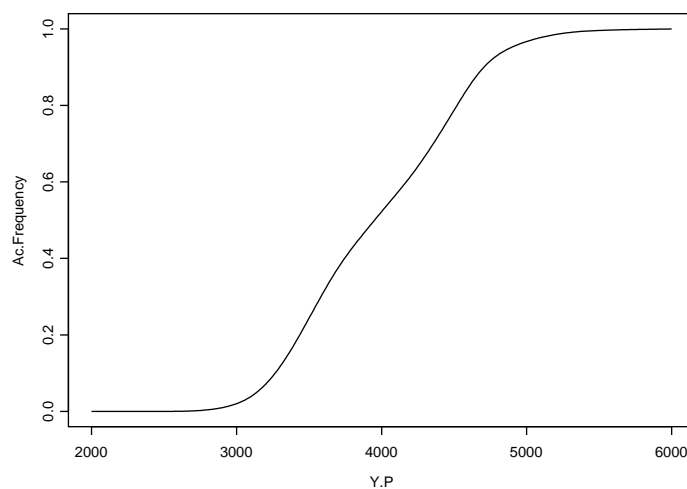


Figura 5.8: Tensão de Escoamento de Aço - Estimação da Distribuição

Da análise resultante, apresentamos as figuras (5.7) e (5.8), referentes, respectivamente, à estimativa da função de densidade e função de distribuição. Percebe-se assim claramente a bimodalidade sugerida pelo histograma, nas regiões de  $3.500 \text{ kg/cm}^2$  e de  $4.700$

$kg/cm^2$ . Referindo-se aparentemente a dois tipos de aço distintos, o que de fato aconteceu, já que a composição da amostra estudada incluía os aço tipo A36 e A572, cujas tensões de escoamento teóricas são :  $2.531 kg/cm^2$  e  $3.515 kg/cm^2$ . Apesar de o gráfico da função de distribuição não permitir uma percepção clara do fenômeno da bimodalidade, permite, com extrema facilidade, estimativas de percentis necessárias aos tipos de inferência desejados.

Verifica-se assim que as técnicas de suavização através do núcleo estimador, tanto da densidade quanto da distribuição, são ferramentas estatísticas importantes no campo de análise de dados.

# Capítulo 6

## Conclusões e Pesquisas Futuras

Os métodos não paramétricos de estimação funcional são bastante úteis principalmente para se obter informações de um conjunto de dados cuja distribuição não pode ser aproximada por modelos paramétricos. Temos abordado no presente trabalho o problema de estimação da janela ótima na definição do núcleo estimador da função de distribuição. Temos visto que, quando a função de distribuição é duas vezes continuamente diferenciável, os resultados são promissores. Quando a última condição não é satisfeita, o método fica comprometido, deixando em aberto o problema de estimação da função na vizinhança do ponto em que a dupla diferenciabilidade não é satisfeita. Geralmente esse problema acontece em pontos de fronteira. Uma possível abordagem desse problema é redefinir o estimador em tais pontos. Alguns artigos [13] e [12], por exemplo, abordam esse problema.

Com relação ao método em si, será buscado o aprimoramento do conceito de  $\Lambda$ , para permitir a captação de informações contidas em frequências que em sua definição atual não são atingidas, principalmente na presença de multimodalidade. Uma boa análise exploratória dos dados poderá revelar a necessidade de depurar o método em situações reais.

Resolvidas estas questões, a pesquisa será aprofundada através do estudo de núcleos estimadores adequados à estimação de função em presença de dados censurados e principalmente nos métodos automáticos de escolha da janela, com ênfase nos do tipo “plug-

in”, que aparentam adequar-se melhor à estimação funcional. No âmbito da estimação de taxa de falhas, a complexidade do problema exige um aprofundamento maior na pesquisa de estimadores adequados e de métodos computacionalmente convenientes à escolha da janela ótima. Parece-nos que na literatura esta questão tem sido tratada a partir de suavização dos estimadores não paramétricos usuais, ou seja, o estimador de Kaplan-Meier e o de Nelson-Aalen. O uso de janela local é também um tema que deverá ser abordado no prosseguimento de nosso trabalho, de maneira a estudarmos a exequibilidade de se utilizar mais de uma janela na análise de um mesmo conjunto de dados.

O uso de núcleo na estimação da densidade invariante de processos markovianos será uma decorrência natural em nossos estudos, abrindo um campo de aplicações teóricas e práticas bastante significativo. Acreditamos que a extensão do método estudado no caso i.i.d. ao caso markoviano é de extrema relevância, mas precisa de ferramentas técnicas mais avançadas, sendo que estudos em nível de doutoramento poderiam propiciar a continuação deste trabalho.

# Apêndice

## Programa em C

```
/* determinacao da janela otima para o estimador de distribuicoes H */
/* atraves de integracao por Simpson baseadas em amostras normais */

#include<stdlib.h>
#include<math.h>
#include<stdio.h>
#define IA 16807
#define IM 2147483647
#define AM (1.0/IM)
#define IQ 127773
#define IR 2836
#define MASK 123459876

int xjmax;
double pi=3.14159265;
double cgauss;
long idum;
char aqui[30];
int normal_impar,nrep,al;
double normal_y,amostran[1000];
double coef_simpson(int i);
double pow(double x, double y);
double sqr(double x);
double normal2(double media, double desv);
double erfff(float x);
double gauss(double x);
double teste_ad();
```

```

void linha(int x0, int y0, int x1, int y1);
void linha1(int x0, int y0, int x1, int y1);
void linha2(int x0, int y0, int x1, int y1);
void linha3(int x0, int y0, int x1, int y1);
void linha4(int x0, int y0, int x1, int y1);
void putp(int x0, int y0, int cor);
void putp1(int x0, int y0, int cor);
void putp2(int x0, int y0, int cor);
void putp3(int x0, int y0, int cor);
void putp4(int x0, int y0, int cor);
float alea(long idum);
domimag()
{
    int i,iii,iiii,j,i1,i2,n,cor,xasp,yasp,a1,a2,b1,b2,
        jm,mx,my,car,band,flagg,ilamb,jj,flag_ker,
        h_opt_i,g_opt_i,ise_i,ise_g_i,xj,ise_emp_i;
    double x1,x2,y1,y2,z1,z2,r1,r2,k1,k2,k3,k4,u1,u2,v1,v2,rr,
        norma,x1min,x2min,x1max,x2max,esc,h,d,asp,aux,r_k,sumfg,Fng,
        stud_emp,stud_g,sump1,sump0,sumpproc,proc0,sumpproc0;

    double c,hc,gc,h_opt,g_opt,pi=3.14159265;
    double x,y,z,ii,delta,ise,ise_g,sump,argx,
        soma,soman,constante,lamb,phi2,soma_real,soma_imag;
    long fd[1000], /* dist. acumulada de f */
        fdp[1000], /* f.d.p. f da v.a. X */
        fdn[1000]; /* fdp normal */
    int
        vetor_hc[1000], /* histograma de valores gerados-H chapeu */
        vetor_h_opt[1000], /* histograma de valores gerados-H optimo */
        vetor_ise[1000], /* histograma de valores gerados-ISE para H */
        vetor_gc[1000], /* histograma de valores gerados-G chapeu */
        vetor_g_opt[1000], /* histograma de valores gerados-G optimo */
        vetor_ise_g[1000], /* histograma de valores gerados-ISE para G */
        vetor_ise_emp[1000];/* histograma de valores gerados-ISE empirica */
    double vetor_phi2[1000],vetor_lambda[1000];
    long p[4],q[4],tamv,xi,yi,zi,bloco,nn,iphi2,kk;
    double xmin,ymin,zmin,xmax,ymax,zmax,dx,dy,dz,
        nexp,nrep2,nrep4,theta,quiq,cgraf,
        delta_lambda,corte;
    double int_bow, var_ker, cte_ker, inv3, inv5;
    double cte_ker_g, g_c, Lambda ;
    double m_lambda, s_lambda, d_lambda, s2_lambda;
    double m_hc, s_hc, d_hc, s2_hc;
    double m_hopt, s_hopt, d_hopt, s2_hopt;

```

```

double m_ise, s_ise, d_ise, s2_ise;
double m_gc, s_gc, d_gc, s2_gc;
double m_gopt, s_gopt, d_gopt, s2_gopt;
double m_ise_g, s_ise_g, d_ise_g, s2_ise_g;
double m_pc, s_pc, d_pc, s2_pc;
double m_emp, s_emp, d_emp, s2_emp;
double pv_emp;
double pv_g;
double ise_emp;
double hc0,gc0,h_opt0,g_opt0;
double mse_h,mse_g,mse_hopt,mse_gopt;
double Di_emp,m_Di_emp, d_Di_emp, s2_Di_emp,s_Di_emp;
double Di_g,m_Di_g, d_Di_g, s2_Di_g,s_Di_g;
double m_proc, s_proc, d_proc, s2_proc,mse_proc;
double desvio_amostra,media_amostra2,media_amostra;
double sumpproc0s, sumpproc0s2, m_proc0, d_proc0;
FILE *anormal, *s, *vetor;

setcolor(1);
cor=1;
inv3=1.0/3.0;
inv5=1.0/5.0;
cgauss=sqrt(2.0 * pi);
nexp=1000; /* numero de experimentos (numero de amostras) */
nrep=500; /* numero de repeticoes (tamanho de cada amostra) */
hc0=0.25/sqrt(pi);
gc0=0.2115711;
h_opt0=1.589573060*pow(nrep,-inv3);
g_opt0=1.05922384105*pow(nrep,-inv5);
proc0=0.001;

flag_ker=0; /* 0:Nucleo Normal 1:Nucleo de Epanechnikov */
idum=1; /* semente aleatoria */
anormal=fopen("anormal.dat","w");
s=fopen("saida.dat","w");
vetor=fopen("vetor.dat","w");
normal_impar=0;
flagg=0; /* 1:aparecem os graficos intermediarios */
tamv=640;
nn=3*nrep/5;
theta=0.4;
cgraf=10;
bloco=32;
nrep4=nrep/4.0;
nrep2=nrep*2.0;

```

```

constante=nrep*(1.56/(sqrt(2.0*3.14159265)))/50.0;
if(flag_ker==0){/*normal*/
  int_bow=0.56418958323;
  var_ker=1.0;
  r_k=0.282093392;
}
else if(flag_ker==1){/* Epanechnikov */
  int_bow=0.257142857;
  var_ker=0.2;
  r_k=0.6;
}
cte_ker=int_bow/sqr(var_ker);
cte_ker_g=r_k/sqr(var_ker);
for(i=0;i<1000;i++){vetor_lambda[i]=0;vetor_hc[i]=0;}
for(i=0;i<1000;i++){vetor_h_opt[i]=0;}
for(i=0;i<1000;i++){vetor_ise[i]=0;}
s_lambda=0.0;
s2_lambda=0.0;
s_hc=0;
s2_hc=0;
s_hopt=0;
s2_hopt=0;
s_ise=0;
s2_ise=0;
s_gc=0;
s2_gc=0;
s_gopt=0;
s2_gopt=0;
s_ise_g=0;
s2_ise_g=0;
s_emp=0;
s2_emp=0;
mse_h=0;
mse_g=0;
mse_hopt=0;
mse_gopt=0;
mse_proc=0;
s_proc=0;
s2_proc=0;
s_Di_emp=0;
s2_Di_emp=0;
s_Di_g=0;
s2_Di_g=0;
sumpproc0=0;
sumpproc0s=0;

```

```

sumpproc0s2=0;
for(iiii=0;iiii<nexp;iiii++){
    for(i=0;i<1000;i++){fdp[i]=0;fd[i]=0;}
    if(flagg){
        {setcolor(6);linha(0,470,639,470);}
        {setcolor(1);linha(0,110,639,110);}
        {setcolor(3);linha(319,471,319,479);}
        for(i=0;i<50;i++){putp((int)(4*i)-1,471,3);}
        for(i=0;i<50;i++){putp((int)(4*i)-1,472,3);}
        for(i=0;i<50;i++){putp((int)(4*i)-1,473,3);}
    }
    for(i=0;i<4;i++){p[i]=0;}
    media_amostra=0;
    media_amostra2=0;
    for(jj=0;jj<nrep;jj++){/*geracao amostra c/ nrep valores normais*/
        x=normal2(0,1);
        amostran[jj]=x;
        media_amostra+=x;
        media_amostra2+=sqr(x);
        fprintf(anormal,"%g\n",x);
        quiq=(x+5)/cgraf;
        xi=(int)(tamv*quiq+0.5); if((xi<1000)&&(xi>=0))fdp[xi]++;
        if(flagg)putp(xi,(int)((470-(fdp[xi])/1.0)),2);
    }
    desvio_amostra = sqrt( (media_amostra2 - sqr(media_amostra) /
nrep ) / (nrep-1.0) );
    media_amostra/=nrep;
    for(i=0;i<nrep;i++) /* bolha*/
        for(j=i+1;j<nrep;j++)
            if(amostran[i]>amostran[j]){
                aux=amostran[i];
                amostran[i]=amostran[j];
                amostran[j]=aux;
            }
    fprintf(anormal,"\n");
    if(flagg){
        soma=0; /* display na tela da amostra normal */
        soman=0;
        for(i=0;i<1000;i++){
            x=(i/64.0-5.0);
            soman+=constante*exp(-x*x/2);
            soma+=fdp[i];
            fd[i]=soma;
            fdn[i]=soman;
        }
    }
}

```

```

        putp(i,(int)((470-360*(fd[i])/nrep)),4);
        putp(i,(int)((470-360*(fdn[i])/nrep2)),1);
    }
}
i=-1;
do{
    soma=0;
    j=0;
    do{
        i++;
        soma+=fdp[i];
        j++;
    }while(j<bloco);
if(flagg){
    setcolor(3);
    linha(i-bloco,(int)(470-360*soma/nn),i,(int)(470-
360*soma/nn));
    linha(i-bloco,470,i-bloco,(int)(470-360*soma/nn));
    linha(i,470,i,(int)(470-360*soma/nn));
}
}while(i<tamv);
clear();
}
/* calcular phi2(lambda) para a amostra normal acima */

if(flagg){
    {setcolor(6);linha(0,470,639,470);}
    {setcolor(1);linha(0,110,639,110);}
}
lamb=0.000;
iphi2=0;
delta_lambda=0.02;
do{
    soma_real=0.0;
    soma_imag=0.0;
    for(iii=0;iii<nrep;iii++){
        soma_real+=cos(lamb*amostran[iii]);
        soma_imag+=sin(lamb*amostran[iii]);
    }
    phi2=soma_real*soma_real+soma_imag*soma_imag;
    phi2=phi2/(nrep*nrep);
    vetor_phi2[iphi2]=phi2;
    iphi2++;
    if(flagg)
        putp((int)((lamb)*100),(int)(470-360*(phi2)),2);
}
}

```

```

        lamb+=delta_lambda;
    }while(lamb<=6.4); /* fim do calculo de phi em [0,6.4] */
    corte=3.0/nrep;
    iphi2=0;
    while(vetor_phi2[iphi2]>corte)iphi2++;
    hc=0.0;
    for(i=0;i<=iphi2;i++){
        lamb=i*delta_lambda;
        hc+=lamb*lamb*(vetor_phi2[i]-1.0/nrep);
    }
Lambda=iphi2*delta_lambda;
hc*=delta_lambda/pi;
fprintf(vetor,"%g ",hc);
if((int)(hc*500)<1000)vetor_hc[(int)(hc*500)]++;
if(flagg)
    putp((int)(hc*500/3),
        (int)(270-1*(vetor_hc[(int)(hc*500)])),3);
vetor_lambda[iphi2]++;
if(flagg)
    putp((int)((iphi2)*delta_lambda*100),
        (int)(470-1*(vetor_lambda[iphi2])),4);
h_opt=cte_ker/(hc*nrep);
h_opt=pow(h_opt,inv3);
fprintf(vetor,"%g\n",h_opt);
h_opt_i=(int)(h_opt*100+0.5);
if(h_opt*100<1000)vetor_h_opt[h_opt_i]++;
printf("%g ",h_opt);
if(flagg==0)
    putp(h_opt_i, vetor_h_opt[h_opt_i], 5);
delta=0.01;
ise=0.0;
x=-5.0;
while(x<=5.0){
    sump=0.0;
        for(i=0;i<nrep;i++){ /* nucleo estimador em x */
            argx=(x-amostran[i])/h_opt;
            sump+=erfff(argx);
        }
        sump = sump/nrep - erfff(x);
    ise+= sqr(sump);
    x+=delta;
}
ise*=delta;
printf("%g\n",ise);
ise_i=(int)(ise*10000+0.5);

```

```

        if(ise_i<1000)vetor_ise[ise_i]++;
        if(flagg==0)
            putp(ise_i, 200 + vetor_ise[ise_i], 6);

        /* process capability (Polanski) */
sump1=0.0;
x=3.291; /* integral de -3.291 a +3.291 e' 0.999 */
        for(i=0;i<nrep;i++){
            argx=(x-amostran[i])/h_opt;
            sump1+=erfff(argx);
        }
            sump1 = sump1/nrep;
sump0=0.0;
x=-3.291;
        for(i=0;i<nrep;i++){
            argx=(x-amostran[i])/h_opt;
            sump0+=erfff(argx);
        }
            sump0 = sump0/nrep;
            sumpproc=1.0-sump1+sump0;

            argx=( 3.291-media_amostra)/desvio_amostra;
sump1=erfff(argx);
            argx=(-3.291-media_amostra)/desvio_amostra;
sump0=erfff(argx);
            sumpproc0s+=1.0-sump1+sump0;
            sumpproc0s2+=sqr(1.0-sump1+sump0);
            sumpproc0+=sqr(1.0-sump1+sump0- 0.001);

delta=0.01;
ise_emp=0.0;
x=-5.0;
while(x<=5.0){
    sump=0.0;
        for(i=0;i<nrep;i++){ /* empirica */
            if(x>amostran[i])sump+=1.0;
        }
            sump = sump/nrep - erfff(x);
            ise_emp+= sqr(sump);
            x+=delta;
}
ise_emp*=delta;
    printf("%g\n", ise_emp);
ise_emp_i=(int)(ise_emp*10000+0.5);

```

```

if(ise_emp_i<1000)vetor_ise_emp[ise_emp_i]++;
if(flagg==0)
    putp(ise_emp_i, 400 + vetor_ise_emp[ise_emp_i], 7);

gc=0.0;
for(i=0;i<=iphi2;i++){
    lamb=i*delta_lambda;
    gc+=sqr(lamb*lamb)*(vetor_phi2[i]-1.0/nrep);
}
gc*=delta_lambda/pi;
if(gc*500<1000)vetor_gc[(int)(gc*500)]++;
if(flagg)
    putp((int)(400+gc*500/3),
        (int)(270-1*(vetor_gc[(int)(gc*500)])),8);
g_opt=cte_ker_g/(gc*nrep);
g_opt=pow(g_opt,inv5);
g_opt_i=(int)(g_opt*100+0.5);
if(g_opt_i<1000)vetor_g_opt[g_opt_i]++;
printf("%g ",g_opt);
if(flagg==0)
    putp(400+g_opt_i, vetor_g_opt[g_opt_i], 9);
delta=0.01;
ise_g=0.0;
Fng=0.0;
x=-5.0; xj=0; xjmax=800;
while(x<=5.0){
    sumfg=0.0;
    for(i=0;i<nrep;i++){ /* nucleo estimador densidade em x */
        argx=(x-amostran[i])/g_opt;
        sumfg+=gauss(argx);
    }
    sumfg/=(nrep*g_opt);
    Fng+=coef_simpson(xj)*sumfg*delta/3.0;
    ise_g+= sqr(Fng-erfff(x));
    x+=delta;
    xj++;
}
ise_g*=delta;
printf("%g\n",ise_g);
ise_g_i=(int)(ise_g*10000+0.5);
if(ise_g_i<1000)vetor_ise_g[ise_g_i]++;
if(flagg==0)
    putp(400+ise_g_i, 200 + vetor_ise_g[ise_g_i], 10);

```

```

s_lambda+=(Lambda);
s2_lambda+=sqr(Lambda);
    s_hc+=hc;
    s2_hc+=sqr(hc);
    s_hopt+=h_opt;
    s2_hopt+=sqr(h_opt);
    s_ise+=ise;
    s2_ise+=sqr(ise);
    s_gc+=gc;
    s2_gc+=sqr(gc);
    s_gopt+=g_opt;
    s2_gopt+=sqr(g_opt);
    s_ise_g+=ise_g;
    s2_ise_g+=sqr(ise_g);
    s_emp+=ise_emp;
    s2_emp+=sqr(ise_emp);
    mse_h+=sqr(hc-hc0);
    mse_g+=sqr(gc-gc0);
    mse_hopt+=sqr(h_opt-h_opt0);
    mse_gopt+=sqr(g_opt-g_opt0);
s_proc+=sumpproc;
s2_proc+=sqr(sumpproc);
mse_proc+=sqr(sumpproc-proc0);

Di_emp = ise_emp - ise;
    Di_g = ise_g - ise;
    s_Di_emp+=Di_emp;
    s2_Di_emp+=sqr(Di_emp);
    s_Di_g+=Di_g;
    s2_Di_g+=sqr(Di_g);
}
m_lambda = s_lambda/nexp;
d_lambda = sqrt( (s2_lambda - sqr(s_lambda) / nexp ) / (nexp-1.0) );

m_hc = s_hc/nexp;
d_hc = sqrt( (s2_hc - sqr(s_hc) / nexp ) / (nexp-1.0) );

m_hopt = s_hopt/nexp;
d_hopt = sqrt( (s2_hopt - sqr(s_hopt) / nexp ) / (nexp-1.0) );

m_ise = s_ise/nexp;
d_ise = sqrt( (s2_ise - sqr(s_ise) / nexp ) / (nexp-1.0) );

```

```

m_gc = s_gc/nexp;
d_gc = sqrt( (s2_gc - sqr(s_gc) / nexp ) / (nexp-1.0) );

m_gopt = s_gopt/nexp;
d_gopt = sqrt( (s2_gopt - sqr(s_gopt) / nexp ) / (nexp-1.0) );

m_ise_g = s_ise_g/nexp;
d_ise_g = sqrt( (s2_ise_g - sqr(s_ise_g) / nexp ) / (nexp-1.0) );

m_emp = s_emp/nexp;
d_emp = sqrt( (s2_emp - sqr(s_emp) / nexp ) / (nexp-1.0) );

m_Di_emp = s_Di_emp/nexp;
d_Di_emp = sqrt( (s2_Di_emp - sqr(s_Di_emp) / nexp ) / (nexp-1.0) );

m_Di_g = s_Di_g/nexp;
d_Di_g = sqrt( (s2_Di_g - sqr(s_Di_g) / nexp ) / (nexp-1.0) );

m_proc = s_proc/nexp;
d_proc = sqrt( (s2_proc - sqr(s_proc) / nexp ) / (nexp-1.0) );

m_proc0 = sumpproc0s/nexp;
d_proc0 = sqrt( (sumpproc0s2 - sqr(sumpproc0s) / nexp ) / (nexp-1.0) );

stud_emp = m_Di_emp * sqrt(nexp) / d_Di_emp;
stud_g = m_Di_g * sqrt(nexp) / d_Di_g;
pv_emp = 1.0 - erfff(stud_emp);
pv_g = 1.0 - erfff(stud_g);

mse_h/=nexp;
mse_g/=nexp;
mse_hopt/=nexp;
mse_gopt/=nexp;
mse_proc/=nexp;
sumpproc0/=nexp;

printf("nrep = %d nexp = %g\n",nrep,nexp);
printf("media_lambda = %g desv_lambda = %g\n",m_lambda,d_lambda);
printf("media_hc = %g desv_hc = %g\n",m_hc,d_hc);
printf("media_hopt = %g desv_hopt = %g\n",m_hopt,d_hopt);
printf("media_ise = %g desv_ise = %g\n",m_ise,d_ise);
printf("media_gc = %g desv_gc = %g\n",m_gc,d_gc);
printf("media_gopt = %g desv_gopt = %g\n",m_gopt,d_gopt);
printf("media_ise_g = %g desv_ise_g = %g\n",m_ise_g,d_ise_g);
printf("media_emp = %g desv_emp = %g\n",m_emp,d_emp);

```

```

printf("mse_h = %g   mse_g = %g\n",mse_h,mse_g);
printf("mse_hopt = %g   mse_gopt = %g\n",mse_hopt,mse_gopt);
printf("pvalor_emp = %g   pvalor_g = %g\n",pv_emp,pv_g);
printf("stud_emp = %g   stud_g = %g\n",stud_emp,stud_g);
printf("m_Di_emp = %g   m_Di_g = %g\n",m_Di_emp,m_Di_g);
printf("d_Di_emp = %g   d_Di_g = %g\n",d_Di_emp,d_Di_g);
printf("media_proc = %g   desv_proc = %g\n",m_proc,d_proc);
printf("mse_proc = %g\n",mse_proc);
printf("sumpproc0 = %g\n",sumpproc0);
printf("media_proc0 = %g   desv_proc0 = %g\n",m_proc0,d_proc0);

fprintf(s,"nrep = %d   nexp = %g\n",nrep,nexp);
fprintf(s,"media_lambda = %g   desv_lambda = %g\n",m_lambda,d_lambda);
fprintf(s,"media_hc = %g   desv_hc = %g\n",m_hc,d_hc);
fprintf(s,"media_hopt = %g   desv_hopt = %g\n",m_hopt,d_hopt);
fprintf(s,"media_ise = %g   desv_ise = %g\n",m_ise,d_ise);
fprintf(s,"media_gc = %g   desv_gc = %g\n",m_gc,d_gc);
fprintf(s,"media_gopt = %g   desv_gopt = %g\n",m_gopt,d_gopt);
fprintf(s,"media_ise_g = %g   desv_ise_g = %g\n",m_ise_g,d_ise_g);
fprintf(s,"media_emp = %g   desv_emp = %g\n",m_emp,d_emp);
fprintf(s,"mse_h = %g   mse_g = %g\n",mse_h,mse_g);
fprintf(s,"mse_hopt = %g   mse_gopt = %g\n",mse_hopt,mse_gopt);
fprintf(s,"pvalor_emp = %g   pvalor_g = %g\n",pv_emp,pv_g);
fprintf(s,"stud_emp = %g   stud_g = %g\n",stud_emp,stud_g);
fprintf(s,"m_Di_emp = %g   m_Di_g = %g\n",m_Di_emp,m_Di_g);
fprintf(s,"d_Di_emp = %g   d_Di_g = %g\n",d_Di_emp,d_Di_g);
fprintf(s,"media_proc = %g   desv_proc = %g\n",m_proc,d_proc);
fprintf(s,"mse_proc = %g\n",mse_proc);
fprintf(s,"sumpproc0 = %g\n",sumpproc0);
fprintf(s,"media_proc0 = %g   desv_proc0 = %g\n",m_proc0,d_proc0);

fim;;
fclose(anormal);
fclose(s);
fclose(vetor);
printf("fim do programa\n");
}

float alea(long idum0){
    long k;
    float ans;
    idum0 ^= MASK;
    k=(idum0)/IQ;
    idum0=IA*(idum0-k*IQ)-IR*k;
    if(idum0<0) idum0+=IM;
}

```

```

    ans=AM*(idum0);
    idum0^=MASK;
    idum=idum0;
    return(ans);
}

double sqr(double x){
    return(x*x);
}

double pow(double x, double y){ /* x**y */
    return(exp(y*log(x)));
}

double coef_simpson(int i){
    if((i==0)|| (i==xjmax))return(1.0);
    else if(i%2==0)return(2.0);
    else return(4.0);
}

double normal2(double media, double desv){
    double u1,u2,v1,v2,s,rr,x;
    if(normal_impar==0){
        do{
            u1=alea(idum);
            u2=alea(idum);
            v1=2*u1-1;
            v2=2*u2-1;
            s=v1*v1+v2*v2;
        }while(s>1);
        rr=sqrt((-2*log(s))/s);
        x=rr*v1;
        normal_y=rr*v2;
        normal_impar=1;
        return(x*desv+media);
    }
    else{
        normal_impar=0;
        return(normal_y*desv+media);
    }
}

double teste_ad(){ /* Teste de Normalidade de Anderson-Darling */
    double termo,soma,media,dp,aux1,aux2,pp[1000];
    int i,j;

```

```

media=aux2=0.0;
for(i=0;i<nrep;i++){
    media+=amostran[i];
    aux2+=sqr(amostran[i]);
}
media/=nrep;
dp=sqrt((aux2-nrep*sqr(media))/(nrep-1.0));
for(i=0;i<nrep;i++){
    pp[i]=erfff((amostran[i]-media)/dp);
}
termo=0.0;
for(i=1;i<=nrep;i++){
    termo+=(2.0*i-1.0)*(log(pp[i-1])+log(1.0-pp[nrep+1-i-1]));
}
termo/=nrep;
termo=-nrep-termo;
return(termo);
}

double erfff(float x){/* integral da funcao de Gauss de -inf. ate' x */
    float t,z,ans;
    z=fabs(x);
    z/=sqrt(2.0);
    t=1.0/(1.0+0.5*z);
    ans=t*exp(-z*z
-1.26551223
+t*(1.00002368
+t*(.37409196
+t*(.09678418
+t*(-.18628806
+t*(.27886807
+t*(-1.13520398
+t*(1.48851587
+t*(-.82215223
+t*.17087277
))))))));
    return x>=0.0? 1.0-ans/2.0 : ans/2.0;
}

double gauss(double x){
    return(exp(-x*x/2.0)/cgauss);
}

```

*Rotina para Geração de Números Aleatórios - Distribuição Gama:*

```
/* Algoritmo Geracao Numeros Aleatorios - Gamma */
/*****

float gamdev(int ia){
    int j;
    float am,e,s,v1,v2,x,y;

    if(ia<1) printf("Error in routine gamdev\n");
    if(ia<6){
        x=1.0;
        for(j=1;j<=ia;j++) x*=alea(idum);
        x=-log(x);
    }
    else{
        do{
            do{
                do{
                    v1=alea(idum);
                    v2=2.0*alea(idum)-1.0;
                }while(v1*v1+v2*v2>1.0);
                y=v2/v1;
                am=ia-1;
                s=sqrt(2.0*am+1.0);
                x=s*y+am;
            }while(x<=0);
            e=(1.0+y*y)*exp(am*log(x/am)-s*y);
        }while (alea(idum)>e);
    }
    return x;
}

double integral_gamma2(double x){
    return( 1.0 - (1.0+x)*exp(-x) );
}

double densidade_gamma2(double x){
    return( x*exp(-x) );
}
```

*Rotina para Geração de Números Aleatórios - Distribuição Bimodal:*

```
/* Algoritmo Geracao Numeros Aleatorios - Bimodal */
/*****/

for(jj=0;jj<nrep;jj++){/*geracao de uma amostra de nrep valores normais*/
    if(alea(idum)>0.5)x=normal2(-1.5,1);
    else x=normal2(1.5,1);
    amostran[jj]=x;
    media_amostra+=x;
    media_amostra2+=sqr(x);
    fprintf(anormal,"%g\n",x);
    quiq=(x+5)/cgraf;
    xi=(int)(tamv*quiq+0.5); if((xi<1000)&&(xi>=0))fdp[xi]++;
    if(flagg)putp(xi,(int)((470-(fdp[xi])/1.0)),2);
}
```

# Referências Bibliográficas

- [1] ALTMAN, N.; LEGER, C. *Bandwidth selection for kernel distribution function estimation*. J. Statist. Plan. Inf., 46, pag. 195-214, 1995.
- [2] ATUNCAR, G. S.; DAMASCENO, E. C. ; MENDONÇA, P. P. *Choosing the bandwidth in nonparametric functional estimation*. Relatório Técnico - Departamento de Estatística UFMG, 1998.
- [3] ATUNCAR, G. S.; OLIVEIRA, P. J. *Escolha da janela ótima em estimação funcional: caso markoviano*. Relatório de Iniciação Científica - CNPQ. Departamento de Estatística UFMG, 1999.
- [4] AZZALINI, A. *A note on the estimation of a distribution function and quantiles by a kernel method*. Biometrika, 68, pag. 326-328, 1981.
- [5] BOWMAN, A. W. *An alternative method of cross validation for the smoothing of density estimates*. Biometrika, 71, pag. 353-360, 1984.
- [6] BOWMAN, A. W.; HALL, P.; PRVAN, T. *Bandwidth selection for the smoothing of distribution functions*. Biometrika, 85, pag. 799-808, 1998.
- [7] BRILLINGER, D. R. *Time series data analysis and theory*. Holt, Rinehart and Winston. New York, 1981.
- [8] CHIU, S. T. *Bandwidth selection for kernel density estimation*. The Annals of Statistics, 33, pag. 1883-1905, 1991.

- [9] CHUNG, K. L. *A course in probability theory*, second edition, Academic Press. San Diego, 1974.
- [10] DAMASCENO, E. C. *Escolha do parâmetro de suavidade em estimação funcional*. Dissertação de mestrado. Departamento de Estatística-UFMG, 2000.
- [11] DEVROYE, L. *A course in density estimation*. Birkhauser. Boston, 1987.
- [12] DIGGLE, P. *A kernel method for smoothing point process data*. Appl. Statist., 34, pag. 138-147, 1985.
- [13] DIGGLE, P.; MARRON, J. S. *Equivalence of smoothing parameter selectors in density and intensity estimation*. Journal of the American Statistical Association, 83, pag. 793-800, 1988.
- [14] FALK, M. *Relative efficiency and deficiency of kernel type estimators of smooth distribution functions*. Statist. Neer., 37, pag. 73-83, 1983.
- [15] FERREIRA DA SILVA, J. B. G. et all. *Evaluation of the probability distribution of the strength of transmission line steel towers based on tower test results*. CIGRE-International Conference on Large High Voltage Electrical System. Report, 22, pag. 305-309. Paris, 1990.
- [16] GLICK, N. *Consistency conditions for probability estimators and integrals of density estimators*. Utilitas Mathematica, 6, pag. 61-74, 1974.
- [17] HALL, P. *Large sample optimality of least squares cross-validation in density estimation*. Ann. Statist., 11, pag. 1156-1174, 1983.
- [18] HALL, P.; MARRON, J. S. *Extend to which least-squares cross-validation minimises integrated square error in nonparametric density estimation*. Probability Theory Related Fields, 74, pag. 567-581, 1987.
- [19] HALL, P.; MARRON, J. S.; PARK, B. *Smoothed cross-validation*. Unpublished manuscript, 1989.

- [20] HILL, P. D. *Kernel estimation of a distribution function*. Communications in Statistics, Theory and Methods, 14, pag. 605-620, 1985.
- [21] HJORT, N. L.; WALKER, S. G. *A note on kernel density estimators with optimal bandwidths*. Statistics and Probability Letters, 54, pag. 153-159, 2001.
- [22] JAMES, B. R. *Probabilidade: um curso em nível intermediário*, seg. ed. Projeto Euclides - IMPA. Rio de Janeiro, 1996.
- [23] JONES, M. C. *The performance of kernel density functions in kernel distribution function estimation*. Statistics and Probability Letters, 9, pag. 129-132, 1990.
- [24] KLEIN, J. P.; MOESCHBERGER, M. L. *Survival analysis: techniques for censored and truncated data*. Springer. New York, 1997.
- [25] MARRON, J. S.; WAND, M. P. *Exact mean integrated squared error*. The Annals of Statistics, 20, pag. 712-736, 1992.
- [26] MATHEWS, J. H. *Numerical methods for mathematics, science and engineering*. Prentice Hall. Englewood Cliffs, 1992.
- [27] NADARAYA, E. A. *Some new estimates for distribution functions*. Theory Prob. Applic., 9, pag. 497-500, 1964.
- [28] PADGETT, W. J. *A kernel-type estimator of a quantile function from right-censored data*. Journal of the American Statistical Association, 39, pag. 215-222, 1986.
- [29] PARK, B.; MARRON, J. S. *Comparison of data-driven bandwidth selectors*. J. Amer. Statist. Assoc., 85, pag. 66-72, 1990.
- [30] POLANSKY, A. M. *A smooth nonparametric approach to process capability*. Division of Statistics, Northern Illinois University, 1997.
- [31] POLANSKY, A. M. *Bandwidth selection for kernel distribution function*. Division of Statistics, Northern Illinois University, 1997.

- [32] PRESS, W. H. et all. *Numerical recipes in C: the art of scientific computing*. Cambridge University Press. New York, 1988.
- [33] REISS, R. -D. *Nonparametric estimation of smooth distribution functions*. Scandinavian Journal of Statistics, 8, pag. 116-9, 1981.
- [34] RUDEMO, M. *Empirical choice of histograms and kernel density estimators*. Scandinavian Journal of Statistics, 9, pag. 65-78, 1982.
- [35] RUDIN, W. *Real and complex analysis*, third edition. McGraw-Hill. New York, 1987.
- [36] SCOTT, D. W.; FACTOR, L. E. *Monte Carlo study of three data-based nonparametric probability density estimators*. J. Amer. Statist. Assoc., 76, pag. 9-15, 1981.
- [37] SCOTT, D. W.; TERRELL, G. R. *Biased and unbiased cross-validation in density estimation*. J. Amer. Statist. Assoc., 82, pag. 1131-1146, 1987.
- [38] SHIRAHATA, S.; CHU, I. -S. *Integrated squared errors of kernel-type estimators of distribution functions*. Annals of the Institute of Statistical Mathematics, 44, pag. 579-591, 1991.
- [39] SILVERMAN, B. W. *Density estimation for statistics and data analysis*. Chapman and Hall. London, 1986.
- [40] SIMONOFF, J. F. *Smoothing methods in statistics*. Springer. New York, 1996.
- [41] STONE, C. J. *An asymptotically optimal window selection rule for kernel density estimates*. Ann. Statist., 12, pag. 1285-1297, 1984.
- [42] SUN, L. *Bandwidth choice for hazard rate estimators from left truncated data and right censored data*. Statistics and Probability Letters, 36, pag. 101-114, 1997.
- [43] SWANEPOEL, J. W. H. *Mean integrated squared error properties and optimal kernels when estimating a distribution function*. Com. Statist. Theory Methods, 17, pag. 3785-3799, 1981.

- [44] YAMATO, H. *Uniform convergence of an estimator of a distribution function*. Bull. Math. Statist., 15, pag. 69-78, 1973.
- [45] WAND, M. P.; JONES, M. C. *Kernel smoothing*. Chapman e Hall. London, 1995.
- [46] WATSON, G. S.; LEADBETTER, M. R. *Hazard analysis II*. Sankhyā Ser. A, 26, pag. 101-116, 1964.
- [47] WINTER, B. B. *Strong uniform consistency of integrals of density estimators*. Canad. J. Statist., 1, pag. 247-253, 1973.
- [48] WOODROOFE, M. *On choosing a delta sequence*. Ann. Math. Statist., 41, pag. 1665-1671, 1970.