O MÉTODO DE VALIDAÇÃO CRUZADA NA ESTIMAÇÃO DE INTENSIDADES DE PROCESSOS DE POISSON NÃO HOMOGÊNEOS

Este exemplar corresponde à redação final da tese devidamente corrigida e defendida pelo Sr. Benito Orlando Olivares Aguilera e aprovada pela Comissão Julgadora.

Campinas, 15 de julho de 1993.

Prof. Dr. Mauro S. de F. Marques OK

Dissertação apresentada ao Instituto de Matemática, Estatística e Ciência da Computação, UNICAMP, como requisito parcial para a obtenção do título de MESTRE em Estatística.

UNICAMP

Aos três vértices do triângulo:

Alvaro, Gonzalo e Bastián.

Agradecimentos

Ao Prof. Dr. Mauro S. de F. Marques, pela orientação neste trabalho.

À Profa. Dra. Irene Morsoletto Ferreira pelo incentivo a trabalhar na área dos Processos Estocásticos.

Ao CNPq e à FAEP, pelo importante apoio financeiro.

À minha esposa, Alejandra, pela compreensão e apoio.

Índice

CAPÍTULO 1 - INTRODUÇÃO. 1

CAPÍTULO 2 - PROCESSOS PONTUAIS UNIVARIADOS. 8

2.A. Definições e Resultados Básicos 9

2.B. Representação Minimal de um Processo Pontual. . . . 18

2.C. Processos de Tipo Poisson 26

2.D. Integrais Estocásticas para Processos Pontuais. . . 29

2.E. Processos com Compensador Determinístico e Contínuo 31

2.F. Alguns Exemplos de Aplicações dos Processos Pontuais. 36

CAPÍTULO 3 - ESTIMAÇÃO NÃO PARAMÉTRICA PARA CERTOS PROCESSOS PONTUAIS 44

3.A. Estimação Não Paramétrica de Funções de Densidade . 45

3.B. A Decomposição Variância-Vício Quadrático 48

3.C. Validação Cruzada de Mínimos Quadrados. 53

3.D. Estimação no Modelo de Aalen. 57

3.E. O Estimador Kernel no Modelo de Aalen 62

3.F. Validação Cruzada para Estimar a Intensidade de um Processo de Poisson Não Homogêneo 72

3.G. Otimalidade Assintótica da Janela obtida por Validação Cruzada 75

CAPÍTULO 4 - ASPECTOS PRÁTICOS NA ESTIMAÇÃO DE INTENSIDADE DE PROCESSOS DE POISSON NÃO HOMOGÊNEOS . . . 88

4.A. Simulação do Processo de Poisson Não Homogêneo. . . 88

4.B. Aspectos Práticos do Método de Validação Cruzada. . 91

4.C. Aspectos Computacionais 95

4.D. Análise dos Resultados. 96

APÊNDICE A 101

A.1. Teoria Básica de Processos 101

A.2. Tempos de Parada 103

A.3. Processos Predizíveis. 104

A.4. Tempos Predizíveis e Totalmente Inacessíveis . . . 106

A.5. Martingales. 107

A.6. Decomposição de Doob-Meyer 108

A.7. Resultados Determinísticos 110

APÊNDICE B - GRÁFICOS 112

APÊNDICE C - PROGRAMAS. 124

BIBLIOGRAFIA. 146

Notações

$\mathbf{I}(A)$: função indicadora de A.

$a \wedge b$: mínimo entre a e b.

$\sigma(X)$: σ-álgebra gerada por X.

(q. c.): quase certamente; com probabilidade um.

$\prod_{s<u\leq t}$: produto sobre (s, t].

$\blacksquare$: fim de prova.

$o(h)$: ordem de h.

$\xrightarrow{P}$: convergência em probabilidade.

Capítulo 1

Introdução

As origens da teoria moderna de processos pontuais não são fáceis de identificar, mas os problemas reais que puderam ser modelados mediante algum tipo de processo pontual são inumeráveis. Eles incluem:

i) análise de histórias de vida e problemas de renovação;

ii) problemas de contagem;

iii) física de partículas e processo populacionais; e

iv) engenharia de comunicações.

As duas primeiras classificações já podem ser encontradas em trabalhos de séculos passados e são as de nosso interesse neste trabalho. As duas últimas pertencem essencialmente a este século.

Análise de histórias de vida está ligada com intervalos entre eventos. Nesta teoria, as histórias de vida individuais são assumidas como trajetórias independentes de processos estocásticos com espaço de estados discreto. Os estados do processo correspondem às diversas categorias para um indivíduo, um componente técnico, etc.; então transições entre estados correspondem a ocorrências dos eventos de interesse.

Freqüentemente, o objeto de estudo é a taxa ou intensidade à qual eventos ocorrem, assim, um modelo estatístico para dados de histórias de vida inclui uma especificação de como as intensidades dependem do tempo e de outros fatores observáveis (covariáveis). O estudo da situação mais simples, na qual existem somente dois estados: "vivo" ou "morto" (ou "funcionando" e "não funcionando"), é chamado análise de sobrevivência ou análise de tempo de falha. Neste caso a intensidade do evento "morte" é simplesmente a função risco da distribuição do tempo de sobrevivência.

Uma situação usual em análise de sobrevivência é aquela em que não é possível observar histórias de vida completas. Este fenômeno, chamado censura pode, por exemplo, ser devido a um término planejado de um ensaio clínico ou à remoção planejada de um componente no estudo da distribuição do tempo de vida.

Uma pormenorizada revisão da teoria de análise de histórias de vida, incluindo uma suficientemente completa discussão bibliográfica é encontrada em Andersen e Borgan (1985). Outras referências importantes são Brémaud (1981), Karr (1986) e Daley e Vere-Jones (1988).

Em aplicações bioestatísticas foi muitas vezes impossível justificar determinados modelos assumidos paramétricos, daí que muitas tentativas surgiram para incorporar métodos não paramétricos à análise de histórias de vida. Pioneiros nesse esforço foram Kaplan e Meier (1958), que propuseram seu famoso estimador de produto limite para a

função distribuição de sobrevivência, que hoje é conhecido justamente como estimador de Kaplan-Meier. Geralmente, os modelos para histórias de vida tem sido formulados por meio de variáveis aleatórias e os métodos estatísticos necessários foram derivados usando resultados de variáveis aleatórias independentes e identicamente distribuídas, o que nem sempre é realista. Em análise de histórias de vida o essencial são os fenômenos aleatórios ocorrendo no tempo de forma em geral dependente, logo, parece natural estudar esta teoria em termos de processos estocásticos.

As extensões em tal sentido foram iniciadas pelo fundamental trabalho de Aalen (1975) que mostrou que a teoria dos processos pontuais proporciona um marco geral no qual os tempos de falha censurados e as observações censuradas de uma cadeia de Markov não homogênea podem ser analisados mais facilmente. Aalen estudou um estimador da intensidade acumulada, que é uma generalização daquele proposto por Nelson (1969, 1972). Para seu estudo, Aalen usa fortemente a teoria de martingales e integrais estocásticas. Cabe citar aqui dois resultados que justificam a aplicação de tais teorias ao estudo de processos pontuais. O primeiro é a decomposição de Doob-Meyer que pode ser parafraseada como:

> "A diferença entre um processo pontual e seu compensador é um martingale".

O segundo afirma que:

"A integral estocástica de um processo predizível em relação a um martingale é um martingale".

Aalen também introduz um modelo para processos pontuais chamado modelo de intensidade multiplicativa, e prova a utilidade do modelo mediante vários exemplos.

Posteriormente, Ramlau-Hansen (1893 a), baseado nas técnicas de estimação existentes para funções densidades de probabilidade e na teoria de martingales, propõe estimar a intensidade de um processo pontual com intensidade multiplicativa por meio de um alisamento através de uma função kernel.

Como no caso de densidades, o método envolve a escolha de um parâmetro de alisamento (janela) h . O valor da janela h é determinante na estimação, dado que como mostrado por Rosenblatt (1971), este valor modifica significativamente as expressões da variância e o vício, e conseqüentemente a suavidade do estimador. Assim, um valor pequeno de h produz um estimador com muito "ruído" (alta variância), enquanto um valor muito grande produz sobrealisamento (alto vício).

Diversos métodos para escolher uma janela ótima foram desenvolvidos. Um dos métodos que obteve maior atenção devido

à suas atrativas propriedades assintóticas foi o de validação cruzada por mínimos quadrados, proposto por Rudemo (1982) e Bowman (1984), para o caso de estimação de densidades. A idéia básica do método é encontrar a janela que minimiza uma certa função dos dados chamada "score"; pode ser mostrado que essa janela também minimiza assintoticamente o erro quadrático integrado.

Um forte teorema sobre a otimalidade assintótica da janela de validação cruzada é dado por Hall (1983) no contexto de estimação de densidades. Este teorema foi provado para o caso de estimação da intensidade de um processo de Poisson não homogêneo por Marron e Brooks (1991).

Neste trabalho fizemos uma revisão dos aspectos envolvidos na otimalidade assintótica da janela de validação cruzada tentando simplificar algumas provas e comprovando a eficácia do método mediante exemplos com trajetórias simuladas para o caso mais simples de modelo de intensidade multiplicativa: o processo de Poisson não homogêneo.

No Capítulo 2 é estudada a teoria básica dos processos pontuais usando o enfoque de martingales introduzido por Brémaud (1980). A decomposição de Doob-Meyer é um dos

principais resultados ali utilizados. Grande parte do capítulo é dedicada aos processos de tipo Poisson. Conclui-se o Capítulo 2 com alguns exemplos de aplicações.

O Capítulo 3 é dedicado principalmente ao estudo da estimação da intensidade de um processo de Poisson não homogêneo mediante funções kernel proposta por Ramlau-Hansen (1983 a). Também no Capítulo 3 o problema da escolha da "melhor" (em algum sentido) janela é estudado através do método de validação cruzada de mínimos quadrados e os resultados de otimalidade assintótica são dados em detalhe.

O Capítulo 4 trata dos aspectos práticos na aplicação do método de validação cruzada para estimar a intensidade de um processo de Poisson não homogêneo. Um algoritmo proposto por Lee e Kim (1991) e uma correção do estimador para efeitos de fronteira mostrada em Schuster (1985) são adaptados para a situação aqui estudada. Simulações são feitas para exemplificar graficamente a citada otimalidade para diferentes formas funcionais de intensidades e diferentes tamanhos de amostra.

Finalmente, o Apêndice A contém todos os conceitos necessários a uma leitura do presente trabalho; o Apêndice B

contém os gráficos das simulações feitas e o Apêndice C os programas computacionais utilizados.

Capítulo 2

Processos Pontuais Univariados

Um processo pontual na reta é um modelo matemático para as ocorrências de algum fenômeno em instantes aleatórios $\tau_i, i=1,2,\ldots$.

Para tais processos existem quatro descrições equivalentes que os definem:

i) uma seqüência crescente de variáveis aleatórias não negativas;

ii) um processo de contagem;

iii) uma medida aleatória de contagem; e

iv) uma seqüência de intervalos aleatórios.

Nesta dissertação somente abordaremos as duas primeiras definições. As caracterizações iii) e iv) são referidas a Karr (1986) e Daley e Vere-Jones (1988) respectivamente.

Um resultado de fundamental importância é dado neste capítulo. Ele diz que existe um processo estocástico (chamado de *compensador*) cuja diferença com o processo pontual associado é um martingale. Este resultado, conhecido como *decomposição de Doob-Meyer* permitirá posteriormente um enfoque via teoria de martingales para a análise estatística de processos pontuais.

Um outro resultado a utilizar afirma que a integral estocástica de um processo predizível em relação a um martingale é também um martingale. A relevância desse fato ficará clara no decorrer do trabalho.

2.A. DEFINIÇÕES E RESULTADOS BÁSICOS.

A seguir apresentaremos as definições e alguns resultados importantes da teoria de processos pontuais. Algumas provas foram omitidas e somente referenciadas por exigirem técnicas além do objetivo da presente dissertação. Os conceitos não usuais estão definidos no Apêndice A.

Aqui assumimos dado um espaço de probabilidade completo $(\Omega, \mathcal{F}, P)$ e uma filtragem $F = \{\mathcal{F}_t;\ t \geq 0\}$.

Um processo pontual sobre $[0,\infty)$ pode ser visto como uma seqüência de variáveis aleatórias não negativas ou como seu processo de contagem associado.

Formalmente, temos: seja $T = (\tau_n;\ n \geq 1)$ uma seqüência de tempos de parada em relação a F, possuindo as seguintes propriedades com probabilidade um:

$$\begin{aligned} &i)\ \tau_1 > 0 \\ &ii)\ \tau_n < \tau_{n+1} \ \textit{sobre}\ \{\tau_n < \infty\} \\ &iii)\ \tau_n = \tau_{n+1} \ \textit{sobre}\ \{\tau_n = \infty\}. \end{aligned} \tag{2.1}$$

DEFINIÇÃO 2.1. A seqüência de tempos de parada $T = (\tau_n;\ n\geq 1)$ satisfazendo as condições (2.1) é chamada um *processo pontual* (univariado).

O processo pontual $T = (\tau_n;\ n\geq 1)$ fica totalmente caracterizado pelo seu processo de contagem associado.

DEFINIÇÃO 2.2. O processo de contagem $N = (N_t;\ t\geq 0)$ associado a um processo pontual $T = (\tau_n;\ n\geq 1)$ e definido por

$$N_t = \sum_{n\geq 1} \mathbf{I}(\tau_n \leq t),\ t\geq 0 \qquad (2.2)$$

também é chamado de processo pontual.

Em aplicações, somente o caso em que o processo pontual é não explosivo, isto é, o limite pontual $\tau_\infty =: \lim_{n\to\infty} \tau_n$ é igual a $+\infty$, é considerado. Neste caso, $N_t < \infty$ para $t < \infty$.

A definição anterior incorre, segundo Brémaud (1981, p. 19), num "abuso de notação, mas inócuo, dado que N_t e τ_n carregam a mesma informação".

Obtém-se da definição do processo $N = (N_t;\ t\geq 0)$ que ele é adaptado a $\mathbf{F}$, isto é, N_t é $\mathcal{F}_t$ -mensurável para cada $t\geq 0$. Devido a isso usaremos também a notação $N = (N_t, \mathcal{F}_t;\ t\geq 0)$ para esse processo. A Definição 2.2 implica que $N_0 = 0$ e $N_t = N_{t-} + (1\ ou\ 0)$, ou seja, as trajetórias de N são funções do tipo escada, contínuas à direita com limites à esquerda e possuem um número enumerável de saltos de tamanho um.

As σ-álgebras $\mathcal{F}_t$ que aparecem na definição do processo de contagem podem ser interpretadas como toda a informação disponível até o instante t, isto é, elas contém todos os eventos observáveis em $[0,t]$.

Um importante exemplo de filtragem é aquela definida por $\mathbf{F}^N = (\mathcal{F}_t^N; t \geq 0)$, onde

$$\mathcal{F}_T^N = \sigma(N_s; s \leq t)$$

é conhecida na literatura como *história interna* do processo.

O processo pontual T pode ser facilmente obtido a partir do processo de contagem N fazendo

$$\tau_n = \inf\{t: N_t = n\}. \qquad (2.3)$$

A seguinte figura permite visualizar as expressões (2.2) e (2.3), onde N_t é interpretado como a contagem das ocorrências do evento de interesse em $[0,t]$ e as variáveis aleatórias τ_n correspondem aos tempos de salto do processo $N = (N_t; t \geq 0)$.

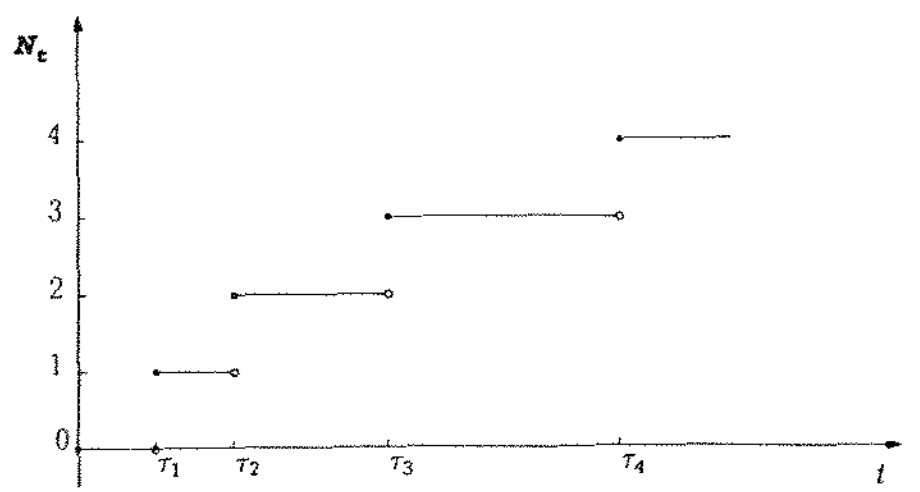


Figura 1. Relação entre processo pontual e processo de contagem.

Dois exemplos clássicos de processos pontuais de grande importância na literatura são dados a seguir. O primeiro trata de um processo cuja grande aplicabilidade em diversas áreas da ciência faz dele, talvez, o processo mais conhecido da teoria de processos pontuais.

EXEMPLO 1. (PROCESSO DE POISSON HOMOGÊNEO).

Um processo de Poisson $N = (N_t;\ t\geq 0)$ com parâmetro λ é um processo pontual cujos tempos entre saltos, definidos como $\tau_i - \tau_{i-1}$, $i = 1, 2, \ldots$ e $\tau_0 = 0$, são variáveis aleatórias i.i.d. com distribuição exponencial de parâmetro λ.

Aqui, $N_t - N_s \sim Poisson\ [\lambda(t-s)]$, $\forall\, s<t$, e $(N_t;\ t\geq 0)$ tem incrementos independentes.

Uma generalização do processo de Poisson homogêneo será dada na seção 2.C.

EXEMPLO 2. (PROCESSO FUNÇÃO DE DISTRIBUIÇÃO EMPÍRICA).

Sejam $X_1, X_2, \ldots, X_n$ variáveis aleatórias i.i.d., não negativas, com função de distribuição contínua F.

Definamos

$$N_t^{(n)} = \sum_{i=1}^{n} \mathbf{I}(X_i \leq t].$$

Então, $N^{(n)} = (N_t^{(n)};\ t\geq 0)$ é, obviamente, um processo pontual com tempos de salto $\tau_i = X_{(i)}$, $i = 1, \ldots, n$ e $\tau_i = \infty$ para $i = n+1, \ldots$, onde $X_{(i)}$ é a i-ésima estatística de ordem da amostra $X_1, X_2, \ldots, X_n$.

Podemos escrever $N_t^{(n)} = nF_n(t)$, onde

$$F_n(t) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{I}(X_i \leq t)$$

é a conhecida função de distribuição empírica ou amostral, muito utilizada na teoria estatística de estimação.

Dado que as trajetórias de um processo pontual $N = (N_t;\ t \geq 0)$ são não decrescentes e que o processo é adaptado, obtém-se

$$E(N_t/\mathcal{F}_s) \geq E(N_s/\mathcal{F}_s) = N_s,\ \forall\ s \leq t,$$

ou seja, N é um submartingale. Também, N é localmente limitado: $N_{\tau_n} \leq n,\ \forall n$, e portanto estão satisfeitas as condições do Teorema A.2 (N é um submartingale de classe D, segundo a Definição A.1).

Temos assim a seguinte decomposição de Doob-Meyer (D-M) para um processo pontual.

TEOREMA 2.1. Um processo pontual $N = (N_t, \mathcal{F}_t;\ t \geq 0)$ admite, para todo $t < \tau_\infty$ a única decomposição

$$N_t = M_t + A_t, \qquad (2.4)$$

onde $M = (M_t, \mathcal{F}_t;\ t < \tau_\infty)$ é um martingale (local) e $A = (A_t, \mathcal{F}_t;\ t \geq 0)$ é um processo crescente predizível.

PROVA. Ver Liptser e Shiryayev (1977, p. 238).

A importância do Teorema 2.1 reside na propriedade de martingale do processo ($N-A$), o que permite a aplicação

formal desta teoria ao estudo dos processos pontuais.

O próximo exemplo mostra a decomposição de Doob-Meyer para o processo de Poisson.

EXEMPLO 3.

Seja $N = (N_t, \mathcal{F}_t^N;\ t \geq 0)$ um processo de Poisson de parâmetro $\lambda > 0$.

Pode-se escrever $N_t = (N_t - \lambda t) + \lambda t$ e prova-se facilmente que o processo $(N_t - \lambda t, \mathcal{F}_t^N;\ t \geq 0)$ é um martingale. Para tanto, usa-se a seguinte igualdade: $N_t - \lambda t = (N_t - N_s) + N_s - \lambda(t-s) - \lambda s$ e a propriedade de incrementos independentes.

Agora o processo $(\lambda t;\ t \geq 0)$, por ser determinístico, é predizível e claramente é crescente. Então, do Teorema 2.1 temos que $M_t = N_t - \lambda t$ e $A_t = \lambda t$ é a decomposição D-M para o processo de Poisson.

O processo A_t pode ser escrito como

$$A_t = \int_0^t \lambda ds,$$

onde λ pode ser interpretado como a ocorrência média de pontos em $[0,1]$, ou

$$\lambda = \lim_{\Delta t \to 0} \frac{1}{\Delta t} P\,(N_{t+\Delta t} - N_t = 1).$$

Este resultado será generalizado na Seção 2.C.

DEFINIÇÃO 2.3. O processo crescente e predizível $\mathbf{A} = (\mathbf{A}_t, \mathcal{F}_t;\ t \geq 0)$ que aparece na decomposição D-M (2.4) é chamado o *compensador* do processo pontual $\mathbf{N} = (\mathbf{N}_t, \mathcal{F}_t;\ t \geq 0)$.

Dado que as trajetórias do compensador $\mathbf{A} = (\mathbf{A}_t,\ t \geq 0)$ são, com probabilidade um, funções não decrescentes, o número de saltos de $\mathbf{A}_t,\ t \geq 0$, é finito ou enumerável. Também pode-se mostrar que a magnitude dos saltos não excede de um [Lema 18.8, Liptser e Shiryayev (1977)], isto é,

$$\sup_{t \leq \tau_\infty} \Delta \mathbf{A}_t \leq 1, \tag{2.5}$$

onde $\Delta \mathbf{A}_t = \mathbf{A}_t - \mathbf{A}_{t-}$.

Uma classe importante de processos pontuais é aquela cujo compensador é contínuo. A estrutura de tais processos é descrita no seguinte lema.

LEMA 2.1. Uma condição necessária e suficiente para que o compensador $(\mathbf{A}_t,\ t \geq 0)$ de um processo pontual $\mathbf{N} = (\mathbf{N}_t, \mathcal{F}_t;\ t \geq 0)$ seja (q.c.)-contínuo em $[0, \tau_\infty)$ é que o processo seja quase-contínuo à esquerda [ver Definição A.8].

PROVA. A condição de quase-continuidade à esquerda, no presente contexto, quer dizer que para qualquer seqüência não decrescente de tempos de parada $(\sigma_n; n \geq 1)$, tem-se $\lim_n \mathbf{N}_{\sigma_n \wedge \tau_\infty} = \mathbf{N}_{\lim_n \sigma_n \wedge \tau_\infty}$ (q.c.). [Ver também a Proposição A.3].

NECESSIDADE: Seja $\sigma = \lim_n \sigma_n$. Então, por (2.4)

$$E[N_{\tau_k \wedge \sigma} - A_{\tau_k \wedge \sigma}] = E[N_{\tau_k \wedge \sigma_n} - A_{\tau_k \wedge \sigma_n}] .$$

Logo,

$$E[N_{\tau_k \wedge \sigma} - N_{\tau_k \wedge \sigma_n}] = E[A_{\tau_k \wedge \sigma} - A_{\tau_k \wedge \sigma_n}] .$$

Como ambas as seqüências $(N_{\tau_k \wedge \sigma} - N_{\tau_k \wedge \sigma_n};\ n \geq 1)$ e $(A_{\tau_k \wedge \sigma} - A_{\tau_k \wedge \sigma_n};\ n \geq 1)$ são limitadas, pode-se aplicar o Teorema da Convergência Dominada para obter:

$$E[\lim_n (N_{\tau_k \wedge \sigma} - N_{\tau_k \wedge \sigma_n}) = E[\lim_n (A_{\tau_k \wedge \sigma} - A_{\tau_k \wedge \sigma_n})] = 0 ,$$

pela continuidade de A_t, $t \geq 0$.

Logo, temos que $N_{\tau_k \wedge \sigma} = \lim_n N_{\tau_k \wedge \sigma_n}$ (q.c.) e isto implica em

$$N_{\tau_\infty \wedge \sigma} = \lim_k \lim_n N_{\tau_k \wedge \sigma_n}$$

$$= \lim_n \lim_k N_{\tau_k \wedge \sigma_n} = \lim_n N_{\tau_\infty \wedge \sigma_n} ,$$

o qual prova a necessidade.

SUFICIÊNCIA. Consideremos o processo $\Pi^{(k)} = (\Pi_t^{(k)}, \mathcal{F}_t;\ t \geq 0)$ onde

$$\Pi_t^{(k)} = E(N_{\tau_k} / \mathcal{F}_t) - N_{t \wedge \tau_k} .$$

É fácil verificar que $\Pi^{(k)}$ é um supermartingale, sabendo que N é um submartingale.

Devido à quase-continuidade do processo N ,

$$E\left[\Pi^{(k)}_{\sigma_n}\right] = E\left[N_{\tau_k}\right] - E\left[N_{\sigma_n \wedge \tau_k}\right] \xrightarrow[n \to \infty]{} E\left[N_{\tau_k}\right] - E\left[N_{\sigma \wedge \tau_k}\right] =$$

$$= E\left[E(N_{\tau_k}/\mathcal{F}_\sigma) - N_{\sigma \wedge \tau_k}\right]$$

$$= E\left[\Pi^{(k)}_\sigma\right]$$

Pelo Teorema 3.11 em Liptser e Shiryayev (1978), $\Pi^{(k)}$ admite uma decomposição D-M da forma

$$\Pi^{(k)}_t = M^{(k)}_t + B^{(k)}_t,$$

onde $B^{(k)} = (B^{(k)}_t, \mathcal{F}_t;\ t \geq 0)$ é um processo predizível contínuo. Decorre da unicidade da decomposição D-M que $B^{(k)}_t = -A_{t \wedge \tau_k}$. Aqui, para $t < \tau_\infty$ o compensador A_t tem trajetórias (q.c.)-contínuas. Isto junto com a igualdade $A_{\tau_\infty} = A_{\tau_\infty -}$ prova que $P(\Delta A_t \neq 0,\ t \leq \tau_\infty) = 0$, ou seja, A é (q.c.)-contínuo.

■

A equivalência formulada na Proposição A.3 que relaciona quase-continuidade à esquerda e tempos totalmente inacessíveis, permite deduzir o seguinte corolário.

COROLÁRIO. O compensador $(A_t,\ t \geq 0)$ do processo pontual $N = (N_t,\ \mathcal{F}_t;\ t \geq 0)$ é (q.c.)-contínuo em $[0, \tau_\infty)$ se e somente se os tempos de salto do processo $(N_t,\ t \geq 0)$, são totalmente inacessíveis.

2.B. REPRESENTAÇÃO MINIMAL DE UM PROCESSO PONTUAL.

Seja $N = (N_t, \mathcal{F}_t;\ t \geq 0)$ um processo pontual. Sabe-se que as variáveis N_t são $\mathcal{F}_t$ -mensuráveis, mas elas podem ser também mensuráveis com relação a σ-álgebras menores como por exemplo a σ-álgebra $\mathcal{F}_t^N = \sigma(N_s;\ s \leq t)$ definida na Seção 2.A.

É claro que a filtragem $F^N = \{\mathcal{F}_t^N;\ t \geq 0\}$ é a "menor" família de σ-álgebras com relação à qual as variáveis N_t, $t \geq 0$, são mensuráveis.

Com essa filtragem pode-se definir o processo $\bar{N} = (N_t, \mathcal{F}_t^N;\ t \geq 0)$, que é também um processo pontual. Pelo Teorema 2.1, tem-se a correspondente decomposição de Doob-Meyer

$$N_t = \bar{M}_t + \bar{A}_t, \quad t < \tau_\infty \tag{2.6}$$

Tal decomposição é chamada a *representação minimal* do processo N .

Um problema importante na teoria de processos pontuais é a determinação explícita do compensador do processo dada uma certa filtragem. Uma diversidade de métodos para tal efeito foram desenvolvidos; uma discussão detalhada destes pode ser encontrada em Gill (1980).

O seguinte teorema pode ser usado para calcular o compensador em termos das distribuições condicionais regulares e mostrar a relação estreita entre o compensador e a função risco.

TEOREMA 2.2. Seja $F_1(t) = P(\tau_1 \leq t)$ e sejam $F_i(t) = P(\tau_i \leq t/\tau_{i-1}, \ldots, \tau_1)$, $i \geq 2$ as funções de distribuição condicional regular. Então a fórmula

$$\bar{A}_t = \sum_{i \geq 1} \bar{A}_t^{(i)} \tag{2.7}$$

onde

$$\bar{A}_t^{(i)} = \int_0^{t \wedge \tau_i} \frac{dF_i(u)}{1-F_i(u-)}, \quad i \geq 1. \tag{2.8}$$

define o compensador $\bar{A} = (\bar{A}_t, \mathcal{F}_t^N;\ t < \tau_\infty)$ do processo pontual $\bar{N} = (N_t, \mathcal{F}_t^N;\ t \geq 0)$.

Note-se que no caso das F_i serem absolutamente contínuas, o integrador em (2.8) coincide com a *função risco* utilizada em análise de sobrevivência e definida como

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

onde f é a densidade correspondente à função de distribuição F.

Para provar o Teorema 2.2, precisamos de um Lema prévio.

LEMA 2.2. Seja $N_t = \sum_{n \geq 1} \mathbf{I}(\tau_n \leq t)$, $t \geq 0$ um processo pontual e seja $\theta = \theta(\omega)$ um tempo de parada em relação a F^N tal que $P(\theta < \tau_\infty) = 1$. Então existem funções mensuráveis $\varphi_n = \varphi_n(t_1, \ldots, t_n)$, $n \geq 1$ e uma constante φ_0 tal que

$$\theta(\omega) = \sum_{n\geq 1} \varphi_{n-1}(\tau_1, \ldots, \tau_{n-1}) \mathbf{I}(\tau_{n-1} \leq \theta \leq \tau_n) \quad (com\ \tau_0 \equiv 0) . \quad (2.9)$$

PROVA. Usaremos o fato que a σ-álgebra $\mathcal{F}_\theta^N$ coincide com a σ-álgebra $\sigma(N_{\theta\wedge t};\ t \geq 0)$. Definamos a σ-álgebra $\mathcal{F}_\theta^N = \sigma(N_{\theta\wedge t};\ t \geq 0)$ para qualquer tempo de parada θ com relação a $(\mathcal{F}_s^N; s \geq 0)$.

Aqui, θ é $\mathcal{F}_\theta^N$-mensurável e existirá um conjunto contável $S \subset [0, \infty)$ e uma função mansurável $\varphi(X_n; n \geq 1)$ tal que

$$\theta(\omega) = \varphi(N_{t\wedge\theta(\omega)}(\omega);\ t \in S);$$

então,

$$\theta(\omega) = \sum_{n\geq 1} \varphi(N_{t\wedge\theta}; t \in S) \mathbf{I}(\tau_{n-1} \leq \theta \leq \tau_n) .$$

Notemos agora que no conjunto $\{\tau_{n-1} \leq \theta \leq \tau_n\}$

$$N_{t\wedge\theta} = N_{t\wedge\tau_{n-1}} = \sum_{k=1}^{n-1} \mathbf{I}(\tau_k \leq t)$$

e conseqüentemente nesse conjunto

$$\varphi(N_{t\wedge\theta}; t \in S) = \varphi\left(\sum_{k=1}^{n-1} \mathbf{I}(\tau_k \leq t);\ t \in S\right).$$

A função do lado direito na última igualdade pode ser representada como $\varphi_{n-1}(\tau_1, \ldots, \tau_{n-1})$, onde φ_{n-1} é uma função mensurável de n-1 variáveis. Logo,

$$\theta(\omega) = \sum_{n\geq 1} \varphi(N_{t\wedge\theta};\, t\in S)\, \mathbb{I}(\tau_{n-1}\leq\theta\leq\tau_n)$$

$$= \sum_{n\geq 1} \varphi_{n-1}(\tau_1, \ldots, \tau_{n-1})\, \mathbb{I}(\tau_{n-1}\leq\theta\leq\tau_n)$$

■

COROLÁRIO. Existem funções mensuráveis $\theta_1, \theta_2, \ldots$ e uma constante θ_0 tal que, no conjunto $\{\theta<\tau_n\}$ tem-se $\theta = \theta_{n-1}(\tau_1, \ldots, \tau_{n-1})$; $n\geq 1$. Em particular, $\theta\wedge\tau_k = \theta_{n-1}\wedge\tau_k$, $k\leq n$.

PROVA DO TEOREMA 2.2.

Pela decomposição de Doob-Meyer, é suficiente provar que para cada n, $n = 1, 2, \ldots$, os processos

$$M^{(n)} = (N_{t\wedge\tau_n} - \bar{A}_{t\wedge\tau_n},\ \mathcal{F}_t^N;\ t\geq 0)$$

onde $\bar{A}_t$ é definida por (2.7) e (2.8), são martingales uniformemente integráveis e que o processo $\bar{A} = (\bar{A}_t, \mathcal{F}_t^N;\ t\geq 0)$ é predizível.

Para provar a propriedade de martingale de $M^{(n)}$ usamos a Proposição A.4; logo, é suficiente estabelecer que para qualquer $\mathcal{F}_t^N$ -tempo de parada θ ,

$$E(N_{\theta\wedge\tau_n} - \bar{A}_{\theta\wedge\tau_n}) = 0$$

ou equivalentemente,

$$E(N_{\theta\wedge\tau_n}) = E(\bar{A}_{\theta\wedge\tau_n})\,. \qquad (2.10)$$

Temos:

$$E\left[\bar{A}_{\theta\wedge\tau_n}\right] = E\left[\sum_{i=1}^{n}\bar{A}^{(i)}_{\theta\wedge\tau_n}\right]$$

$$= E\left[\sum_{i=1}^{n}\bar{A}^{(i)}_{\theta\wedge\tau_i}\right] = E\left[\sum_{i=1}^{n}\bar{A}^{(i)}_{\theta_{i-1}\wedge\tau_i}\right]$$

$$= E\left[\sum_{i=1}^{n}E\left[\bar{A}^{(i)}_{\theta_{i-1}\wedge\tau_i}/\tau_{i-1},\ldots,\tau_1\right]\right] \tag{2.11}$$

onde θ_{i-1} é definido como no Corolário do Lema 2.2.

Da fórmula (2.8),

$$E\left(\bar{A}^{(i)}_{\theta_{i-1}\wedge\tau_i}/\tau_{i-1},\ldots,\tau_1\right) = \int_0^{\infty}\left[\int_0^{\theta_{i-1}\wedge s}\frac{dF_i(u)}{1-F_i(u-)}\right]dF_i(s)$$

$$= \int_0^{\theta_{i-1}}\left[\int_0^{s}\frac{dF_i(u)}{1-F_i(u-)}\right]dF_i(s) + \left[1-F_i(\theta_{i-1})\right]\int_0^{\theta_{i-1}}\frac{dF_i(u)}{1-F_i(u-)} \tag{2.12}$$

A seguinte igualdade é uma fórmula de integração por partes para integrais de Stieltjes [ver Teorema A.4.].

$$Z(t)B(t) = \int_0^t Z(u-)\,dB(u) + \int_0^t B(u)\,dZ(u)\,.$$

Aplicando esta expressão a

$$Z(s) = F_i(s)$$

e

$$B(s) = \int_0^s \frac{dF_i(u)}{1-F_i(u-)},$$

obtemos

$$\int_0^{\theta_{i-1}} \left[\int_0^s \frac{dF_i(u)}{1-F_i(u-)} \right] dF_i(s) = F_i(\theta_{i-1}) \int_0^{\theta_{i-1}} \frac{dF_i(u)}{1-F_i(u-)} - \int_0^{\theta_{i-1}} \frac{F_i(u-)\,dF_i(u)}{1-F_i(u-)}$$

$$= [F_i(\theta_{i-1}) - 1] \int_0^{\theta_{i-1}} \frac{dF_i(u)}{1-F_i(u-)} + \int_0^{\theta_{i-1}} [1-F_i(u-)] \frac{dF_i(u)}{1-F_i(u-)}$$

$$= [F_i(\theta_{i-1}) - 1] \int_0^{\theta_{i-1}} \frac{dF_i(u)}{1-F_i(u-)} + [F_i(\theta_{i-1}) - F_i(0)].$$

(2.13)

Substituindo (2.13) em (2.12) e usando o fato que $F_i(0) = 0$, $i=1,2,\ldots$, resulta

$$E\left(\overline{A}^{(i)}_{\theta_{i-1}\wedge\tau_i} / \tau_{i-1}, \ldots, \tau_1\right) = F_i(\theta_{i-1}). \qquad (2.14)$$

Então, substituindo em (2.11), temos que

$$E\left(\overline{A}_{\theta_{i-1}\wedge\tau_i}\right) = E\left(\sum_{i=1}^{n} F_i(\theta_{i-1})\right). \qquad (2.15)$$

Por outra parte, segue do Corolário ao Lema 2.2 que $\theta\wedge\tau_i = \theta_{i-1}\wedge\tau_i$ e $\theta\wedge\tau_{i-1} = \theta_{i-1}\wedge\tau_{i-1}$. Logo, podemos escrever

$$
\begin{aligned}
E\ (N_{\theta\wedge\tau_n}) &= E\left(N_{\theta\wedge\tau_1}+\left[N_{\theta\wedge\tau_2}-N_{\theta\wedge\tau_1}\right]+\cdots+\left[N_{\theta\wedge\tau_n}-N_{\theta\wedge\tau_{n-1}}\right]\right)\\
&= E\left(N_{\theta\wedge\tau_1}\right)+E\left(\sum_{i=2}^{n}\left[N_{\theta\wedge\tau_i}-N_{\theta\wedge\tau_{i-1}}\right]\right)\\
&= E\left(N_{\theta_0\wedge\tau_1}+E\left(\sum_{i=2}^{n}\left[N_{\theta_{i-1}\wedge\tau_i}-N_{\theta_{i-1}\wedge\tau_{i-1}}\right]\right)\right)\\
&= E\left[\mathbf{I}(\tau_1\le\theta_0)\right]+E\left(\sum_{i=2}^{n}\left[\sum_{k=1}^{i}\mathbf{I}(\tau_k\le\theta_{i-1})-\sum_{k=1}^{i-1}\mathbf{I}(\tau_k\le\theta_{i-1})\right]\right)\\
&= E\left[\mathbf{I}(\tau_1\le\theta_0)\right]+E\left(\sum_{i=2}^{n}\mathbf{I}(\tau_i\le\theta_{i-1})\right)\\
&= F_1(\theta_0)+\sum_{i=2}^{n}E\left[F_i(\theta_{i-1})\right]
\end{aligned}
\tag{2.16}
$$

Comparando esta última expressão com (2.15) obtemos a propriedade de martingale de $M^{(n)}$, isto é, $E\left(N_{\theta\wedge\tau_n}-\bar{A}_{\theta\wedge\tau_n}\right)=0$.

Da desigualdade

$$
\left|N_{t\wedge\tau_n}-\bar{A}_{t\wedge\tau_n}\right| \le N_{t\wedge\tau_n}+\bar{A}_{t\wedge\tau_n} \le n+\bar{A}_{\tau_n}
$$

e de (2.10), obtemos que $E(\bar{A}_{\tau_n}) = E(N_{\tau_n}) \le n$. Portanto, o martingale $M^{(n)}$ é uniformemente integrável.

Só falta verificar que o processo $\bar{A}_t = \sum_{i\ge1}\bar{A}_t^{(i)}$, $t\ge0$ é predizível. Para este fim é suficiente verificar que cada um dos processos $\bar{A}_t^{(i)}$, $i = 1,2,\ldots$ é predizível.

Seja para cada i, $i = 1,2,\ldots$, $\tau_i(t,\omega) \equiv t\wedge\tau_i(\omega)$. Pelo Teorema A.1, vemos que o conjunto aleatório

$$\{(t,\omega : \tau_i(t,\omega) > x\} = \{t > x\} \times \{\omega : \tau_i(\omega) > x\}$$

tem a forma de um gerador da σ-álgebra predizível; logo, $\tau_i(t,\omega)$ é predizível. A função $\overline{A}_t^{(i)}$ é monótona crescente e contínua à direita; assim, tem uma única inversa $(\overline{A}_t^{(i)})^{-1}$ para a qual $(\overline{A}_t^{(i)})^{-1} \geq y$ se e somente se $t \geq (\overline{A}_y^{(i)})^{-1}$.

Em particular $\{\tau_i(t,\omega) \geq (\overline{A}_y^{(i)})^{-1}\}$ é predizível. Portanto, o processo $\overline{A}_t^{(i)}$ é predizível, $i = 1, 2, \ldots$.

■

COROLÁRIO. Se as funções $F_i(t)$, $i \geq 1$ são (q.c.)-contínuas (resp. absolutamente contínuas), o compensador $(\overline{A}_t,\ t \geq 0)$ tem trajetórias (q.c.)-contínuas (resp. absolutamente contínuas).

O exemplo que segue mostra que a escolha da filtragem pode afetar drasticamente a forma do compensador.

EXEMPLO 4.

Seja $W = (W_t, \mathcal{F}_t;\ t \geq 0)$ um processo de Wiener e $\tau = inf\ \{t : W_t = 1\}$. Consideremos o processo pontual $(N = N_t;\ t \geq 0)$ com relação à filtragem $F = \{\mathcal{F}_t;\ t \geq 0\}$, onde $N_t = \mathbb{I}(\tau \leq t)$. A Definição A.6 mostra que N é predizível e como N é um processo crescente, N é o seu próprio compensador, isto é, na decomposição D-M, $N_t = A_t + M_t$ com $A_t = N_t$ e $M_t = 0$, $t \geq 0$.

Consideremos agora a representação minimal de $N = (N_t;\ t \geq 0)$, ou seja, o processo $\overline{N} = (N_t, \mathcal{F}_t^N;\ t \geq 0)$. Notar

que só mudamos a filtragem, que agora é $\mathcal{H} = \{\mathcal{F}_t^N; t \geq 0\}$, com $\mathcal{F}_t^N = \sigma(N_s; s \leq t)$.

Dado que

$$F(t) = P(\tau \leq t) = \sqrt{\frac{2}{\pi}} \int_{t^{-\frac{1}{2}}}^{\infty} e^{-\frac{y^2}{2}} dy$$

[ver Liptser e Shiryayev (1977), p. 32], usando o Teorema 2.2 obtemos que

$$\bar{A}_t = \int_0^{t \wedge \tau} \frac{dF(u)}{1-F(u)} = -\ln[1-F(t \wedge \tau)]$$

$$\bar{A}_t = -\ln\left[1-\sqrt{\frac{2}{\pi}} \int_{(t \wedge \tau)^{-\frac{1}{2}}}^{\infty} e^{\frac{-y^2}{2}} dy\right].$$

Ressaltamos aqui a natureza diferente dos dois compensadores: enquanto o F-compensador, A_t , é uma função descontínua, o $\mathcal{H}$-compensador, $\bar{A}_t$, é uma função absolutamente contínua.

2.C. PROCESSOS DO TIPO POISSON.

Dentro da classe dos processos pontuais, existem alguns de extrema importância que possuem uma estrutura relativamente simples, permitem elegantes formalizações de suas propriedades e são muito utilizados em aplicações.

Tais processos são caracterizados por serem seus compensadores absolutamente contínuos em relação a uma medida de Lebesgue-Stieltjes, isto é, o compensador $\mathbf{A} = (A_t, \mathcal{F}_t; t \geq 0)$ tem a forma

$$A_t = \int_0^t \lambda_s db(s) \qquad (2.17)$$

onde $\lambda = (\lambda_t, \mathcal{F}_t; t \geq 0)$ é um processo estocástico predizível não negativo e $b(t)$ é uma função determinística não negativa, não decrescente e contínua à direita.

Aqui, a decomposição de Doob-Meyer diz que

$$M_t = N_t - \int_0^t \lambda_s db(s) \qquad (2.18)$$

é um martingale local.

DEFINIÇÃO 2.4. Os processos pontuais cujos compensadores satisfazem à fórmula (2.17) são chamados *processos do tipo Poisson* e $\lambda = (\lambda_t, \mathcal{F}_t; t \geq 0)$ é denominado *processo intensidade* de N.

No caso particular $\lambda_t \equiv \lambda$, $b(t) = t$ obtemos o processo de Poisson homogêneo.

Aalen (1978) define o processo intensidade como

$$\lambda_t = \lim_{\Delta t \to 0} \frac{1}{\Delta t} E\left(N_{t+\Delta t} - N_t / \mathcal{F}_t\right)$$

ou equivalentemente

$$\lambda_t = \lim_{\Delta t \to 0} \frac{1}{\Delta t} P\ (N_{t+\Delta t} - N_t = 1 / \mathcal{F}_t)\ .$$

Uma das vantagens dos processos do tipo Poisson é que neles torna-se simples o problema não trivial de calcular o compensador $\overline{A}_t$ da representação minimal (2.6) a partir do compensador $A = (A_t, \mathcal{F}_t; t \geq 0)$ do processo pontual $N = (N_t, \mathcal{F}_t; t \geq 0)$.

Esse é o resultado do seguinte teorema.

TEOREMA 2.3. Seja $N = (N_t, \mathcal{F}_t; t \geq 0)$ um processo do tipo Poisson, isto é, seu compensador é da forma $A_t = \int_0^t \lambda_s db(s)$ e seja $N_t = \overline{A}_t + \overline{M}_t$ sua representação minimal.

Então,

$$\overline{A}_t = \int_0^t \overline{\lambda}_s db(s)\ , \tag{2.19}$$

onde

$$\overline{\lambda}_t = E\ (\lambda_t / \mathcal{F}_t^N)\ .$$

Neste caso,

$$A_t = \int_0^t \lambda_s \overline{\lambda}_s^{+} d\overline{A}_s \quad (P\text{-q.c.}) \text{ em } [0, \tau_\infty) \tag{2.20}$$

onde

$$\overline{\lambda}_t^+ = \begin{cases} (\overline{\lambda}_t)^{-1}, & \overline{\lambda}_t > 0 \\ 0 \quad , & \overline{\lambda}_t = 0. \end{cases}$$

IDÉIA DA PROVA. A prova aparece em detalhe em Liptser e Shiryayev (1977, p. 250) e consiste basicamente em mostrar os três resultados a seguir:

$$i) \int_0^t \overline{\lambda}_s db(s) < \infty \quad (q.c.) \text{ em } [0, \tau_\infty)$$

$$ii) \; B_t = \int_0^t \overline{\lambda}_s db(s) \text{ é predizível}$$

$$iii) \; B_t = \overline{A}_t \; (q.c.).$$

COROLÁRIO. Se A_t é uma função determinística, então $\overline{A}_t = A_t$.

PROVA. Direta, fazendo $b(t) = A_t$ e $\lambda_t \equiv 1$.

■

A relação (2.20) pode não ser verdadeira sem a hipótese $A_t = \int_0^t \lambda_s db s$. De fato, no Exemplo 4 mostra-se que A_t é descontínua enquanto $\overline{A}_t$ é absolutamente contínua, logo (2.20) não é satisfeita.

2.D. INTEGRAIS ESTOCÁSTICAS PARA PROCESSOS PONTUAIS.

Para completar a base teórica necessária ao nosso

estudo dos processos pontuais, introduziremos os processos aleatórios conhecidos como integrais estocásticas, cuja definição geral e análise de suas propriedades podem ser encontradas em Meyer (1976) e Elliot (1982).

Em nosso caso, trabalharemos só com integrais da forma

$$\int_0^t f_s(\omega)\, dN_s \qquad ou \qquad \int_0^t f_s(\omega)\, d\Lambda_s,$$

onde $f = (f_t, \mathcal{F}_t; t \geq 0)$ é um processo não negativo de uma classe a ser explicitada posteriormente. N e Λ são o processo pontual e seu compensador respectivamente.

Dado que tanto $N = (N_t; t \geq 0)$ quanto $\Lambda = (\Lambda_t; t \geq 0)$ são processos de variação limitada, as integrais acima podem ser consideradas simplesmente integrais de Lebesgue-Stieltjes em $(0, t]$. Assim, por exemplo, $\int_0^t f_s dNs$ é a representação de Lebesgue-Stieltjes da soma dos valores de f nos tempos de salto de N , no intervalo $(0, t]$.

A integral estocástica de um processo predizível com relação a um martingale é um martingale [ver Teorema A.3]. Esse fato será de grande utilidade na análise estatística dos processos pontuais, a ser estudada no próximo capítulo, pois as estatísticas que aparecem tem uma representação integral com respeito a processos pontuais e a martingales, que permitirá o uso de fórmulas para momentos e de um teorema central do limite para martingales a fim de descrever suas propriedades.

2.E. PROCESSOS COM COMPENSADOR DETERMINÍSTICO E CONTÍNUO.

Um tipo de processos que aparece freqüentemente em aplicações é aquele cujo compensador é contínuo e determinístico, por exemplo o processo de Poisson. A estrutura de tais processos é apresentada a seguir.

Seja $(N = N_t, \mathcal{F}_t; t \geq 0)$ um processo pontual tendo compensador determinístico $(A_t; t \geq 0)$. Denotemos $\sigma_\infty = inf\{t \geq 0 : A_t = \infty\}$, com a convenção $\sigma_\infty = \infty$ se $\lim_{t \to \infty} A_t < \infty$.

TEOREMA 2.5. O processo pontual $N = (N_t, \mathcal{F}_t; t < \sigma_\infty)$ com compensador determinístico é um processo com incrementos independentes e (q.c.) para $s \leq t < \sigma_\infty$, $\lambda \in \mathbb{R}$ a função característica condicional é dada por

$$E\left(\exp\left[i\lambda(N_t - N_s)\right] / \mathcal{F}_s\right) = \prod_{s < u \leq t} \left[1 + (e^{i\lambda} - 1)\Delta A_u\right] \exp\left[(e^{i\lambda} - 1)\left(A_t^c - A_s^c\right)\right], \tag{2.21}$$

onde

$$A_t^c = A_t - \sum_{s \leq t} \Delta A_s.$$

PROVA. Pode-se escrever

$$
\begin{aligned}
e^{i\lambda N_t} - e^{i\lambda N_s} &= \sum_{s<u\le t} \left[e^{i\lambda N_u} - e^{i\lambda N_{u-}} \right] \\
&= \sum_{s<u\le t} \left[e^{i\lambda (N_{u-}+1)} - e^{i\lambda N_{u-}} \right] \\
&= \sum_{s<u\le t} e^{i\lambda N_{u-}} (e^{i\lambda} - 1)(N_u - N_{u-}) \\
&= (e^{i\lambda} - 1) \int_s^t e^{i\lambda N_{u-}} dN_u \\
&= (e^{i\lambda} - 1) \int_s^t e^{i\lambda N_{u-}} d[N_u - \mathbb{A}_u] + (e^{i\lambda} - 1) \int_s^t e^{i\lambda N_{u-}} d\mathbb{A}_u ,
\end{aligned}
\tag{2.22}
$$

onde na penúltima igualdade usamos o fato que $\Delta N_u = N_u - N_{u-} = 0$ *ou* 1 .

Por ser $e^{i\lambda N_{u-}}$ predizível e $(N_u - \mathbb{A}_u)$ um martingale, temos que a primeira integral é um martingale uniformemente integrável para $t < \sigma_m$.

Logo,

$$
E\left(\int_s^t e^{i\lambda N_{u-}} d[N_u - \mathbb{A}_u] / \mathcal{F}_s \right) = 0 \quad (q.c.), \quad s \le t < \sigma_m . \tag{2.23}
$$

Façamos $V_s(t) = E\left(e^{i\lambda N_t} / \mathcal{F}_s\right)$, $s \le t$ e notemos que decorre do Teorema da Convergência Dominada que $V_s(t) = E\left(e^{i\lambda N_{t-}} / \mathcal{F}_s\right)$. Com isto, encontramos de (2.22) e (2.23) que

$$V_s(t) = V_s(s) + (e^{i\lambda} - 1) \int_s^t V_s(u-)\, d\Lambda_u \,. \qquad (2.24)$$

Dado que $V_s(t) \neq 0$, podemos definir

$$U_s(t) = \frac{V_s(t)}{V_s(s)} = E\left(e^{i\lambda(N_t - N_s)} / \mathcal{F}_s\right).$$

Obtêm-se assim a equação

$$U_s(t) = 1 + (e^{i\lambda} - 1) \int_s^t U_s(u-)\, d\Lambda_u \,, \qquad (2.25)$$

a qual, pelo Teorema A.5, tem a solução dada em (2.21).

■

COROLÁRIO. Se $\Lambda_t = t$, o processo pontual $N = (N_t, \mathcal{F}_t; t \geq 0)$ é um processo de Poisson e

$$E\left(e^{i\lambda(N_t - N_s)} / \mathcal{F}_s\right) = \exp\left[(e^{i\lambda} - 1)(t - s)\right].$$

Daremos a seguir um teorema que ressalta a relevância dos processos de Poisson dentro da teoria dos processos pontuais e que é de grande importância na simulação de processos pontuais. O teorema afirma que todo processo pontual $N = (N_t, \mathcal{F}_t; t \geq 0)$ pode ser transformado num processo de Poisson mediante uma adequada mudança na escala de tempo.

TEOREMA 2.6. Sejam $\sigma(t) = \inf\{s \geq 0 : \mathbb{A}_s > t\}$, $\Pi_t = N_{\sigma(t)}$ e a σ-álgebra $\mathcal{E}_t = \mathcal{F}_{\sigma(t)}$. Assuma que o compensador $(\mathbb{A}_t;\ t \geq 0)$ é (q.c.)-contínuo e que

$$P\left(\lim_{t \to \infty} \mathbb{A}_t = \infty\right) = 1.$$

Então, o processo $\Pi = (\Pi_t, \mathcal{E}_t; t \geq 0)$ é um processo de Poisson.

PROVA. É claro da definição de Π_t que o processo Π tem trajetórias do tipo escada, contínuas à direita, onde os saltos são inteiros, pois $\Delta\Pi_t = N_{\sigma(t)} - N_{\sigma(t-)}$.

Assim, para provar o teorema é suficiente mostrar que a magnitude dos saltos de Π é um e que (q.c.)

$$E\left(e^{i\lambda(\Pi_t - \Pi_s)} / \mathcal{E}_s\right) = \exp\left[(e^{i\lambda} - 1)(t - s)\right]. \tag{2.26}$$

Por (2.22),

$$\begin{aligned} e^{i\lambda\Pi_t} - e^{i\lambda\Pi_s} = (e^{i\lambda} - 1) \int_{\sigma(s)}^{\sigma(t)} e^{i\lambda N_u} d[N_u - \mathbb{A}_u] \ + \\ + (e^{i\lambda} - 1) \int_{\sigma(s)}^{\sigma(t)} e^{i\lambda N_u} d\mathbb{A}_u \ . \end{aligned} \tag{2.27}$$

O processo

$$\left(\int_0^t e^{i\lambda N_u} d[N_u - \mathbb{A}_u] , \mathcal{F}_t ; t \geq 0 \right)$$

é um martingale local, logo, pela igualdade $E(\mathbb{A}_{\sigma(t)}) = t$, tem-se

$$E\left(\int_{\sigma(s)}^{\sigma(t)} e^{i\lambda N_u} d[N_u - \mathbb{A}_u] / \mathcal{F}_{\sigma(s)} \right) = 0 . \qquad (2.28)$$

Ainda, usando-se a continuidade da função $\mathbb{A}_t$ e a definição dos tempos $\sigma(t)$, temos:

$$\int_{\sigma(s)}^{\sigma(t)} e^{i\lambda N_u} d\mathbb{A}_u = \int_{\sigma(s)}^{\sigma(t)} e^{i\lambda N_u} d\mathbb{A}_u = \int_s^t e^{i\lambda N_{\sigma(u)}} du = \int_s^t e^{i\lambda \Pi_u} du . \qquad (2.29)$$

Fazendo $V_s(t) = E\left(e^{i\lambda \Pi_t} / \mathcal{E}_s\right)$ e usando (2.27)-(2.29) da mesma forma que na prova do Teorema 2.5, obtém-se o resultado (2.26).

Agora estabeleceremos a magnitude dos saltos. Para tanto, diferenciamos ambos os lados de (2.26) com respeito a λ e tomando $\lambda = 0$, encontramos que o processo $(\Pi_t - t, \mathcal{E}_t ; t \geq 0)$ é um martingale. Conseqüentemente, se $\theta_1 = inf\{t : \Pi_t > 0\}$, então

$$E(\Pi_{t \wedge \theta_1}) = E(t \wedge \theta_1) \leq E\,\theta_1 .$$

Dado que $P(\theta_1 > t) = P(\Pi_t = 0)$ e $E\left(e^{i\lambda \Pi_t}\right) = \exp[(e^{i\lambda} - 1)t]$, obtém-se facilmente que $P(\Pi_t = 0) = e^{-t}$ e $E(\theta_1) = 1$.

Aqui, $E(\Pi_{t \wedge \theta_1}) \leq 1$ e pelo Lema de Fatou, $E(\Pi_{\theta_1}) \leq 1$.

Mas, $\Pi_{\theta_1} \geq 1$ (q.c.), logo,

$$P(\Pi_{\theta_1} = 1) = 1,$$

ou seja, o valor do primeiro salto do processo Π é um (q.c.). O mesmo pode ser estabelecido para os saltos restantes.

■

Um caso particular do teorema anterior, que permitirá a simulação das trajetórias de um processo de Poisson não homogêneo no Capítulo 3, é dado a seguir.

COROLÁRIO. Com as notações do Teorema 2.6, seja $N = (N_t, \mathcal{F}_t; t \geq 0)$ um processo de Poisson não homogêneo com função intensidade $\lambda(t)$, então o processo $\Pi = (\Pi_{\sigma(t)}, \mathcal{E}_t; t \geq 0)$ é um processo de Poisson homogêneo com função intensidade $\lambda \equiv 1$.

2.F. ALGUNS EXEMPLOS DE APLICAÇÕES DOS PROCESSOS PONTUAIS.

Sendo um processo pontual um modelo para pontos indistingüíveis distribuídos aleatoriamente em $[0, \infty)$, onde os pontos representam tempos de ocorrência de eventos, os mesmos têm sido utilizados para modelar uma ampla variedade de situações em diversas áreas tais como engenharia de comunicações, análise de sobrevivência, física de partículas, pesquisa operacional, etc.

Encerramos este capítulo com alguns exemplos de

aplicações de processos pontuais no contexto de teoria de comunicações e de análise de sobrevivência.

EXEMPLO 5. (TRANSMISSÃO LASER).

Um processo pontual é apropriado para carregar informação acerca de algum outro processo e nosso interesse poderia ser recobrar parte ou a totalidade dessa informação contida numa seqüência de pontos. Como um exemplo, consideremos uma fonte laser modulada por uma corrente elétrica que proporciona a energia de ativação e contém a informação a ser transmitida [ver Figura 2].

Denotemos por S_t o sinal de modulação; a emissão de fótons é proporcional a alguma função não negativa $f(S_t)$ deste sinal satisfazendo $f(0) = 0$. No extremo receptor deste sistema de comunicação simplificado existe um dispositivo de conversão fotoeletrônica (fotodetector) cuja saída é um fluxo de elétrons que deveria ser aproximadamente proporcional ao fluxo médio de fótons incidentes, isso no caso de não existir ruído na conversão.

O fluxo de elétrons é observável; o processo de contagem associado $N = (N_t, \mathcal{F}_t; t \geq 0)$ tem estatísticas que podem ser descritas em termos de sua intensidade: N é um processo pontual com intensidade $\lambda_t = \mu + B_t f(S_t)$. Aqui μ é a intensidade correspondente ao ruído de conversão: se $S_t = 0$, alguns elétrons são emitidos pelo cátodo e formam um processo de Poisson com intensidade μ . B_t é um ruído multiplicativo, independente do sinal S_t e pode ser considerado como devido à absorção e turbulência na

atmosfera, entre a fonte laser e o receptor.

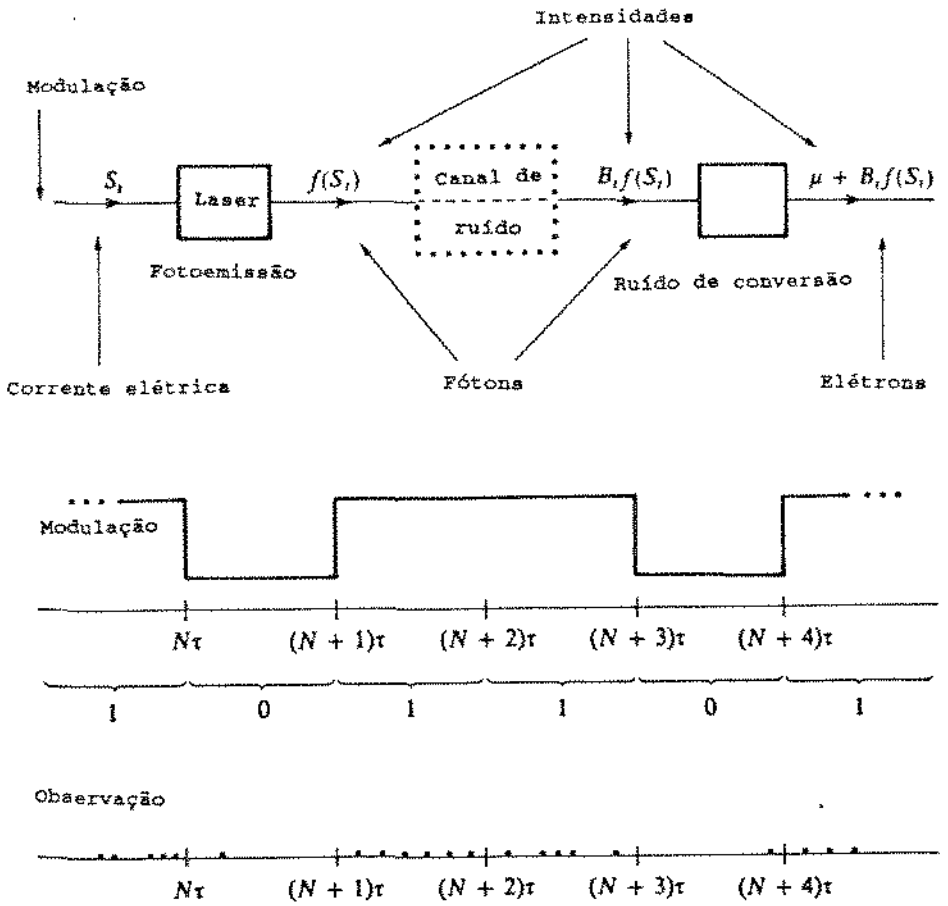


Figura 2. Esquema simplificado de um dispositivo de comunicação laser.

O modelo acima não é muito exato do ponto de vista físico; o problema do ruído na conversão atualmente é quase eliminado pela utilização de fibras ópticas como guias de onda.

EXEMPLO 6.

Sejam $T_1, T_2, \ldots, T_r$ variáveis aleatórias contínuas não negativas e assumamos que elas representam os tempos de vida de r diferentes indivíduos. Definamos para $t \geq 0$, N_t como

o número de indivíduos que morrem até a idade t , isto é,

$$N_t = \sum_{i=1}^{r} \mathbf{I}(T_i \leq t)$$

Supondo que dois indivíduos não morrem ao mesmo tempo, $N = (N_t, t \geq 0)$ é um processo pontual em relação a $\mathcal{F}_t^N$.

Notar que a observação de N equivale a observar as estatísticas de ordem $(T_{(1)}, \ldots, T_{(r)})$, mas não proporciona informação referente a qual indivíduo morreu primeiro, qual morreu segundo, etc.

Seja $f(t)$ a função densidade de T_i , o compensador $A = (A_t; t > 0)$ de N pode ser obtido usando o Teorema 2.2. Assim,

$$A_t = \sum_{i \geq 1} \int_0^{t \wedge T_i} \frac{f(u)}{1 - F(u)} du, \quad i \geq 1 . \tag{2.30}$$

EXEMPLO 7. (ANÁLISE DE SOBREVIVÊNCIA COM CENSURA).

Suponha que um grupo de n pacientes indexados por $i = 1, \ldots, n$, afetados por uma certa doença (letal) é controlado num hospital desde o instante de diagnóstico da doença até o instante de morte ou até alguma data fixa de término do estudo. Assim, para cada paciente i observamos uma duração da doença $\overline{T}_i$, a qual é seu verdadeiro tempo de sobrevivência T_i , (isto é, o tempo desde o diagnóstico até a morte) ou um tempo *censurado* (isto é, o tempo desde o diagnóstico até a data de término).

Seja

$$D_i = \begin{cases} 1, & \text{se } \overline{T}_i = T_i \\ 0, & \text{caso contrário.} \end{cases}$$

Assumimos ainda que se os pares (T_i, D_i), $i=1,\ldots,n$ são independentes, podemos definir o seguinte processo:

$$N_t^i = \mathbb{I}(T_i \leq t, D_i = 1), \quad i=1,\ldots,n \tag{2.31}$$

(na verdade este é um processo pontual *multivariado*, para cada i).

Pela definição, N^i é zero antes de $\overline{T}_i$ e salta a um em $\overline{T}_i$, se $\overline{T}_i$ é um verdadeiro tempo de sobrevivência.

Para encontrar o correspondente processo intensidade, raciocinamos da seguinte maneira: em qualquer tempo t, sabemos que para o *i-ésimo* paciente foi observada morte ou censura, ou bem ele está vivo e sem censura. Para os dois primeiros casos a probabilidade condicional de observar um salto de N^i em $(t, t+dt)$, dado todo o passado estrito até t é zero.

Para o último caso, esta probabilidade condicional será $\alpha_i(t)\,dt$, onde $\alpha_i(t)$ é a função risco do Exemplo 6 para o tempo de sobrevivência T_i do paciente i.

A definição de processo de intensidade formulada por Aalen (1978), e que foi dada na Seção 2.C para complementar a Definição 2.4, pode ser reescrita como

$$\lambda_t dt = P\,(dN_t = 1/\mathcal{F}_{t^-})\,. \qquad (2.32)$$

Definamos

$$Y_t^i = \mathbb{I}(\overline{T}_i \geq t)\,. \qquad (2.33)$$

Então, temos que

$$P\left(dN_t^i = 1/\mathcal{F}_{t^-}\right) = \alpha_i(t)\,Y_t^i \qquad (2.34)$$

($\mathcal{F}_{t^-}$ representa toda a informação disponível sobre a evolução da doença até antes de t). De (2.32) e (2.34) vemos que o processo N^i tem processo intensidade

$$\lambda_t^i = \alpha_i(t)\,Y_t^i,\ i=1,2,\ldots,n. \qquad (2.35)$$

Em muitos casos, uma hipótese razoável de assumir é a homogeneidade da população, ou seja, as intensidades α_i são idênticas para todos os indivíduos. Denotemos por α este valor comum.

Se definirmos um novo processo somando os processos "individuais", isto é,

$$N_t = \sum_{i=1}^{n} N_t^i \qquad (2.36)$$

obteremos um processo pontual univariado que conta o número total de mortes em $[0,t]$. De (2.32) e (2.35) e a suposição de não haver mortes simultâneas, segue que $N = (N_t, t\geq 0)$ tem processo intensidade λ dado por

$$\lambda_t = \sum_{i=1}^{n} \lambda_t^i = \alpha(t) Y_t \qquad (2.37)$$

onde

$$Y_t = \sum_{i=1}^{n} Y_t^i .$$

É interessante notar que a expressão de λ_t é simplesmente o produto de intensidades de morte (risco) para um só indivíduo e o número de indivíduos "em risco" até antes do tempo t .

Um processo cuja intensidade tem a forma (2.37) é chamado de *processo com intensidade multiplicativa* introduzido por Aalen (1975), que mostrou sua aplicabilidade a uma ampla gama de situações. Retomaremos o estudo deste tipo de processo na Seção 3.D.

EXEMPLO 8. (PROCESSO DE RENOVAÇÃO)

Um processo pontual $N_t = \sum_{n\geq 1} \mathbb{I}(\tau_n \leq t)$ é um *processo de renovação* com distribuição entre chegadas F se os tempos entre saltos $\tau_i - \tau_{i-1}, i=1,2,\ldots$ e $\tau_0 = 0$, são variáveis aleatórias i.i.d. com distribuição F.

Comparando este exemplo com o Exemplo 1, notamos que o processo de renovação é uma generalização do processo de Poisson, correspondendo este último ao caso particular em que F é a distribuição exponencial.

Uma aplicação direta da fórmula (2.8) conduz a

$$\Lambda_t = -\sum_{n\geq 1} \log[1-F(t\wedge\tau_n)], \quad t\geq 0$$

como o compensador do processo de renovação acima descrito.

A importância do processo de renovação reside no fato que muitos processos estocásticos contém imersos processos de renovação cujos pontos são tipicamente os tempos de ocorrência de eventos nos quais o processo estocástico (chamado de processo regenerativo) se renova probabilisticamente. Exemplos de tais processos encontram-se em teoria de confiabilidade, teoria de filas, etc.

Capítulo 3

Estimação Não Paramétrica para certos Processos Pontuais.

Em um modelo de processo pontual, o objeto de interesse é a taxa ou intensidade com que eventos ocorrem, pois o desenvolvimento no tempo do processo $N = (N_t; t \geq 0)$ está governado pelo seu processo intensidade $\lambda = (\lambda_t; t \geq 0)$, formalizado na Definição 2.4 e interpretado como sendo $\lambda_t dt$ a probabilidade condicional que um salto de N ocorra num intervalo $(t, t+dt)$, dado tudo que ocorreu até o instante anterior a t, isto é, $\lambda_t dt = P(dN_t = 1/\mathcal{F}_{t-})$.

Aqui faremos suposições apenas sobre a forma funcional da intensidade, impondo certas condições de regularidade. Estaremos portanto diante de um problema não paramétrico ou com espaço paramétrico de dimensão infinita, no qual precisamos construir estimadores da intensidade baseados num conjunto de observações do processo pontual em questão.

Grande parte das técnicas de estimação existentes para intensidades de processos pontuais estão baseadas nas técnicas de estimação de funções densidade de probabilidade. Uma ampla revisão de tal assunto é feita em Bean and Tsokos (1980).

A fim de facilitar o estudo posterior do problema de estimar intensidades de processos pontuais, começaremos este

capítulo fazendo uma breve discussão de alguns aspectos da estimação não paramétrica de densidades.

3.A. ESTIMAÇÃO NAO PARAMÉTRICA DE FUNÇÕES DE DENSIDADE.

A presente seção é de natureza puramente estatística e pretende estabelecer uma familiaridade com os conceitos envolvidos em estimação não paramétrica de densidades.

A formulação do problema de estimar uma densidade é seguinte: dados $X_1,\ldots,X_n$, variáveis aleatórias independentes, identicamente distribuídas com função de densidade desconhecida f, deve-se construir um estimador $\hat{f}$ para f.

Um dos primeiros estimadores de densidades na literatura foi aquele proposto por Rosenblatt (1956) na forma

$$\hat{f}(x) = \hat{f}_{n,h}(x) = \frac{F_n(x+h) - F_n(x-h)}{2h}, \qquad (3.1)$$

onde $h = h_n \sim n^{-\alpha}$ para $0<\alpha<1$.

Aqui $F_n(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{I}(X_i \leq x)$ é a função de distribuição empírica que foi introduzida no Exemplo 2.2 e representa a proporção de pontos amostrais em $(-\infty, x]$.

Na determinação do estimador $\hat{f}$ aparece o problema da escolha de h de forma que $h = h_n \longrightarrow 0$ quando $n \longrightarrow \infty$.

Usando como medida de discrepância o Erro Quadrático Integrado Médio (EQIM) definido por

$$EQIM(\hat{f}) = E\left(\int_{-\infty}^{\infty}[\hat{f}(x) - f(x)]^2 dx\right), \qquad (3.2)$$

Rosenblatt (1956) prova que a escolha ótima de h, no sentido de minimizar $EQIM(\hat{f})$ é $h = kn^{-1/5}$, para certo k que depende somente de $\int (f'')^2$.

É fácil verificar que o estimador de Rosenblatt pode ser escrito como

$$\hat{f}(x) = \hat{f}_{n,h}(x) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{h}W\left(\frac{x - X_i}{h}\right), \qquad (3.3)$$

onde $W(u) = \frac{1}{2}\mathbb{I}(|u|<1)$.

Graficamente isto significa que o estimador pode ser obtido somando "caixas" de amplitude $2h$ e altura $(2hn)^{-1}$ centradas sobre cada observação.

Uma das desvantagens do estimador anteriormente definido é que não é uma função contínua e tem derivada zero em todo ponto interior de (X_i, X_{i+1}), o que não é em geral verdade para a densidade f a ser estimada.

Parzen (1962) propôs uma generalização do estimador de Rosenblatt substituindo a função peso W em (3.3) por uma função K, limitada e com integral igual a um, usualmente assumida como uma densidade de probabilidade simétrica e denominada função *kernel*.

O "estimador de função kernel" proposto por Parzen, que daqui para diante chamaremos simplesmente de *estimador kernel*, tem a forma

$$\hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x-X_i}{h}\right) = \int_{-\infty}^{\infty} K_h(x-y)\, dF_n(y)\,, \qquad (3.4)$$

onde $K_h(u) = h^{-1}K(u/h)$, F_n é a função de distribuição empírica e h é um parâmetro positivo que chamaremos de *janela*, também referido na literatura como *parâmetro de alisamento* (Rudemo, 1982) ou *amplitude de banda* (Marron, 1985).

Analogamente ao estimador de Rosenblatt, o estimador kernel pode ser visualizado graficamente como uma soma de curvas centradas nas observações, como mostra a Figura 3.

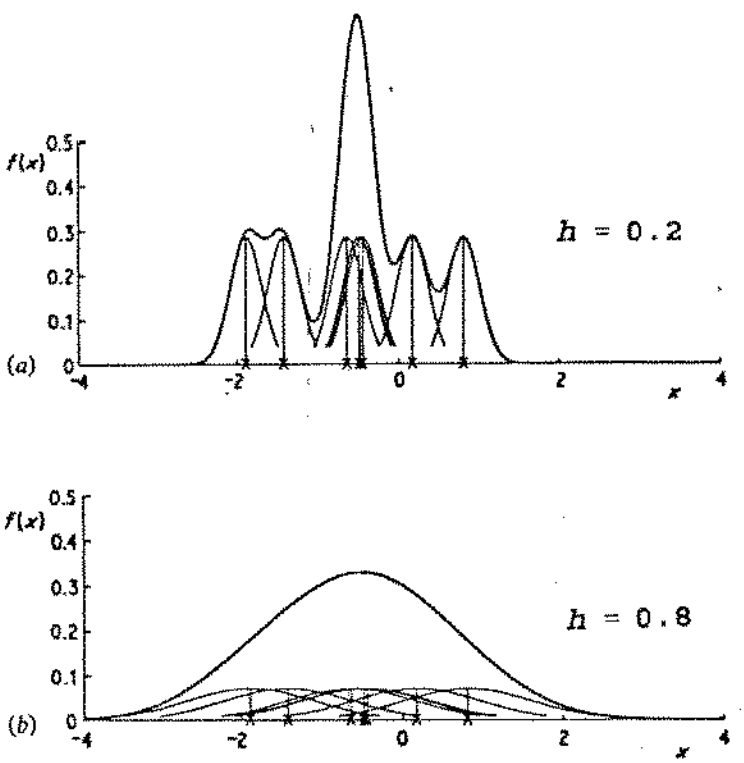


Figura 3. Interpretação gráfica do estimador de kernel.

A função kernel determina as formas das curvas e a janela, sua amplitude.

Segue da definição do estimador kernel que $\hat{f}$ herdará todas as propriedades (continuidade, diferenciabilidade, etc.) de função kernel. Também é claro da definição que variações na janela resultarão em variações na forma e conseqüentemente na "qualidade" do estimador.

Segundo os resultados obtidos por Anderson (1969) e Epanechnikov (1969), a escolha da função kernel é de importância secundária, enquanto o problema da escolha da janela é de crucial importância na estimação.

Em certas aplicações, é possível escolher a janela de uma forma de certo modo subjetiva, como em Silverman (1978) e Crestana (1991), mas outras aplicações requerem uma escolha automática da janela, isto é, uma escolha totalmente baseada nos dados e sem nenhuma especificação a priori.

3.B. A DECOMPOSIÇÃO VARIÂNCIA-VÍCIO QUADRÁTICO.

Para ter uma idéia quantitativa da qualidade ("performance") de um estimador fazem-se necessárias medidas de discrepância ou proximidade do estimador $\hat{f}$ da verdadeira densidade f. Uma medida bastante natural para esse propósito é o Erro Quadrático Médio (EQM) definido localmente em x por

$$EQM_x(\hat{f}) = E[\hat{f}(x) - f(x)]^2. \qquad (3.5)$$

Claramente, o EQM pode ser decomposto como

$$EQM_x(\hat{f}) = [E\hat{f}(x) - f(x)]^2 + var\,\hat{f}(x), \qquad (3.6)$$

isto é, a soma do vício quadrático e a variância. Existe uma forte conexão entre estas duas quantidades e uma não pode ser reduzida sem aumentar o valor da outra. A magnitude de ambas é controlada pelo valor da janela.

Uma outra medida de proximidade usual é o Erro Quadrático Integrado Médio (**EQIM**), definido em (3.2). Notamos que

$$EQIM(\hat{f}) = \int EQM_x(\hat{f})\,dx. \qquad (3.7)$$

Analogamente ao $EQM_x(\hat{f})$ existe também para $EQIM(\hat{f})$ uma decomposição em soma do vício quadrático integrado e a variância integrada.

Usando a decomposição variância-vício quadrático acima mencionada, Rosenblatt (1971) apresentou uma descrição bastante intuitiva das implicações em escolher valores muito pequenos ou muito grandes para a janela: um valor pequeno de h resulta em uma alta variância, enquanto um valor grande de h resulta em alto vício. As expressões de aproximação que clarificam esse comportamento são mostradas a continuação. Por simplicidade assumiremos que

$$\int K(t)\,dt=1,\quad \int tK(t)\,dt=0 \quad e \quad \int t^2K(t)\,dt=c_2\neq 0,$$

onde convencionamos que uma integral sem limites significa integração sobre a reta toda.

Convém lembrar que o kernel está sob controle do usuário para propósitos práticos. Indicaremos a dependência de h do estimador escrevendo $\hat{f}_h$ em lugar de $\hat{f}$.

Temos,

$$vício\left[\hat{f}_h(x)\right] = E\hat{f}_h(x) - f(x)$$

Fazendo a mudança $y = x-ht$ e usando a hipótese $\int K=1$, temos que

$$vício\left[\hat{f}_h(x)\right] = \int K(t)\,[f(x-ht) - f(x)]\,dt.$$

Uma expansão em série de Taylor-McLaurin produz

$$f(x-ht) = f(x) - htf'(x) + \frac{1}{2}h^2t^2f''(x) + o(h^3)$$

e pelas suposições feitas sobre K,

$$\begin{aligned} vício\left[\hat{f}_h(x)\right] &= \\ &= -hf'(x)\int tK(t)\,dt + \frac{1}{2}h^2f''(x)\int t^2K(t)\,dt + o(h^3) \\ &= \frac{1}{2}h^2f''(x)\,c_2 + o(h^3). \end{aligned} \tag{3.8}$$

Logo, o vício quadrático integrado vai ser dado aproximadamente por

$$\int \text{vício}^2[\hat{f}_h(x)]dx \approx \frac{1}{4}h^4c_2^2\int [f'(x)]^2dx. \qquad (3.9)$$

Para a variância:

$$\text{var}[\hat{f}_h(x)] = n^{-1}\int h^{-2}K^2\left(\frac{x-y}{h}\right)f(y)\,dy - n^{-1}\{f(x) + \text{vício}[\hat{f}_h(x)]\}^2 ;$$

usando $y = x-ht$ e a expressão aproximada (3.8) do vício, obtemos

$$\text{var}\,\hat{f}_h(x) \approx (nh^{-1})\int f(x-ht)K^2(t)\,dt - n^{-1}\{f(x) + o(h^2)\}^2.$$

Assumindo que h seja pequeno e n grande e expandindo $f(x-ht)$ em série de Taylor-McLaurin, obtemos

$$\begin{aligned}\text{var}\,\hat{f}_h(x) &\approx (nh)^{-1}\int [f(x) - htf'(x) + \ldots]K^2(t)\,dt + o(n^{-1}) \\ &= (nh)^{-1}f(x)\int K^2(t)\,dt + o(n^{-1}) \\ &\approx (nh)^{-1}f(x)\int K^2(t)\,dt.\end{aligned} \qquad (3.10)$$

Assim, a variância integrada pode ser aproximada como:

$$\int \text{var}\,\hat{f}_h(x)\,dx \approx (nh)^{-1}\int K^2(t)\,dt \qquad (3.11)$$

e portanto,

$$EQIM_x(\hat{f}_h) \approx \frac{1}{4}h^4c_2^2\int [f''(x)]^2dx + \frac{1}{nh}\int K^2(t)\,dt. \qquad (3.12)$$

Vemos de (3.12) que se tentarmos diminuir o vício tomando um valor h muito pequeno, estaremos incrementando a variância integrada. Analogamente, se tomarmos um valor h muito grande, estaremos diminuindo a variância, mas aumentando o vício.

O resultado de variar a janela é claramente ilustrado na Figura 3. Para a janela $h = 0.2$ aparece uma alta variância em forma de ruído no estimador. Isto acontece porque quando $h \longrightarrow 0$ o estimador chega a ser uma soma de funções delta de Dirac.

Para uma janela maior como $h = 0.8$, aparece um valor alto de vício que se traduz em um sobrealisamento, perdendo-se detalhes importantes na estimação, como por exemplo uma possível bimodalidade.

Depreende-se da análise anterior que é desejável encontrar uma janela que equilibre os efeitos do vício e da variância do estimador.

Vários métodos têm sido propostos para escolher uma janela ótima em algum sentido. O método conhecido como "plug-in" foi introduzido por Woodroofe (1970); um outro método, relativo à informação de Kullback-Leibler foi dado por Duin (1976) e um método prático de escolha, baseado na segunda derivada do estimador, foi proposto por Silverman (1978). Na seção seguinte estudaremos o método de "validação cruzada", sugerido por Rudemo (1982) e Bowman (1984). Ele é totalmente baseado nos dados, possui propriedades assintóticas atrativas e será aquele que utilizaremos na estimação de intensidades.

3.C. VALIDAÇÃO CRUZADA DE MÍNIMOS QUADRADOS.

Embora o método de validação cruzada de mínimos quadrados tenha sido formulado em anos recentes, a idéia na qual está baseado é muito simples: dado qualquer estimador kernel $\hat{f}_h$ de uma densidade f, o erro quadrático integrado (EQI) pode ser escrito como

$$EQI(h) = \int (\hat{f}_h - f)^2 = \int \hat{f}_h^2 - 2\int \hat{f}_h f + \int f^2, \qquad (3.13)$$

como o último termo não depende de h; a escolha ótima da janela, no sentido de minimizar o EQI, corresponde àquela que minimiza a quantidade R definida por

$$R(\hat{f}_h) = \int \hat{f}_h^2 - 2\int \hat{f}_h f. \qquad (3.14)$$

Esta minimização não pode ser efetuada na prática, pois o termo $\int \hat{f}_h f$ é desconhecido.

O princípio básico da validação cruzada de mínimos quadrados é construir um estimador de $R(\hat{f}_h)$ a partir dos dados e então minimizar este estimador em relação a h para obter uma escolha da janela.

Definamos $\hat{f}_{hi}$ como o estimador kernel construído com todas as observações, exceto X_i, isto é,

$$\hat{f}_{hi}(x) = \frac{1}{(n-1)h} \sum_{j \neq i} K\left(\frac{x - X_j}{h}\right) \qquad (3.15)$$

O estimador $\hat{f}_{hi}$ é chamado na literatura estimador "leave-one-out".

Definamos agora a *função de validação cruzada* ou *score* como sendo

$$VC(h) = \int \hat{f}_h^2 - \frac{2}{n} \sum_i \hat{f}_{hi}(X_i) \qquad (3.16)$$

$VC(h)$ depende somente dos dados e é um estimador razoável dos termos em $EQI(h)$ que dependem de h, como será mostrado a seguir. Logo, a janela que minimiza $VC(h)$ deveria ser "próxima" da janela que minimiza o EQI de $\hat{f}_h(x)$.

A justificativa da afirmação anterior baseia-se nos seguintes resultados:

$$\int f \hat{f}_h = \int \frac{1}{n} \sum_{i=1}^{n} K_h(x - X_i) f(x)\, dx$$

$$= \frac{1}{n} \sum_{i=1}^{n} \int K_h(x - X_i) f(x)\, dx,$$

onde $K_h(u) = \frac{1}{h} K(u/h)$.

Então,

$$E\left(\int f \hat{f}_h\right) = E\left(\int K_h(x - X_i) f(x)\, dx\right)$$

$$= \int\int K_h(x - y) f(x) f(y)\, dx dy$$

$$= E[K_h(X - Y)],$$

onde X e Y são variáveis aleatórias independentes com densidade f . Conseqüentemente,

$$E\left(\int f\hat{f}_h\right) = E\left[\frac{1}{n(n-1)}\sum_i\sum_{j\neq i} K_h(X_i - X_j)\right]$$

$$= E\left[\frac{1}{n}\sum_i \hat{f}_{hi}(X_i)\right], \tag{3.17}$$

pela própria definição de $\hat{f}_{hi}$.

Assim,

$$E\,R(\hat{f}_h) = E\int \hat{f}_h^2 - 2E\int f\hat{f}_h$$

$$= E\int \hat{f}_h^2 - \frac{2}{n}E\sum_i \hat{f}_{hi}(X_i)$$

$$= E[VC(h)]. \tag{3.18}$$

Tomando esperança em (3.13) e usando (3.7), (3.17) e (3.18) temos que

$$EQIM(\hat{f}_h) = E[R(\hat{f})] + \int f^2$$

$$= E[VC(h)] + \int f^2. \tag{3.19}$$

Esta expressão mostra que minimizar $E[VC(h)]$ corresponde exatamente a minimizar o erro quadrático integrado médio (EQIM) de $\hat{f}_h$. Assumindo que o mínimo de $VC(h)$ é próximo do mínimo de $E[VC(h)]$, justifica-se a afirmação que minimizar $VC(h)$ fornece uma boa escolha de

janela ótima. Esta última afirmação foi mostrada por Hall (1983) na forma a seguir: sejam $\hat{h}_0$ a janela que minimiza $EQI(h)$ e $\hat{h}_{VC}$ a janela que minimiza $VC(h)$. Sob a hipótese que a densidade $f(x)$ tem segunda derivada contínua, $\hat{h}_{VC}$ é assintoticamente equivalente a $\hat{h}_0$ no sentido que

$$\frac{\hat{h}_{VC}}{\hat{h}_0} \xrightarrow{P} 1$$

e

$$\frac{EQI(\hat{h}_{VC})}{EQI(\hat{h}_0)} \xrightarrow{P} 1.$$

Este resultado de otimalidade assintótica da janela de validação cruzada será adaptado ao caso de intensidade multiplicativa, na seção 3.G.

Para expressar o score VC de uma forma mais adequada a seu cálculo computacional, definamos $K^{(2)}$ como a convolução do kernel consigo mesmo, isto é, $K^{(2)} = K*K$ e suponhamos K simétrico. Fazendo $u = \frac{x}{h}$, obtemos

$$\int \hat{f}_h^2(x)\,dx = \int \left[\frac{1}{nh}\sum_i K\left(\frac{x-X_i}{h}\right)\right] dx$$

$$= \frac{1}{n^2 h}\sum_i\sum_j \int K\left(\frac{X_i}{h}-u\right)K\left(u-\frac{X_j}{h}\right)du \quad (K \text{ limitado})$$

$$= \frac{1}{n^2 h}\sum_i\sum_j K^{(2)}\left(\frac{X_i-X_j}{h}\right).$$

(3.20)

Por outra parte,

$$\frac{1}{n}\sum_i \hat{f}_{hi}(X_i) = \frac{1}{n}\sum_i \frac{1}{(n-1)}\sum_j \frac{1}{h}K\left(\frac{X_i - X_j}{h}\right)$$

$$= \frac{1}{n(n-1)}\sum_i\sum_j \frac{1}{h}K\left(\frac{X_i - X_j}{h}\right) - \frac{1}{(n-1)h}K(0)\,. \tag{3.21}$$

Substituindo (3.20) e (3.21) em (3.16) resulta

$$VC(h) = \frac{1}{n^2 h}\sum_{i=1}^{n}\sum_{j=1}^{n} K^{(2)}\left(\frac{X_i - X_j}{h}\right) - \frac{2}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1}^{n} K\left(\frac{X_i - X_j}{h}\right) + \frac{2}{(n-1)h}K(0)\,. \tag{3.22}$$

Esta última expressão será utilizada no programa computacional que calcula a janela ótima no Capítulo 4.

3.D. ESTIMAÇÃO NO MODELO DE AALEN.

Na Seção 2.F, foi introduzido o modelo de Aalen de intensidade multiplicativa.

Para um processo pontual $N = (N_t; t \geq 0)$ com decomposição D-M,

$$M_t = N_t - \int_0^t \lambda_s ds \quad , \tag{3.23}$$

o processo intensidade é da forma

$$\lambda_t = \alpha(t) Y_t, \quad t \in [0,1] \tag{3.24}$$

com α determinística e Y_t correspondente a um processo

predizível e observável.

Por simplicidade matemática e sem perda de generalidade, nesta seção os processos serão restritos ao intervalo de tempo $[0,1]$.

Aalen assume o martingale em (3.23) como de quadrado integrável com processo variância [ver Apêndice A]

$$\langle M \rangle_t = \int_0^t \lambda_s ds \tag{3.25}$$

Usando (3.23) e (3.24) podemos escrever simbolicamente

$$dN_t = \alpha(t) Y_t dt + dM_t \tag{3.26}$$

Consideremos o problema de estimar a função intensidade acumulada

$$\beta(t) = \int_0^t \alpha(s)\, ds. \tag{3.27}$$

Interpretando dM_t como um "ruído", (3.26) sugere que um estimador para $\beta(t)$ seria $\int_0^t \frac{dN_s}{Y_s}$, pois

$$dN_t \approx \alpha(t) Y_t dt \Rightarrow \frac{1}{Y_t} dN_t \approx \alpha(t)\, dt.$$

Em muitos casos, é possível acontecer que $Y_s = 0$; sendo assim, é conveniente modificar $\beta(t)$ definindo o processo $\beta^* = (\beta_t^*;\ t \in [0,1])$ como

$$\beta_t^* = \int_0^t \alpha(s)\, J(s)\, ds \qquad (3.28)$$

onde $J(s) = \mathbb{I}(Y_s > 0)$ e com a convenção que $\frac{J(s)}{Y_s} = 0$ se $Y_s = 0$.

Assume-se, como uma condição técnica, que

$$E\left[\int_0^1 \frac{J(s)\,\alpha(s)}{Y_s}\, ds\right] < \infty$$

com o propósito de garantir a existência de certas integrais estocásticas relacionadas ao estimador a ser estudado.

A modificação feita sobre $\beta(t)$ sugere como estimador de β_t^* a integral (estocástica) de Stieltjes

$$\beta_t = \int_0^t \frac{J(s)}{Y_s}\, dN_s, \qquad (3.29)$$

que pode ser escrita alternativamente como

$$\beta_t = \sum_{i:\, \tau_i \le t} \frac{J(\tau_i)}{Y_{\tau_i}}, \qquad (3.30)$$

onde $\tau_1, \tau_2, \ldots$ são os tempos de salto do processo N .

Portanto β_t é uma função tipo escada, crescente, contínua à direita e com incrementos $(Y_{\tau_i})^{-1}$ nos tempos observados de salto de N . O estimador β_t é conhecido como *estimador de Nelson-Aalen*, pois é uma generalização daquele proposto por Nelson (1969, 1972) para estimar a função risco acumulado.

TEOREMA 3.1. O processo $\hat{\beta}-\beta^*$ é um martingale de quadrado integrável com processo variância

$$\langle \hat{\beta}-\beta^* \rangle_t = \int_0^t \left[\frac{\alpha(s) J(s)}{Y_s} \right] ds. \qquad (3.31)$$

PROVA.

$$\begin{aligned}
\hat{\beta}_t - \beta_t^* &= \int_0^t \frac{J(s)}{Y_s} dN_s - \int_0^t \alpha(s) J(s) ds \\
&= \int_0^t \frac{J(s)}{Y_s} [dN_s - \alpha(s) Y_s ds] \\
&= \int_0^t \frac{J(s)}{Y_s} d\left[N_s - \int_0^s \alpha(u) Y_u du \right] \\
&= \int_0^t \frac{J(s)}{Y_s} dM_s
\end{aligned} \qquad (3.32)$$

(onde a última passagem segue de (3.23)).

Aqui, $J(s)/Y_s$ é predizível (por ser contínuo à esquerda e adaptado à $\mathcal{F}$) e M é um martingale de quadrado integrável. Logo, pelo Teorema A.3, $\hat{\beta}-\beta^*$ é também um martingale de quadrado integrável com variância dada por (3.31).

■

O erro quadrático médio do estimador β é

$$\eta(t) = E\left[\beta_t - \beta_t^*\right]^2 = E\left[\int_0^t \frac{J(s)}{Y_s} dM_s\right]^2 .$$

De acordo com o Teorema 3.1, $\beta - \beta^* \in \mathbf{M}^2$ e portanto $(\beta - \beta^*)^2$ é um submartingale que admite a seguinte decomposição D-M

$$(\beta - \beta^*) = M' + \langle \beta - \beta^* \rangle , \tag{3.33}$$

onde M' é um martingale de média zero.

Segue que

$$\eta(t) = E[\langle \beta - \beta^* \rangle_t] = E\left(\int_0^t \frac{\alpha(s) J(s)}{Y_s} ds\right). \tag{3.34}$$

Definamos

$$\hat{\eta}(t) = \int_0^t \frac{\alpha(s) J(s)}{Y_s^2} dN_s . \tag{3.35}$$

De (3.23) temos que

$$\hat{\eta}(t) = \int_0^t \frac{\alpha(s) J(s)}{Y_s} ds + \int_0^t \frac{J(s)}{Y_s^2} dM_s .$$

Aqui a segunda integral é um martingale de média zero, assim

$$E\hat{\eta}(t) = E\left(\int_0^t \frac{\alpha(s) J(s)}{Y_s} ds\right) = \eta(t) .$$

Portanto $\hat{\eta}(t)$ é um estimador não tendencioso para $\eta(t)$.

3.E. O ESTIMADOR KERNEL NO MODELO DE AALEN.

Na seção anterior foi desenvolvido um estimador β para a intensidade acumulada, $\beta(t) = \int_0^t \alpha(s)\,ds$, com alguma restrição sobre o processo $Y = (Y_t; t \in [0,1])$, mas para muitos propósitos, a entidade de interesse é a intensidade $\alpha(t)$ e não $\beta(t)$ ou $\beta^*(t)$. Precisamos portanto de um estimador para $\alpha(t)$. Ramlau-Hansen (1983), baseou-se no estimador kernel de Parzen-Rosenblatt, dado em (3.4), para propor como estimador de $\alpha(t)$ no modelo de intensidade multiplicativa o seguinte

$$\hat{\alpha}(t) = \frac{1}{h}\int_0^1 K\left(\frac{t-s}{h}\right)d\hat{\beta}_s, \qquad (3.36)$$

onde h é um parâmetro positivo, isto é, a janela, K uma função kernel e $\hat{\beta}$ é o estimador de Nelson-Aalen definido em (3.29).

Por simplicidade matemática, nesta seção assumiremos que a função kernel tem suporte em $[-1,1]$.

O estimador de Ramlau-Hansen é basicamente um alisamento do estimador de Nelson-Aalen para intensidade acumulada, assim como o estimador de Parzen-Rosenblatt era um alisamento da função distribuição empírica.

Analogamente à (3.30) o estimador de Ramlau-Hansen pode ser escrito como

$$\hat{\alpha}(t) = \frac{1}{h} \sum_{i:\tau_i \le t} K\left(\frac{t-\tau_i}{h}\right) \frac{1}{Y_{\tau_i}}. \qquad (3.37)$$

Dado que K é zero fora de $[-1,1]$, somente aqueles valores de i para os quais $t-h \le \tau_i \le t+h$, contribuem na soma; logo, para uma janela h será considerada a estimação de $\hat{\alpha}(t)$ somente em $[h, 1-h]$, dado que para esses valores de t, (3.37) é uma média real dos incrementos $\frac{1}{Y_{\tau_i}}$ de β.

Com o fim de estudar as propriedades de $\hat{\alpha}(t)$, Ramlau-Hansen (1983) define, para $0<h<1/2$ a quantidade

$$\alpha^*(t) = \frac{1}{h} \int_0^1 K\left(\frac{t-s}{h}\right) d\beta_s^*$$

$$= \int_{-1}^{1} K(u)\, \alpha(t-hu)\, J(t-hu)\, du\,, \qquad (3.38)$$

para $t \in [h, 1-h]$.

Notemos que

$$\hat{\alpha}(t) - \alpha^*(t) = \frac{1}{h} \int_0^1 K\left(\frac{t-s}{h}\right) \frac{J(s)}{Y_s} dM_s$$

é uma integral estocástica em relação ao martingale de quadrado integrável M; este fato proporciona uma base para estudar as propriedades estatísticas do estimador de Ramlau-Hansen.

A média de $\hat{\alpha}$, sua variância e um estimador não tendencioso da variância estão contidos na seguinte

proposição.

PROPOSIÇÃO 3.1. (Ramlau-Hansen, 1983). Sejam $f(s) = E[J(s)]$ e $K_h(s) = (1/h)K(s/h)$. Então

$$E\hat{\alpha}(t) = E\alpha^*(t) = (K_h * (\alpha f))(t) \ , \tag{3.39}$$

onde $*$ denota convolução usual. Além disso,

$$\begin{aligned} \sigma^2(t) &= E[\hat{\alpha}(t) - \alpha^*(t)]^2 \\ &= \frac{1}{h}\int_{-1}^{1} K^2(u)\,\alpha(t-hu)\,E\left[\frac{J(t-hu)}{Y(t-hu)}\right]du \ , \end{aligned} \tag{3.40}$$

para $t \in [h, 1-h]$. Um estimador não tendencioso para $\sigma^2(t)$ é dado por

$$\hat{\sigma}^2(t) = \frac{1}{h^2}\int_0^1 K^2\left(\frac{t-s}{h}\right)\left[\frac{J(s)}{Y^2(s)}\right]dN_s \tag{3.41}$$

para $t \in [h, 1-h]$.

PROVA. Da definição de $\hat{\alpha}$ e α^* segue que

$$\hat{\alpha}(t) - \alpha^*(t) = \frac{1}{h}\int_0^1 K\left(\frac{t-s}{h}\right)d(\hat{\beta} - \beta^*)_s \ .$$

Pelo Teorema 3.1, $(\hat{\beta} - \beta^*)$ é um martingale de média zero e como $K(\cdot)$ é predizível (determinístico), tem-se

$$E[\hat{\alpha}(t) - \alpha^*(t)] = 0.$$

Um cálculo direto leva a

$$E\hat{\alpha}(t) = E\alpha^*(t) = \int_0^1 K_h(t-s)\alpha(s)E(J(s))ds$$

$$= (K_h * (\alpha f))(t).$$

Pela fórmula (3.31),

$$E[\hat{\alpha}(t) - \alpha^*(t)]^2 = \frac{1}{h^2} E\left(\int_0^1 K^2\left(\frac{t-s}{h}\right)\left[\frac{\alpha(s)J(s)}{Y(s)}\right]ds\right)$$

$$= \frac{1}{h}\int_{-1}^{1} K^2(u)\alpha(t-hu)E\left[\frac{J(t-hu)}{Y(t-hu)}\right]du .$$

O resultado final da proposição segue do fato que

$$\hat{\sigma}^2(t) = \frac{1}{h}\int_{-1}^{1} K^2(u)\left[\frac{J(t-hu)}{Y(t-hu)}\right]\alpha(t-hu)\,du + \frac{1}{h^2}\int_0^1 K^2\left(\frac{t-s}{h}\right)J(s)\,dM_s , \tag{3.42}$$

onde M é o martingale de (3.23).

■

Se denotarmos o primeiro termo em (3.42) por $\sigma^{*2}(t)$, teremos que $E\hat{\sigma}^2(t) = E\sigma^{*2}(t)$ (pois o segundo termo é um submartingale de média zero) e que

$$E[\hat{\sigma}^2(t) - \sigma^{*2}(t)]^2 = \frac{1}{h^3}\int_{-1}^{1} K^4(u)\alpha(t-hu)E\left[\frac{J(t-hu)}{Y(t-hu)}\right]du. \tag{3.43}$$

A expressão (3.39) mostra que o estimador kernel é um estimador tendencioso de $\alpha(t)$. Porém, é possível mostrar

que ele é assintoticamente não tendencioso e derivar uma expressão assintótica para sua variância da seguinte forma: se considerarmos uma seqüência (N_n) de processos pontuais, cada um deles com processo intensidade $\lambda_n(t) = \alpha(t) Y_n(t)$, podemos construir uma correspondente seqüência de estimadores kernel, que denotaremos por $\hat{\alpha}_n(t)$, onde assumimos o kernel fixo e uma seqüência de janelas h_n.

Obtém-se assim o seguinte resultado.

PROPOSIÇÃO 3.2. (Ramlau-Hansen, 1983).

a) Se a intensidade α é contínua em t e se $\tilde{J}_n = E J_n \underset{n \to \infty}{\longrightarrow} 1$ uniformemente numa vizinhança de t, então

$$E\hat{\alpha}_n(t) \longrightarrow \alpha(t) . \tag{3.44}$$

b) se são satisfeitas as condições

i) $nE\left[\frac{J_n(s)}{Y_n(s)}\right] \longrightarrow \frac{1}{\tau(s)}$ uniformemente numa vizinhança de t.

ii) α e τ são contínuas no ponto t.

Então

$$\begin{aligned} \sigma_n^2(t) &= E\left[\hat{\alpha}_n(t) - \alpha_n^*(t)\right]^2 \\ &= \frac{1}{nh_n} \frac{\alpha(t)}{\tau(t)} \int_{-1}^{1} K^2(u)\, du + o\left(\frac{1}{nh_n}\right), \end{aligned} \tag{3.45}$$

quando a seqüência de janelas $h_n \longrightarrow 0$.

PROVA. A relação (3.44) segue de (3.39) e do fato que a seqüência de funções (K_{h_n}) , onde $K_{h_n}(u) = \frac{1}{h_n}K(u/h_n)$, é uma seqüência de Dirac quando $h_n \longrightarrow 0$. A relação (3.45) segue diretamente de (3.40) e das suposições feitas.

■

Pode-se mostrar que sob condições não muito fortes, o estimador kernel é um estimador consistente para a intensidade em consideração.

PROPOSIÇÃO 3.3. (Ramlau-Hansen, 1983). Considere a seqüência (N_n) de processos pontuais unidimensionais como na proposição 3.2. Assuma satisfeitas as condições da parte b) na Proposição 3.2. Então

$$E\left[\hat{\alpha}_n(t) - \alpha(t)\right]^2 \longrightarrow 0 \,, \tag{3.46}$$

quando $n \longrightarrow \infty$, $h_n \longrightarrow 0$ e $nh_n \longrightarrow \infty$.

PROVA. Pela Proposição 3.2. b) segue que $\sigma_n^2(t) \longrightarrow 0$, assim é suficiente mostrar que $E\left[\alpha_n^*(t) - \alpha(t)\right]^2 \longrightarrow 0$. Dado que $J_n \xrightarrow{P} 1$ uniformemente numa vizinhança de t e dado que

$$\alpha_n^*(t) - \alpha(t) = \int_{-1}^{1} K(u)\left[\alpha(t - h_n u)\, J_n(t - h_n u - \alpha(t)\right] du,$$

segue que

$$\left[\alpha_n^*(t) - \alpha(t)\right] \xrightarrow{P} 0$$

e que existe uma constante $c > 0$ tal que

$$|\alpha_n^*(t) - \alpha(t)| \leq c.$$

Concluímos que

$$E\left[\alpha_n^*(t) - \alpha(t)\right]^2 \longrightarrow 0$$

quando $n \longrightarrow \infty$.

■

Um resultado importante em estimação por kernel é a normalidade assintótica do estimador, provada por Ramlau-Hansen (1983) utilizando um resultado prévio de Liptser e Shiryayev (1980) sobre a distribuição assintótica de um arranjo triangular de martingales.

Consideremos uma seqüência (N_n) de processos pontuais como na Proposição 3.2, com a correspondente seqüência de martingales dada por

$$M_n(t) = N_n(t) - \int_0^t \lambda_n(s)\, ds \qquad (3.47)$$

onde (λ_n) é a seqüência de processos intensidade. Seja (H_n) uma seqüência de processos predizíveis tais que

$$E\left[\int_0^1 H_n^2(s)\,\lambda_n(s)\right] < \infty$$

e seja

$$\tilde{M}_n(t) = \int_0^t H_n(s)\, dM_n(s) .$$

Pelo Teorema A.3, $(\tilde{M}_n)$ é uma seqüência de martingales e portanto a seguinte proposição é válida.

PROPOSIÇÃO 3.4. Suponha que quando $n \longrightarrow \infty$

i) $\forall \epsilon > 0 : \int_0^1 H_n(s)^2 \mathbf{I}(|H_n(s)| > \epsilon)\, \lambda_n(s)\, ds \xrightarrow{P} 0$

ii) $\int_0^1 H_n(s)^2 \lambda_n(s)\, ds \xrightarrow{P} 1$

Então $\tilde{M}_n(1) \xrightarrow{D} N(0,1)$, onde $N(0,1)$ é a distribuição normal padrão.

PROVA. As condições i) e ii) acima são equivalentes as condições (L_2) e $(C\text{-}12)$ em Liptser e Shiryayev (1980), Corolário 2. O resultado segue pela Nota 1 daquele artigo.

■

Consideremos agora uma seqüência de processos pontuais onde os correspondentes estimadores de kernel são escritos da seguinte forma.

$$\hat{\alpha}_n(t) = \alpha_n^*(t) + \frac{1}{h_n} \int_0^1 K\left(\frac{t-s}{h_n}\right) \frac{J_n(s)}{Y_n(s)}\, dM_n(s),$$

onde, como anteriormente definido,

$$\alpha_n^*(t) = \frac{1}{h_n} \int_0^1 K\left(\frac{t-s}{h_n}\right) \alpha(s) J(s)\, ds.$$

Fixado o valor de t, temos

$$(nh_n)^{1/2}\left[\hat{\alpha}_n(t) - \alpha_n^*(t)\right] = \int_0^1 H_n(s)\, dM_n(s)\,, \qquad (3.48)$$

com

$$H_n(s) = \left(\frac{n}{h_n}\right)^{1/2} K\left(\frac{t-s}{h_n}\right)\frac{J_n(s)}{Y_n(s)}\,.$$

Formulamos a seguir o anunciado resultado de normalidade assintótica para a seqüência de estimadores de kernel.

TEOREMA 3.2. (Ramlau-Hansen, 1983). Assuma que

a) $\dfrac{nJ_n}{Y_n} \xrightarrow{P} \dfrac{1}{\tau}$ uniformemente numa vizinhança de t,

quando $n \longrightarrow \infty$.

b) as funções α e τ são contínuas no ponto t.

Então

$$(nh_n)^{1/2}\left[\hat{\alpha}_n(t) - \alpha_n^*(t)\right]$$

converge em distribuição para uma variável aleatória com distribuição normal com média zero e variância igual a

$$\frac{\alpha(t)}{\tau(t)}\int_{-1}^{1} K^2(u)\, du$$

quando $n \longrightarrow \infty$ e $nh_n \longrightarrow \infty$.

PROVA. Verificaremos serem satisfeitas as condições i) e ii) da Proposição 3.4, das quais segue a normalidade assintótica.

Em primeiro lugar, temos que

$$\{|H_n(s)|>\epsilon\} = \left\{\left|K\left(\frac{t-s}{h_n}\right)\left(\frac{n}{h_n}\right)^{1/2}\frac{J_n(s)}{Y_n(s)}\right|>\epsilon\right\}$$

$$= \left\{\left|K\left(\frac{t-s}{h_n}\right)\frac{nJ_n(s)}{Y_n(s)}\right|>\epsilon(nh_n)^{1/2}\right\}.$$

Dado que $h_n \longrightarrow 0$, $nh_n \longrightarrow \infty$; de a), $\frac{nJ_n}{Y_n} \xrightarrow{P} \frac{1}{\tau}$ uniformemente numa vizinhança de t e $1/\tau$ é limitada nessa vizinhança, concluímos que $\mathbb{I}(|H_n(s)|>\epsilon) \xrightarrow{P} 0$ uniformemente em $[0,1]$. Então, pela definição de H_n , tem-se

$$\int_0^1 H_n^2(s)\,\mathbb{I}(|H_n(s)|>\epsilon)\,Y_n(s)\,\alpha(s)\,ds \xrightarrow{P} 0.$$

Assim, a condição i) é válida. Por outro lado, de a) e b) temos que

$$\int_0^1 H_n^2(s)\,\lambda_n(s)\,ds = \frac{1}{h_n}\int_0^1 K^2\left(\frac{t-s}{h_n}\right)\left\{\frac{n\alpha(s)J_n(s)}{Y_n(s)}\right\}ds$$

$$= \int_{-1}^1 K^2(u)\left\{\frac{nJ_n(t-h_nu)}{Y_n(th_nu)}\right\}\alpha(t-h_nu)\,du \xrightarrow{P} \frac{\alpha(t)}{\tau(t)}\int_{-1}^1 K^2(u)\,du,$$

a qual é a condição ii) e o teorema fica provado.

■

3.F. VALIDAÇÃO CRUZADA PARA ESTIMAR A INTENSIDADE DE UM PROCESSO DE POISSON NÃO HOMOGÊNEO

Na presente seção estudaremos uma adaptação do método de validação cruzada desenvolvido na Seção 3.C para encontrar o valor ótimo de janela do estimador kernel de intensidade de um importante processo pontual: o processo de Poisson não homogêneo.

O presente enfoque teórico é devido a Marron e Brooks (1991). Sejam $\tau_1, \tau_2, \ldots, \tau_{N_T}$ todas as observações no intervalo $[0, T]$ de um processo de Poisson não homogêneo com função intensidade $\lambda(t)$. A fim de impor uma estrutura matemática sobre a função intensidade, assumiremos um caso particular do modelo de Aalen dado na Seção 2.F: o modelo simples de intensidade multiplicativa.

Aqui,

$$\lambda(t) = c\,\alpha(t) \tag{3.49}$$

onde c é uma constante positiva e $\alpha(t)$ uma densidade de probabilidade em $[0, T]$. Denotaremos a intensidade por $\lambda_c(t)$ quando interessar sua dependência com os valores assumidos por c.

Sob esse modelo N_T, o número de observações em $[0, T]$, tem uma distribuição de Poisson com

$$E(N_T) = \int_0^T \lambda(u)\,du = c.$$

Condicionado a N_T, os tempos de ocorrência $\tau_1, \tau_2, \ldots, \tau_N$ têm a mesma distribuição que as estatísticas de ordem correspondentes a N_T variáveis aleatórias independentes com densidade de probabilidade $a(t)$, $t \in [0, T]$. Também temos que no modelo assumido, fazer $c \longrightarrow \infty$ tem o efeito de aumentar as observações em $[0, T]$ sem mudar a forma relativa de $\lambda_c(t)$ no processo limite. Então, dado que precisaremos considerar propriedades de convergência, suporemos que os valores de c provêm de uma seqüência de reais positivos $(c_s : s \geq 1)$ tal que $c_s/s \longrightarrow a$ para alguma constante a, quando $s \longrightarrow \infty$, o que implica que $c_s \longrightarrow \infty$, quando $s \longrightarrow \infty$.

A função intensidade é similar a uma função densidade no sentido que esperamos ter mais (resp. menos) observações num intervalo de acordo com a área sob a curva da função intensidade no intervalo ser maior (resp. menor). Logo, um estimador natural para $\lambda(t)$ é o estimador kernel

$$\begin{aligned} \hat{\lambda}_h(t) &= \sum_{i=1}^{N_T} K_h(t-\tau_i), \quad t \in [0, T] \\ &= \int_0^T K_h(t-s)\, dN_s \end{aligned} \tag{3.50}$$

onde $K_h(u) = (1/h)\, K(u/h)$. De novo, a função kernel $K(\cdot)$ será geralmente uma função densidade de probabilidade simétrica e a janela, h, um parâmetro positivo. Se para a seqüência de funções intensidade $\lambda_{c_s}(t) = c_s a(t)$ indexada por s, construímos uma correspondente seqüência de estimadores kernel $\hat{\lambda}_h^s(t)$ e assumindo que quando $c_s \longrightarrow \infty$, $h \longrightarrow 0$ com

$hc_n \longrightarrow \infty$, obtemos da seção anterior que $\hat{\lambda}_h^n$ é consistente e assintoticamente normal.

Existem duas diferenças entre os estimadores kernel para densidades, dado em (3.4) e para intensidades, em (3.50). Primeiro, $\hat{\lambda}_h$ não inclui um fator de normalização n^{-1} , dado que $\int_r^t \lambda(u)\,du$ é o número esperado em lugar da proporção esperada de observações entre r e t . Segundo, N_T é uma variável aleatória no contexto de estimação de intensidades.

Nosso atual objetivo é estudar a adaptação do método de escolha da janela por validação cruzada, estudado na Seção 3.C, para ser aplicado ao estimador $\hat{\lambda}_h$.

No presente contexto temos que o erro quadrático integrado (EQI) de $\hat{\lambda}_h$ é dado por

$$EQI_\lambda(h) = \int_0^T \hat{\lambda}_h^2 - 2\int_0^T \hat{\lambda}_h \lambda + \int_0^T \lambda^2 . \tag{3.51}$$

Similarmente a (3.15) e (3.16), definimos a função "score" como

$$VC_\lambda(h) = \int_0^T \hat{\lambda}_h^2 - 2\sum_{i=1}^{N_T} \hat{\lambda}_{hi}(\tau_i) \tag{3.52}$$

onde

$$\hat{\lambda}_{hi}(t) = \sum_{j \neq i} K_h(t-\tau_j) . \tag{3.53}$$

Usando uma justificativa idêntica àquela dada para estimação de densidades [ver (3.17)-(3.19)]: só substituir as esperanças por esperanças condicionadas a N_T , é possível

afirmar que $VC_\lambda(h)$ é um estimador razoável dos termos em $EQI_\lambda(h)$ que dependem de h e conseqüentemente, a janela que minimiza $VC_\lambda(h)$ deve, em princípio, ser próxima da janela que minimiza $EQI_\lambda(h)$.

É possível ainda, e será o propósito da seguinte seção, mostrar que a janela escolhida por validação cruzada é assintoticamente ótima.

3.G. OTIMALIDADE ASSINTÓTICA DA JANELA OBTIDA POR VALIDAÇÃO CRUZADA.

O erro quadrático integrado médio (EQIM) de $\hat{\lambda}_h$ pode ser decomposto em um termo variância e um outro de vício quadrático da mesma forma que em (3.12), obtém-se assim

$$\begin{aligned} EQIM_\lambda(h) &= h^{-1}c_g\left(\int K^2(t)\,dt\right) + \\ &\quad + h^4 c_g^2\left[\frac{1}{2}\int t^2 K(t)\,dt\right]^2 \int_0^T [\alpha''(x)]^2 dx + \\ &\quad + o(h^{-1}c_g + h^4 c_g^2) \end{aligned} \tag{3.54}$$

quando $h \longrightarrow 0$, $c_g \longrightarrow \infty$ e $hc_g \longrightarrow \infty$.

Assim, o erro quadrático integrado médio assintótico (EQIMA) é

$$EQIMA_\lambda(h) = h^{-1}c_g\left(\int K^2\right) + h^4 c_g^2\left[\frac{1}{2}\int t^2 K\right]^2 \int_0^T [\alpha''(x)]^2 . \tag{3.55}$$

Um resultado de otimalidade assintótica da janela obtida pelo método de validação cruzada de mínimos quadrados será apresentado nesta seção, mas previamente são formulados

dois lemas que serão utilizados na prova do teorema.

LEMA 3.1. Para o estimador kernel tem-se que

$$\sup_{h \in H_s} \left| \frac{EQI_\lambda(h)}{EQIMA_\lambda(h)} \right| \longrightarrow 1 \quad (q.c.), \text{ quando } s \longrightarrow \infty.$$

O Lema 3.1 afirma que o erro quadrático integrado e o erro quadrado integrado médio assintótico de $\hat{\lambda}_h(t)$ são assintoticamente equivalentes. A prova é remetida a Brooks (1991) e é muito próxima da prova do Lema 3.2 dado à continuação.

LEMA 3.2. Para o estimador kernel tem-se que

$$\sup_{h,b \in H_s} \left| \frac{VC_\lambda(h) - EQI_\lambda(h) - [VC_\lambda(b) - EQI_\lambda(b)]}{EQIMA_\lambda(h) + EQIMA_\lambda(b)} \right| \longrightarrow 0 \quad (q.c.),$$

quando $s \longrightarrow \infty$.

PROVA. Seja $g(1, 2, \ldots, N_T) \longrightarrow (1, 2, \ldots, N_T)$ uma permutação aleatória dos números $(1, 2, \ldots, N_T)$. Definamos $Y_i = \tau_{g(i)}$, ou seja, os Y_i 's são τ_i 's não ordenados.

Dado que os τ_i 's são observações de um processo de Poisson não homogêneo com intensidade $\lambda(t)$, as Y_i 's são variáveis aleatórias i.i.d. com densidade $\alpha(t)\mathbf{I}_{[0,T]}(t)$.

Definamos

$$\hat{\alpha}_h(t) = c_s^{-1}\hat{\lambda}_h(t) = c_s^{-1}\sum_{i=1}^{N_T} K_h(t-\tau_i) ,$$

$$\hat{\alpha}_{hi}(t) = c_n^{-1}\hat{\lambda}_{hi}(t) = c_n^{-1}\sum_{\substack{j=1 \\ j\neq i}}^{N_T} K_h(t-\tau_j) \ ,$$

e

$$VC_\alpha(h) = c_n^{-2}VC_\lambda(h) = \int_0^T \hat{\alpha}_h(t)^2 dt - 2c_n^{-1}\sum_{i=1}^{N_T}\hat{\alpha}_{hi}(\tau_i) \ .$$

O EQI para α pode ser expresso como

$$EQI_\alpha(h) = \int_0^T [\hat{\alpha}_h(u) - \alpha(u)]^2 du = c_n^{-2} EQI_\lambda(h)$$

e

$$EQIMA_\alpha(h) = c_n^{-2} EQIMA_\lambda(h)$$

Na prova do Lema 3.2, precisaremos do seguinte resultado

$$\sup_{h,b\in H_n}\left|\frac{VC_\lambda(h) - EQI_\lambda(h) - [VC_\lambda(b) - EQI_\lambda(b)]}{EQIMA_\lambda(h) + EQIMA_\lambda(b)}\right|$$

$$\leq 4 \sup_{h\in H_n}\left|\frac{c_n^{-1}\sum_{i=1}^{N_T}\hat{\alpha}_{hi}(\tau_i) - \int_0^T \hat{\alpha}\alpha - G}{EQIMA_\alpha(h)}\right| \qquad (3.56)$$

onde $G = \left[\frac{2(N_T-1)}{c_n} - 1\right]c_n^{-1}\sum_{i=1}^{N_T}\alpha(\tau_i) - \left[\frac{N_T(N_T-1)}{c_n^2}\right]\int_0^T \alpha^2$. O resultado é demonstrado a seguir.

Tem-se diretamente das definições que

$$\left| \frac{VC_{\lambda}(h) - EQI_{\lambda}(h) - [VC_{\lambda}(b) - EQI_{\lambda}(b)]}{EQIMA_{\lambda}(h) + EQIMA_{\alpha}(b)} \right|$$

$$= \left| \frac{VC_{\alpha}(h) - EQI_{\alpha}(h) - [VC_{\alpha}(b) - EQI_{\alpha}(b)]}{EQIMA_{\alpha}(h) + EQIMA_{\alpha}(b)} \right|$$

$$= \left| \frac{2\int_0^T \hat{\alpha}_h \alpha - 2c_s^{-1} \sum_{i=1}^{N_T} \hat{\alpha}_{hi}(X_i) - 2\int \hat{\alpha}_b \alpha - 2c_s^{-1} \sum_{i=1}^{N_T} \hat{\alpha}_{bi}(X_i)}{EQIMA_{\alpha}(h) + EQIMA_{\alpha}(b)} \right|$$

$$= 2 \left| \frac{F(h) - F(b)}{EQIMA_{\alpha}(h) + EQIMA_{\alpha}(b)} \right| = 2 \left| \frac{F(h) - G - (F(b) - G)}{EQIMA_{\alpha}(h) + EQIMA_{\alpha}(b)} \right|$$

(onde a penúltima passagem segue fazendo $F(h) = \int_0^T \hat{\alpha}_h \alpha - c_s^{-1} \sum_{i=1}^{N_T} \hat{\alpha}_{hi}(X_i)$).

$$\leq 2 \left\{ \left| \frac{F(h) - G}{EQIMA_{\alpha}(h) + EQIMA_{\alpha}(b)} \right| + \left| \frac{F(b) - G}{EQIMA_{\alpha}(h) + EQIMA_{\alpha}(b)} \right| \right\}$$

$$\leq 2 \left\{ \left| \frac{F(h) - G}{EQIMA_{\alpha}(h)} \right| + \left| \frac{F(b) - G}{EQIMA_{\alpha}(b)} \right| \right\}$$

$$\leq 2 \left\{ \sup_{h \in H_s} \left| \frac{F(h) - G}{EQIMA_{\alpha}}(h) \right| + \sup_{b \in H_s} \left| \frac{F(b) - G}{EQIMA_{\alpha}(b)} \right| \right\}$$

$$= 4 \sup_{h \in H_s} \left| \frac{F(h) - G}{EQIMA_{\alpha}(h)} \right|$$

Então, para provar o Lema 3.2 é suficiente provar que

$$\sup_{h \in H_n} \left| \frac{c_n^{-1} \sum_{i=1}^{N_T} \hat{\alpha}_{hi}(\tau_i) - \int_0^T \hat{\alpha}_h \alpha - G}{EQIMA_\alpha(h)} \right| \longrightarrow 0 \quad q.c., \tag{3.57}$$

quando $n \longrightarrow \infty$.

Adaptando os resultados de Härdle, Marron e Wand (1990) para estimar por kernel as derivadas de uma densidade de probabilidade, definamos

$$U_{ij}(h) = K_h(Y_i - Y_j) - \int k_h(y - Y_j)\,\alpha(y)\,dy - \alpha(Y_i) + \int_0^T \alpha^2(y)\,dy,$$

$$V_i(h) = E\,(U_{ij}/Y_i) = E\,[K_h(Y_i - Y_j)/Y_i]$$

$$- E\left[\int k_h(y - Y_j)\,\alpha(y)\,dy/Y_i\right] - E\,[\alpha(Y_i)/Y_i] + \int_0^T \alpha^2(y)\,dy$$

$$= \int K_h(Y_i - y)\,\alpha(y)\,dy - \iint K_h(y - z)\,\alpha(y)\,\alpha(z)\,dy\,dz - \alpha(Y_i) + \int_0^T \alpha^2 dy,$$

$$W_{ij}(h) = U_{ij}(h) - V_i(h)$$

$$= K_h(Y_i - Y_j) - \int K_h(y - Y_j)\,\alpha(y)\,dy - \int K_h(Y_i - y)\,\alpha(y)\,dy +$$

$$+ \iint K_h(y - z)\,\alpha(y)\,\alpha(z)\,dy\,dz$$

e

$$R_i(h) = \left[\frac{N_T - 1}{c_s} - 1\right] [2\int K_h(Y_i - y)\,\alpha(y)\,dy - \iint K_h(y-z)\,\alpha(y)\,\alpha(z)\,dydz$$

$$- 2\alpha(Y_i) + \int_0^T \alpha^2(y)\,dy] .$$

Obtém-se que $E(U_{ij}) = 0$ e logo $E(V_i/N_T) = E(W_{ij}/Y_i, N_T) = 0$,

para $i, j = 1, 2, \ldots, N_T$.

Para somas sobre $i = 1, 2, \ldots, N_T$, as Y_i 's (observações não ordenadas) podem ser substituídas pelos τ_i 's (observações ordenadas). Logo, obtemos

$$c_s^{-1}\sum_{i=1}^{N_T} V_i + c_s^{-2}\sum_{i=1}^{N_T}\sum_{j\neq i} W_{ij} + c_s^{-1}\sum_{i=1}^{N_T} R_i = c_s^{-1}\sum_{i=1}^{N_T} \hat{\alpha}_{hi}(\tau_i) - \int_0^T \hat{\alpha}_h \alpha - G. \quad (3.58)$$

Aqui, (3.58) é satisfeita se são satisfeitas as três condições seguintes

$$\sup_{h\in E_s} \left| \frac{c_s^{-1}\sum_{i=1}^{N_T} V_i(h)}{EQIMA_s(h)} \right| \longrightarrow 0 \quad q.c. \text{ quando } s \longrightarrow \infty. \quad (3.59)$$

$$\sup_{h\in E_s} \left| \frac{c_s^{-2}\sum_{i=1}^{N_T}\sum_{j\neq i} W_{ij}(h)}{EQIMA_s(h)} \right| \longrightarrow 0 \quad q.c. \text{ quando } s \longrightarrow \infty. \quad (3.60)$$

$$\sup_{h\in E_s} \left| \frac{c_s^{-1}\sum_{i=1}^{N_T} R_i(h)}{EQIMA_s(h)} \right| \longrightarrow 0 \quad q.c. \text{ quando } s \longrightarrow \infty. \quad (3.61)$$

Portanto, a fim de provar o Lema 3.2 é suficiente provar (3.59)-(3.61).

Condicionado a N_T, $\left(\sum_{i=1}^{k} V_i\right)$ e $\left(\sum_{i=1}^{k}\sum_{j=1}^{i-1} W_{ij}\right)$, para $k = 1, 2, \ldots, N_T$, são martingales em relação às σ-álgebras geradas por $\{Y_1, Y_2, \ldots, Y_k\}$.

A desigualdade de Burkholder (1973, pag.40) implica que para alguma constante A

$$E\left[\left(\sup_{k=1,\ldots,N}\sum_{i=1}^{N_T} V_i\right)\right] \leq A\, E\left[\left(\sum_{i=1}^{N_T} E\,(v_i^2/N_T)\right)/N_T\right] + A\sum_{i=1}^{\infty} E\left[|V_i|^{2m}/N_T\right] \tag{3.62}$$

e

$$E\left[\left(\sup_{k=1,\ldots,N_T}\sum_{i=1}^{N_T}\sum_{j=1}^{i-1} W_{ij}\right)^2/N_T\right]$$

$$\leq A\, E\left[\left(\sum_{j=1}^{i-1} E\left[\left(\sum_{j=1}^{i-1} W_{ij}\right)^2/Y_1, Y_2, \ldots, Y_{i-1}, N_T\right]\right)\right] \tag{3.63}$$

$$+ A\sum_{i=1}^{\infty} E\left[\left|\sum_{j=1}^{i-1} W_{ij}\right|^2/N_T\right]$$

Dado que N_T tem distribuição Poisson com média c_s, para cada m existe uma constante A_m tal que $E\,(N_T^m) \leq A_m(EN_T)^m = A_m c_s^m$.

Aqui,

$$E\left[c_s^{-1}\sum_{i=1}^{N_T} V_i\right]^{2m} \leq c_s^{-2m} E\left[E\left[\left(\sup_{k=1,\ldots,N_T}\sum_{i=1}^{k} Y_i\right)^{2m}\right]/N_T\right]$$

$$\leq c_s^{-2m} E\left[A_m N_T^m h^{4m} + A_m N_T\right] \quad [\text{por (3.62)}]$$

$$\leq A_m(c_s^{-m} h^{4m} + c_s^{-2m+1}),$$

De (3.63),

$$E\left[c_s^{-2}\sum_{i=1}^{N_T}\sum_{j=1}^{i-1}W_{ij}\right]^{2m} \le c_s^{-4m}E\,[A_m N_T^{2m}h^{-m}+A_m N_T^{m+2}h^{-2m}]$$

$$\le A_m(c_s^{-2m}h^{-m}+c_s^{-3m+2}h^{-2m}).$$

Como resultado, temos que para $h\in H_s$, m suficientemente grande e algum $\gamma>0$

$$E\left[\frac{c_s^{-1}\sum_{i=1}^{N_T}V_i}{EQIMA_s(h)}\right]^{2m} \le \frac{A_m(c_s^{-m}h^{4m}+c_s^{-2m+1})}{(c_s^{-1}h^{-1}+h^4)^{2m}} \le A_m[h^m+c_s h^{2m}]$$

$$\le A_m c_s^{-\gamma m}. \qquad (3.64)$$

Usando a Desigualdade de Chebychev

$$\sup_{h\in H_s} P\left[\left|c_s^{-1}\sum_{i=1}^{N_T}V_i\right| > c_s^{-\gamma/4}\,EQIMA_s(h)\right] \le A_s c_s^{-\gamma m}. \qquad (3.65)$$

Lembremos que $\#(H_s)\le c_s^p$ [suposição c)] e escolha-se m tal que $m>2(p+2)/\gamma$. Pela suposição d), pode-se mostrar que para a seqüência $(c_s; s\ge 1)$,

$$\sum_{s=1}^{\infty} P\left[\sup_{h\in H_s}\left|\frac{c_s^{-1}\sum_{i=1}^{N_T}V_i}{EQIMA_s(h)}\right| > \epsilon\right]$$

$$\le \sum_{s=1}^{\infty}\#(H_s)\sup_{h\in H_s} P\left[\left|c_s^{-1}\sum_{i=1}^{N_T}V_i\right| > \epsilon\,EQIMA_s(h)\right] \qquad (3.66)$$

$$< \infty.$$

Assim, o Lema de Borel-Cantelli implica que

$$\sup_{h \in H_s} \left| \frac{c_s^{-1} \sum_{i=1}^{N_T} V_i}{EQIMA_\alpha(h)} \right| \longrightarrow 0 \quad q.c. \text{ quando } s \longrightarrow \infty. \tag{3.67}$$

Isto prova a afirmação (3.59). Com o mesmo procedimento pode-se provar (3.60).

Consideremos a expressão (3.61). Dado que N_T tem distribuição Poisson (c_s), pelo Teorema Central do Limite tem-se que $\frac{N_T - c_s}{\sqrt{c_s}}$ é assintoticamente normal quando $s \longrightarrow \infty$.

Usando esse fato e uma expansão de Taylor para R_i, obtém-se que para qualquer $0 < \epsilon < 1/2$

$$c_s^{1/2} h^{-2+\epsilon} \left[c_s^{-1} \sum_{i=1}^{N_T} R_i \right] \xrightarrow{P} 0. \tag{3.68}$$

Aqui, para cada m, existe uma constante c_0 tal que se $c_s > c_0$

$$E\left[\frac{c_s^{-1} \sum_{i=1}^{N_T} R_i}{EQIMA_\alpha(h)} \right]^{2m} \leq \frac{A_m c_s^{-m} h^{4m-2\epsilon m}}{(c_s^{-1} h^{-1} + h^4)^{2m}} \leq A_m h^{(1-2\epsilon)m} \leq A_m c_s^{-\gamma m} \tag{3.69}$$

para $h \in H_s$ e algum $\gamma > 0$. De (3.69) segue o resultado (3.61).

Fica assim provado o Lema 3.2.

■

Enunciaremos agora o teorema de otimalidade assintótica

da janela de validação cruzada, provado por Stone (1984) para estimação de densidades e cuja versão adaptada a estimação de intensidades pertence a Marron e Brooks (1991).

Seja $\hat{h}_0$ qualquer janela que minimiza $EQI_\lambda(h)$ e $\hat{h}_{VC}$ uma janela que minimiza $VC_\lambda(h)$; esses mínimos sempre existem, pois $EQI_\lambda(h)$ e $VC_\lambda(h)$ são funções contínuas e limitadas.

Faremos as seguintes suposições:

a) A função kernel é uma função densidade simétrica, limitada e com suporte compacto (sem perda de generalidade, assumimos o suporte como sendo $[-1,1]$).

b) A função intensidade λ tem duas derivadas contínuas e limitadas.

c) Para cada s , as janelas a considerar provêm de um conjunto H_s tal que

$$\#H_s = \{\textit{número de elementos em } H_s\} \leq c_s^\rho$$

e para cada $h \in H_s$, $\beta^{-1} c_s^{(-1+\gamma)} \leq h \leq \beta c_s^{-\gamma}$, para certas constantes β, γ, ρ positivas.

d) Para alguma constante $a>0$, $\frac{C_s}{s} \longrightarrow a$ quando $s \longrightarrow \infty$.

A suposição b) é uma suposição técnica que permite aplicar expansões de Taylor para o estudo das funções de erro (EQI ou EQIM) de $\hat{\lambda}_h(t)$. A suposição c) indica que H_s é um conjunto cuja cardinalidade cresce algebricamente rápido e com ela o conjunto de possíveis janelas cobre o espectro de janelas consistentes.

TEOREMA 3.3 Suponhamos satisfeitas as condições a)-d) acima, então sob o modelo simples de intensidade multiplicativa

$$\frac{EQI_\lambda(\hat{h}_{VC})}{EQI_\lambda(\hat{h}_0)} \longrightarrow 1 \quad (q.c.), \quad \textit{quando } s \longrightarrow \infty.$$

Daremos aqui uma prova ligeiramente modificada da dada por Marron e Brooks (1991) e que consideramos mais intuitiva.

PROVA. Os Lemas 3.1 e 3.2 juntos implicam que

$$\sup_{h,b \in H_s} \left| \frac{VC_\lambda(h) - EQI_\lambda(h)\ [VC_\lambda(b) - EQI_\lambda(b)]}{EQI_\lambda(h) + EQI_\lambda(b)} \right| \longrightarrow 0 \quad (3.70)$$

(q.c.), quando $s \longrightarrow \infty$.

Agora,

$$\left| \frac{EQI_\lambda(\hat{h}_{VC})}{EQI_\lambda(\hat{h}_0)} - 1 \right| = \left| \frac{EQI_\lambda(\hat{h}_{VC}) - EQI_\lambda(\hat{h}_0)}{EQI_\lambda(\hat{h}_0)} \right|$$

$$\leq \left| \frac{EQI_\lambda(\hat{h}_{VC}) - EQI_\lambda(\hat{h}_0)}{EQI_\lambda(\hat{h}_0)} + \frac{VC_\lambda(\hat{h}_0) - VC(\hat{h}_{VC})}{EQI_\lambda(\hat{h}_0)} \right| =$$

$$= \left| \frac{VC_\lambda(\hat{h}_0) - EQI_\lambda(\hat{h}_0) - [VC_\lambda(\hat{h}_{VC}) - EQI_\lambda(\hat{h}_{VC})]}{EQI_\lambda(\hat{h}_0) + EQI_\lambda(\hat{h}_{VC})} \right| \times$$

$$\times \left| \frac{EQI_\lambda(\hat{h}_0) + EQI_\lambda(\hat{h}_{VC})}{EQI_\lambda(\hat{h}_0)} \right|$$

$$\leq \sup_{h,b \in H_n} \left| \frac{VC_\lambda(h) - EQI_\lambda(h) - [VC_\lambda(b) - EQI_\lambda(b)]}{EQI_\lambda(h) + EQI_\lambda(b)} \right| \times$$

$$\times \left| \frac{EQI_\lambda(h) + EQI_\lambda(b)}{EQI_\lambda(h)} \right| \qquad (3.71)$$

Segue diretamente das definições que

$$\left| \frac{EQI_\lambda(h) + EQI_\lambda(b)}{EQI_\lambda(h)} \right| = \left| \frac{EQI_\lambda(b)}{EQI_\lambda(h)} + 1 \right|$$

$$= \left| \frac{s^{-2}\int_0^T \hat{\lambda}_b^2 - 2s^{-1}(c_n/s)\int_0^T \hat{\lambda}_b \alpha + (c_n/s)^2 \int_0^T \alpha^2}{s^{-2}\int_0^T \hat{\lambda}_h^2 - 2s^{-1}(c_n/s)\int_0^T \hat{\lambda}_h \alpha + (c_n/s)^2 \int_0^T \alpha^2} + 1 \right|$$

Usando a hipótese $c_n/s \longrightarrow a$ quando $s \longrightarrow \infty$, o fato de $\hat{\lambda}_b$ e $\hat{\lambda}_h$ serem limitadas e as suposições sobre α, tem-se que

$$\lim_{s \longrightarrow \infty} \left| \frac{EQI_\lambda(h) + EQI_\lambda(b)}{EQI_\lambda(h)} \right| < \infty .$$

Isto junto com (3.70) implicam de (3.71) o resultado do Teorema 3.3. ■

Um resultado importante de otimalidade foi obtido por Hall (1983) na forma

$$\frac{h_{vc}}{h_0} \xrightarrow{P} 1$$

para estimação de funções densidade de probabilidade.

Nós constatamos não existir na literatura disponível uma prova do resultado

$$\frac{h_{vc}}{h_0} \longrightarrow 1 \quad (q.c.)$$

para estimação de funções intensidade de processos estocásticos.

O Teorema 3.3 proporciona um forte argumento assintótico em favor do método de validação cruzada de mínimos quadrados para a escolha da janela no estimador kernel de intensidade de um processo de Poisson não homogêneo, isto é, o método produz a escolha ótima da janela no sentido de mínimo Erro Quadrático Integrado.

No capítulo seguinte estudaremos os aspectos práticos de tal otimalidade mediante simulação de trajetórias de um processo de Poisson e o cálculo do estimador kernel obtido com a janela de validação cruzada.

Capítulo 4

Aspectos Práticos na Estimação de Intensidade de Processos de Poisson Não Homogêneo.

4.A. SIMULAÇÃO DO PROCESSO DE POISSON NÃO HOMOGÊNEO.

Foi mostrado na Seção 3.G que a janela obtida pelo método de validação cruzada de mínimos quadrados para o estimador kernel de intensidade de um processo de Poisson não homogêneo, é assintoticamente ótima. Uma análise assintótica permite visualizar a estrutura básica de um estimador, mas, dado que os conjuntos reais de dados sempre possuem um número finito de observações, é fundamental que um estimador se comporte razoavelmente bem para um conjunto de dados relativamente grande. Mostraremos que isso acontece no nosso caso.

Para tanto, simulamos trajetórias de um processo de Poisson não homogêneo com determinadas intensidades e sobre elas calculamos o estimador kernel usando a janela obtida por validação cruzada.

As funções intensidade utilizada foram:

1) $\lambda_1(t) = 2t+1, \quad t\in[0,1]$

2) $\lambda_2(t) = 12e^{-2t}, \quad t\in[0,1]$

3) $\lambda_3(t) = 3+sen(2\pi t), \quad t\in[0,1]$

No caso das intensidades $\lambda_1(t)$ e $\lambda_2(t)$, que são

funções inversíveis, a simulação das trajetórias foi feita utilizando o Corolário ao Teorema 2.6, que pode ser reformulado como:

"Os pontos $t_1, t_2, \ldots$ são pontos de salto de um processo de Poisson não homogêneo com $E[N(t)] = \Lambda(t)$ se e somente se $\Lambda(t_1), \Lambda(t_2), \ldots$ são pontos de salto de um processo de Poisson homogêneo com função intensidade $\lambda = 1$ ".

O algoritmo utilizado na geração de n trajetórias de um processo de Poisson com intensidades $\lambda = 1$ usa o resultado de que os tempos entre chegadas do dito processo têm distribuição exponencial com parâmetro 1. O algoritmo é como segue:

ALGORITMO 1.

1) Faça $i = 1$.
2) Faça $s = 0$.
3) Gere um número aleatório u em $[0,1]$.
4) Faça $y = -\log(1-u)$.
5) Faça $s = s+y$.
6) Se $s > \Lambda(1)$, faça $i = i+1$;
Se $i \leq n$, vá para 2;
Se $i > n$, vá para 7;
Se $s < \Lambda(1)$, salvar i e $\Lambda^{-1}(s) = inf\{t; \Lambda(t) > s\}$.
7) Fim.

Para a simulação das trajetórias com intensidade $\lambda_3(t)$, que é uma função não inversível, foi utilizado o método de "thinning" proposto por Lewis e Shedler (1979), através do seguinte teorema.

TEOREMA 4.1. Considere um processo de Poisson não homogêneo $N^* = (N_t^*; t \geq 0)$ com função intensidade $\lambda^*(t)$, tal que o número de pontos $N_{t_0}^*$ em um intervalo fixo $(0, t_0)$ possui distribuição de Poisson com parâmetro $\mu_0^* = \Lambda^*(t_0) - \Lambda^*(0)$. Sejam $\tau_1^*, \tau_2^*, \ldots, \tau_{N_{t_0}}^*$ os pontos de salto do processo N^* em $(0, t_0]$. Suponha que para $t \in (0, t_0]$, $\lambda(t) \leq \lambda^*(t)$. Para $i = 1, 2, \ldots, n$ delete o ponto de salto τ_i^* com probabilidade $1 - \lambda(\tau_i^*)/\lambda^*(\tau_i^*)$. Os pontos de salto restantes determinam a trajetória de um processo de Poisson não homogêneo $N = (N_t; t \geq 0)$ com intensidade $\lambda(t)$ no intervalo $(0, t_0]$.

O algoritmo para o método descrito com $t_0 = 1$, é como segue:

ALGORITMO 2.

1) Gere uma trajetória de um processo de Poisson $(N_t^*; t \geq 0)$ com função intensidade inversível $\lambda^*(t) \geq \lambda(t)$ em $[0, 1]$, segundo o Algoritmo 1. Seja n^* o número de pontos gerados.
2) Denote os pontos (ordenados) obtidos por $\tau_1^*, \tau_2^*, \ldots, \tau_{n^*}^*$. Faça $i = 1$ e $k = 0$.
3) Gere u_i, uniformemente distribuído em $[0, 1]$. Se $u_i \leq \lambda(\tau_i^*)/\lambda^*(\tau_i^*)$, faça $k = k+1$ e $\tau_k = \tau_i^*$. Caso contrário, delete τ_i^*, faça $i = i+1$ e repita 3.
4) Faça $i = i+1$. Se $i \leq n^*$, vá para 3.
5) Se $i > n^*$, retorne $\tau_1, \tau_2, \ldots, \tau_k$ e também k.

Os pontos resultantes (não deletados) representam uma trajetória de um processo de Poisson $(N_t; t \geq 0)$ com função intensidade $\lambda(t)$. A função $\lambda^*(t)$ do algoritmo deveria ser

o mais próxima possível de $\lambda(t)$ e ao mesmo tempo permitir gerar facilmente o processo $(N^*(t); t \geq 0)$. Um caso simples de operar, e que aqui adotaremos, é fazer $\lambda^*(t) = \lambda^* = \max_{0 \leq t \leq 1} \lambda(t)$. Com essa condição, a eficiência obtida medida pelo número de pontos deletados, não é a melhor possível. Em compensação, estaremos gerando trajetórias com uma intensidade constante, o que é operacionalmente muito conveniente.

4.B. ASPECTOS PRÁTICOS DO MÉTODO DE VALIDAÇÃO CRUZADA.

Usando as notações da Seção 3.F, seja

$$\hat{\lambda}_h(t) = \sum_{i=1}^{N} K_h(t-\tau_i), \quad t \in [0,1] \tag{4.1}$$

o estimador kernel de intensidade $\lambda(t)$ de um processo de Poisson não homogêneo. O método de validação cruzada de mínimos quadrados, descrito em 3.C consiste em escolher para o cálculo do estimador aquela janela h que minimiza a função "score"

$$VC_\lambda(h) = \int_0^1 \hat{\lambda}_h^2(t)\,dt - 2\sum_{i=1}^{N} \hat{\lambda}_{hi}(\tau_i) \tag{4.2}$$

Usando uma ligeira modificação da fórmula obtida por Silverman (1983) para o "score" e que foi dada em (3.22), obtemos

$$VC_1(h) =$$

$$= \frac{1}{h}\sum_{i=1}^{N}\sum_{j=1}^{N} K^{(2)}\left(\frac{\tau_i-\tau_j}{h}\right) - \frac{2}{h}\sum_{i=1}^{N}\sum_{j=1}^{N} K\left(\frac{\tau_i-\tau_j}{h}\right) + \frac{2N}{h}K(0)\,. \tag{4.3}$$

Aqui $K^{(2)}$ é a convolução do kernel consigo mesmo e foi assumido K simétrico. Tentar minimizar diretamente (4.3) é altamente ineficiente, pois aparecem $N(N-1)/2$ cálculos das funções $K^{(2)}$ e K para obtenção de cada valor de $VC_1(h)$, o que requer muito tempo computacional.

Um algoritmo eficiente para minimizar $VC(h)$ foi dado no contexto de estimação de densidades por Lee e Kim (1991) para kernels polinomiais simétricos. Dado que o kernel de Epanechnikov é um dos mais utilizados na literatura e satisfaz a condição de ser polinômio simétrico, será este o kernel por nós adotado no trabalho computacional.

O kernel de Epanechnikov é definido por

$$K(t) = \begin{cases} \frac{3}{4\sqrt{5}}\left(1-\frac{1}{5}t^2\right), & se\ |t|<\sqrt{5} \\ 0 \qquad\qquad , & caso\ contrário\,. \end{cases} \tag{4.4}$$

Alguns cálculos levam a:

$$K^{(2)}(t) = \begin{cases} \frac{3}{40}\left(-\frac{1}{500}|t|^5+\frac{1}{5}|t|^3-\frac{2\sqrt{5}}{5}t^2+\frac{8\sqrt{5}}{5}\right), & se\ |t|<2\sqrt{5} \\ 0 \qquad\qquad\qquad\qquad\qquad\qquad , & caso\ contrário\,. \end{cases} \tag{4.5}$$

Adaptaremos agora o algoritmo de Lee e Kim ao nosso

contexto de estimação da intensidade de um processo de Poisson. A idéia básica é extrair h de K e $K^{(2)}$ sem calcular as somas de K e $K^{(2)}$ para cada h .

Suponha que observemos N pontos de salto $\tau_1, \ldots, \tau_N$ de um processo de Poisson não homogêneo com intensidade $\lambda(t)$.

Antes do algoritmo é preciso calcular as diferenças absolutas $\tau_i - \tau_j$ como

$$d_1' = |\tau_1 - \tau_2|, d_2' = |\tau_1 - \tau_3|, \ldots, d_{N-1}' = |\tau_1 - \tau_N|$$

$$d_N' = |\tau_2 - \tau_3|, \ldots, d_{2N-3}' = |\tau_2 - \tau_N|$$

$$d_{N(N-1)/2}' = |\tau_{N-1} - \tau_N|$$

Dado que o kernel é simétrico, é claro que $K(\tau_i - \tau_j/h) = K(\tau_j - \tau_i/h) = K(|\tau_i - \tau_j|/h)$, para todo i, j . Seja o vetor $(d_1, d_2, \ldots, d_{N(N-1)/2})$ uma ordenação crescente do vetor $(d_1', d_2', \ldots, d_{N(N-1)/2}')$. Então construímos o vetor

$$(S_1, S_2, \ldots, S_{N(N-1)/2}),$$

onde $S_i = \sum_{k=1}^{i} d_k$, para $i = 1, 2, \ldots, N(N-1)/2$.

A parte mais trabalhosa na implementação do algoritmo é o cálculo de $VC(h)$ para o kernel escolhido.

Chamando de m o número de pares $i<j$ tais que $\tau_i = \tau_j$ e fazendo $t = (\tau_i - \tau_j)/h$, obtém-se:

$$VC(h) = \frac{3}{100h}\left\{-\frac{1}{100}\sum_{i,j} t^5 \mathbf{I}_{(0,2\sqrt{5})}(t) + \sum_{i,j} t^3 \mathbf{I}_{(0,2\sqrt{5})}(t)\right.$$

$$\left. -2\sqrt{5}\sum_{i,j} t^2 \mathbf{I}_{(0,2\sqrt{5})}(t) + 4\sqrt{5}\left[2\left(\sum_{i,j} \mathbf{I}_{(0,2\sqrt{5})}(t) + m\right) + N\right]\right\} \quad (4.6)$$

$$-\frac{3}{\sqrt{5}h}\left\{-\frac{1}{5}\sum_{i,j} t^2 \mathbf{I}_{(0,\sqrt{5})}(t) + \left(\sum_{i,j} \mathbf{I}_{(0,\sqrt{5})}(t) + m\right)\right\}$$

O algoritmo para obtenção da janela de validação cruzada é como segue:

ALGORITMO 3.

1) Inicializar h (como um número grande).
2) Encontrar os maiores inteiros n_1, n_2 tais que $d_{n_1} < \sqrt{5}h$ e $d_{n_2} < 2\sqrt{5}h$.
3) Salvar h que minimiza

$$VC(h) =$$

$$= \frac{3}{100h}\left\{-\frac{1}{100h^5}S_{n_2}{}^5 + \frac{1}{h^3}S_{n_2}{}^3 - \frac{2\sqrt{5}}{h^2} + 4\sqrt{5}(2n_2 + N)\right\} - \frac{3}{\sqrt{5}h}\left\{-\frac{1}{5h^2}S_{n_1}{}^2 + n_1\right\}.$$

$VC(h)$ será computado para cada h pertencente a um "grid" formado sobre um intervalo adequado de possíveis janelas. Um outro algoritmo para calcular $VC(h)$ foi dado por Silverman (1982), mas necessita de certos cálculos e inversões de Fourier, nem sempre fáceis, e de certas subrotinas de transformada rápida de Fourier, nem sempre disponíveis.

Resta ainda um aspecto a ser tomado em conta na

nossa estimação. Dado que estamos estimando uma função intensidade sobre um intervalo finito, é importante considerar efeitos de fronteira. Em particular, não existem observações fora do intervalo $[0,1]$ e logo o estimador kernel é artificialmente baixo nos extremos. Além disso, quando não são introduzidas correções de fronteira, o estimador kernel da intensidade terá massa fora do intervalo de estudo, e portanto $\int_0^1 \hat{\lambda}_h(t)\,dt$ será menor que o número de observações médias em $[0,1]$. Uma correção, sugerida por Schuster (1985), é a de "imagem de espelho" que faz com que a área das funções kernel que se estendem além dos extremos do intervalo seja revertida sobre $[0,1]$, ficando o total da área contida nesse intervalo.

Então, o estimador kernel corrigido será:

$$\hat{\lambda}_h^*(t) = \sum_{i=1}^{N} [K_h(t-\tau_i) + K_h(t-\tau_i) + K_h(2-t-\tau_i)], \qquad (4.7)$$

para $t \in [0,1]$.

4.C. ASPECTOS COMPUTACIONAIS.

Os programas para geração das trajetórias do processo de Poisson e de validação cruzada para obter a janela ótima foram feitos em Fortran 77. Para a geração de números aleatórios em $[0,1]$ foi utilizado o algoritmo de Wichmann e Hill (1982), que proporciona um gerador de período maior que 2.78×10^{13} e é constituído por três geradores congruenciais multiplicativos simples.

O cálculo da janela de validação cruzada de mínimos

quadrados foi feita utilizando o Algoritmo 3, dado em 4.B, com os valores de h variando sobre $(0,2)$ com incrementos de 0.01. O valor inicial de h foi assumido como 999, mas isto é irrelevante.

A fim de ter um critério para comparar a qualidade dos estimadores foi calculada a soma dos desvios quadráticos sobre cada ponto x_i de um "grid" de $[0,1]$, isto é,

$$SDQ = \sum_i \left(\hat{\lambda}_h(x_i) - \lambda(x_i)\right)^2. \qquad (4.8)$$

O método de ordenação necessário ao programa foi o de "Heapsort", por ser muito rápido e não requerer grande capacidade de memória.

Todos os programas foram executados na estação de trabalho Sun para um número de trajetórias que depende da intensidade. Por exemplo, para $\lambda_1(t) = 2t+1$, $n = 200$ trajetórias originaram em média 407 pontos amostrais (tempos de salto), enquanto para $\lambda_2(t) = 12e^{-2t}$, $n = 200$ trajetórias originaram em média 635 pontos, o qual produziu overflow.

4.D. ANÁLISE DOS RESULTADOS.

Os resultados obtidos utilizando amostras simuladas, cujos gráficos são apresentados no Apêndice B, mostram a eficácia do método de validação cruzada para selecionar a melhor janela possível, no sentido de mínimo erro quadrático integrado.

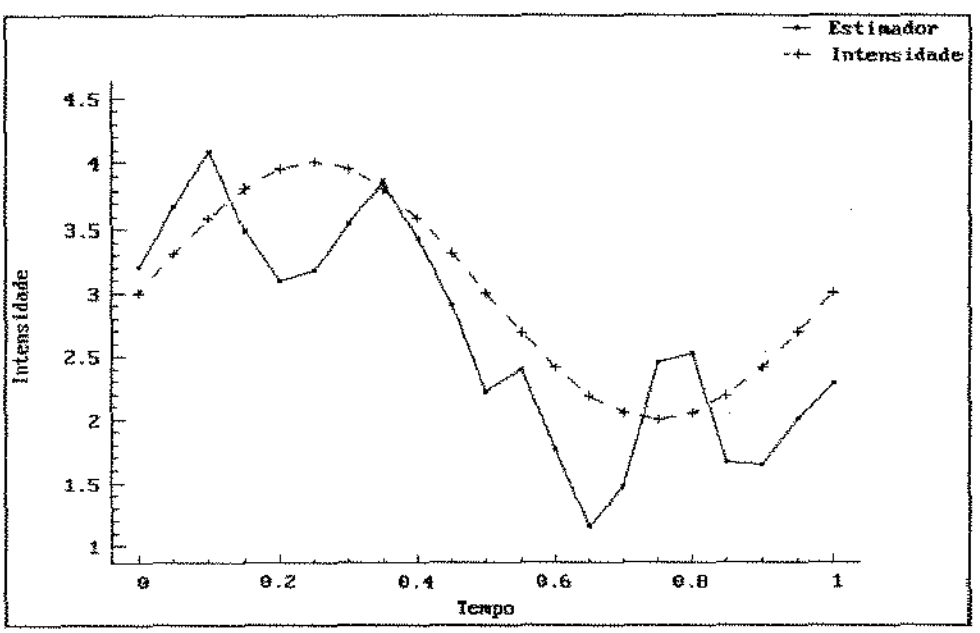
A fim de ilustrar graficamente a otimalidade da janela

de validação cruzada, em termos de suavidade do estimador obtido, damos a continuação os gráficos do estimador kernel para três valores diferentes de janela.

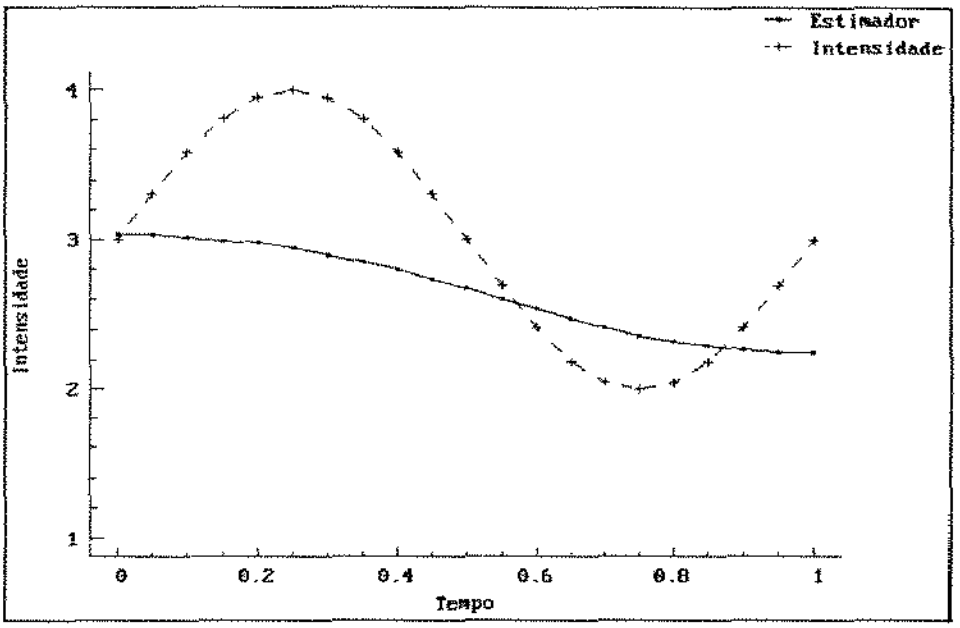
$$\alpha(t) = 3 + sen(2\pi t)$$

$$N = 70 \text{ trajetórias}$$

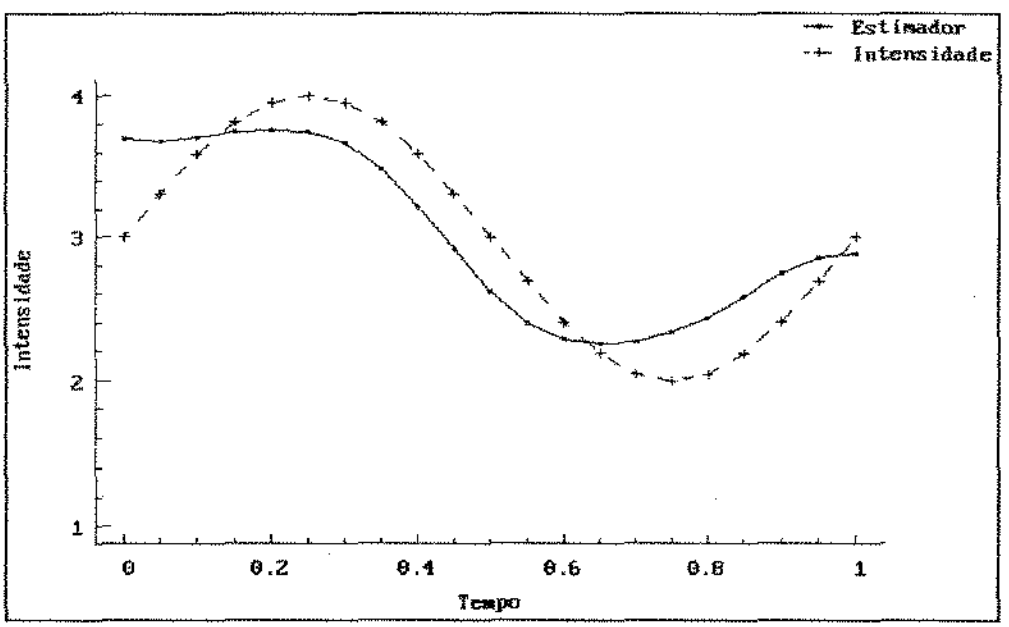
Estimador com Correção de Fronteira.



h = 0.03


Estimador
Intensidade
Intensidade
Tempo
1
2
3
4
0
0.2
0.4
0.6
0.8
1

h = 0.40


Estimador
Intensidade
Intensidade
Tempo
1
2
3
4
0
0.2
0.4
0.6
0.8
1

h = 0.13

Aqui é possível apreciar graficamente o comentário ao resultado (3.12): um valor de janela muito pequeno (h = 0.03) faz com que apareça grande variância (ruído), enquanto um valor de janela muito grande (h = 0.4) implica em um valor alto de vício (sobre-alisamento). O valor fornecido pelo método de validação cruzada (h = 0.13) mostrou-se apropriado, pois balanceia adequadamente os efeitos do vício e a variância, como esperado.

Nos gráficos do Apêndice B, ressalta à primeira vista a importância da correção de fronteira para o estimador. O estimador corrigido é, sem dúvida, superior ao estimador sem correção. Notemos que para $\alpha_1(t) = 2t+1$ (Gráficos B.1-B.3) a correção foi mais notória no extremo esquerdo, para $\alpha_2(t) = 12e^{-2t}$ (Gráficos B.4-B.6) no extremo direito e para $\alpha_3(t) = 3+sen(2\pi t)$ (Gráficos B.7-B.10) em ambos os extremos; ou seja, a correção é influenciada pela condição da intensidade ser alta ou baixa nos extremos (alta intensidade ≡ efeito de correção mais destacado). Na maioria dos casos, na região de baixa intensidade a correção produziu sobrestimação, ou seja, o valor estimado superou o valor real da função intensidade.

A influência do tamanho de amostra n é também ilustrada observando a evolução do estimador para $\alpha_3(t) = 3+sen(2\pi t)$: à medida que aumentamos o número de trajetórias amostrais utilizadas no cálculo do estimador, este, como esperado, vai sendo melhorado. Porém, foi observado em alguns casos que aumentando o número de trajetórias, o estimador pode piorar em um certo instante. Uma possível explicação deste fenômeno é a não linearidade do

estimador kernel: o estimador é muito sensível às variações de h . Comparar, por exemplo, os estimadores $\lambda_2(t) = 12e^{-2t}$ para $N = 50$ e $N = 100$ (Gráficos B.5 e B.6).

Por último, uma grande vantagem do estimador kernel sobre outros estimadores é que, por não estar restrito a uma família funcional pré-especificada, ele se comporta bem para diferentes formas funcionais da intensidade, o que não acontece, por exemplo, com o estimador de "sieves", onde diferentes "sieves" têm comportamento distinto para diferentes formas funcionais de intensidade [ver por exemplo Gomes (1991)].

Apêndice A

O presente apêndice tenta resumir os conceitos, notações e resultados fundamentais ao estudo de processos pontuais usando teoria de martingales. Para se obter uma visão mais completa e profunda do tema pode-se consultar os textos de Elliot (1982), Jacod e Shiryayev (1986) e Liptser e Shiryayev (1977, 1978).

A.1. TEORIA BÁSICA DE PROCESSOS

Suponha $(\Omega, \mathcal{F}, P)$ um espaço de probabilidade completo fixo.

DEFINIÇÃO A.1. Uma família $F = \{\mathcal{F}_t, t \geq 0\}$ de sub-σ-álgebras de $\mathcal{F}$ é uma *filtragem* se

i) **F** é crescente: $\mathcal{F}_s \subset \mathcal{F}_t$, para $s < t$.

ii) **F** é contínua à direita: $\bigcap_{s>t} \mathcal{F}_s = \mathcal{F}_t$.

iii) **F** é completa: $\mathcal{F}_0$ contém todo conjunto P-nulo de $\mathcal{F}$.

A coleção $(\Omega, \mathcal{F}, P, F)$ é chamada *base estocástica*.

DEFINIÇÃO A.2. Seja $(\Omega, \mathcal{F}, P, F)$ uma base estocástica e seja $X = (X_t, t \geq 0)$ um processo estocástico definido sobre esta base.

a) X é *mensurável* se a função $(t, \omega) \to X(t, \omega)$ é mensurável com relação à σ-álgebra produto $\mathcal{B} \times \mathcal{F}$, onde $\mathcal{B}$ é a σ-álgebra de Borel em $[0, \infty)$.

b) X é *adaptado* à filtragem F (ou F-adaptado) se X_t é $\mathcal{F}_t$-mensurável, para todo $t \geq 0$.

c) X é *integrável* se $\sup_{0 \leq t < \infty} E\,|X_t| < \infty$ e é *uniformemente integrável* se $\lim_{\lambda \to \infty} \sup_{0 \leq t < \infty} E\,|X_t|\, \mathbf{I}(|X_t| \geq \lambda) = 0$.

d) X é dito ser *contínuo à direita* (resp. *contínuo à esquerda*, resp. *contínuo à direita com limites à esquerda*) se todas suas trajetórias são contínuas à direita (resp. contínuas à esquerda, resp. contínuas à direita com limites à esquerda).

Quando X é um processo contínuo à direita com limites à esquerda, definem-se os processos $X_- = (X_{t-}; t \geq 0)$, onde $X_{t-} = \lim_{s \uparrow t} X_s$, para $t > 0$ (com $X_{0-} = X_0$) e $\Delta X = (\Delta X_t; t \geq 0)$, onde $\Delta X_t = X_t - X_{t-}$.

A.2. TEMPOS DE PARADA.

DEFINIÇÃO A.3. Uma variável aleatória τ assumindo valores em $[0,\infty]$ é chamada um *tempo de parada* com relação a **F** se $[\tau \le t] = \{\omega : \tau(\omega) \le t\} \in \mathcal{F}_t$, para todo $t \ge 0$.

Se τ é um tempo de parada, denota-se por $\mathcal{F}_\tau$ a σ-álgebra de conjuntos $A \in \mathcal{F}$ tais que

$$A \cap [\tau \le t] \in \mathcal{F}_t, \ \forall t \ge 0,$$

e denota-se por $\mathcal{F}_{\tau-}$ a σ-álgebra gerada por $\mathcal{F}_0$ e todos os conjuntos da forma

$$A \cap [\tau > t], \ A \in \mathcal{F}_t \ \text{e} \ t \ge 0.$$

Algumas propriedades dos tempos de parada são dadas a seguir.

PROPOSIÇÃO A.1. Seja a uma constante. Então,

a) $\tau = a \ge 0$ é um tempo de parada.

b) se τ é um tempo de parada e $a \ge 0$, $\tau + a$ é um tempo de parada.

c) se τ é um tempo de parada e se $A \in \mathcal{F}_\tau$,

$$\tau_A(\omega) = \begin{cases} \tau(\omega), & se\ \omega \in A \\ +\infty, & se\ \omega \notin A \end{cases}$$

é também um tempo de parada.

DEFINIÇÃO A.4. (PROCEDIMENTO DE LOCALIZAÇÃO).

Se $\mathcal{E}$ é uma classe de processos, denota-se por $\mathcal{E}_{loc}$ a *classe localizada* definida como segue: um processo X pertence a $\mathcal{E}_{loc}$ se existe uma seqüência crescente (τ_n) de tempos de parada tal que $\lim \tau_n = +\infty$ (q.c.) e cada processo "parado" $X^{\tau_n} = (X_{\tau_n \wedge t};\ t \geq 0)$ pertence a $\mathcal{E}$.

Por exemplo, se $\mathcal{E}$ é a classe de processos integráveis, $\mathcal{E}_{loc}$ será a classe dos chamados processos localmente integráveis.

A.3. PROCESSOS PREDIZÍVEIS.

Uma classe de processos muito importantes em integração estocástica é a dos processos predizíveis, definida à continuação.

DEFINIÇÃO A.5. a) A σ-*álgebra predizível* é a σ-álgebra $\mathcal{P}$ em $\Omega \times [0, \infty)$ gerada por todos os processos contínuos à esquerda e adaptados.

b) Um processo que é $\mathcal{P}$-mensurável é chamado *predizível*.

TEOREMA A.1. A σ-álgebra predizível também é gerada por qualquer uma das seguintes coleções de conjuntos aleatórios:

i) $A \times \{0\}$ e $[\![0, \tau]\!]$; onde $A \in \mathcal{F}_0$, τ é um tempo de parada e

$$[\![0, \tau]\!] = \{(\omega, t) : 0 \leq t \leq \tau(\omega) ; t \geq 0)\}.$$

ii) $A \times \{0\}$ e $B \times (s, t]$; onde $A \in \mathcal{F}_0$, $B \in \mathcal{F}_s$ e $s < t$.

PROPOSIÇÃO A.2 Se X é um processo predizível e τ é um tempo de parada, $X_\tau \mathbb{I}(\tau < \infty)$ é $\mathcal{F}_{\tau-}$ -mensurável.

Esta proposição ajuda a esclarecer a adoção do termo "predizível". Interpretando $\mathcal{F}_{\tau-}$ como o passado estrito ao tempo aleatório τ , heuristicamente a proposição diz que um processo é predizível quando o conhecimento de seus valores

em $[0, t)$ determina seu valor em t .

A.4. TEMPOS PREDIZÍVEIS E TOTALMENTE INACESSÍVEIS.

DEFINIÇÃO A.6. Um *tempo predizível* é um tempo de parada τ tal que o processo $\mathbf{I}(t \leq \tau)$ é predizível.

Uma classe de tempos de parada "opostos" em algum sentido aos tempos predizíveis é introduzida na seguinte definição.

DEFINIÇÃO A.7. Um tempo de parada σ é chamado *tempo totalmente inacessível* se $P(\sigma = \tau < \infty) = 0$, para todo tempo predizível τ .

Em geral tempos predizíveis correspondem a eventos predizíveis e tempos totalmente inacessíveis a eventos não predizíveis.

DEFINIÇÃO A.8. Um processo $X = (X_t; t \geq 0)$ contínuo à direita com limites à esquerda é chamado *quase contínuo à esquerda* se $\Delta X_\tau = 0$

(q.c.) no conjunto $\{\tau<\infty\}$, para todo tempo τ predizível.

PROPOSIÇÃO A.3. Seja X um processo adaptado contínuo à direita com limites à esquerda. As seguintes afirmações são equivalentes:

a) X é quase-contínuo à esquerda.

b) Existe uma seqüência de tempos totalmente inacessíveis que contém todos os saltos de X .

c) Para qualquer seqüência crescente (τ_n) de tempos de parada com limite τ , tem-se $\lim_{n\to\infty} X_{\tau_n} = X_\tau$ (q.c.), sobre $\{\tau<\infty\}$.

A.5. MARTINGALES.

Um dos conceitos fundamentais na teoria moderna dos processos estocásticos é o conceito de martingale.

DEFINIÇÃO A.9. Um processo $M = (M_t; t\geq 0)$ adaptado, contínuo à direita com limites à esquerda e $E|M_t|<\infty$, $\forall t>0$, é um *martingale* (ou um $\mathcal{F}_t$ -martingale) se $E(M_t/\mathcal{F}_s) = M_s$ (q.c.), para todo $0\leq s\leq t$. M é um *submartingale* se em lugar da igualdade temos $E(M_t/\mathcal{F}_s)\geq M_s$ e um *supermartingale* caso a

desigualdade seja oposta.

Denota-se por $\mathbf{M}$ a classe de todos os martingales uniformemente integráveis e por $\mathbf{M}^2$ a classe de martingales de quadrado integrável.

DEFINIÇÃO A.10. Um *martingale local* (resp. um *martingale localmente de quadrado integrável*) é um processo que pertence à classe localizada $\mathbf{M}_{loc}$ (resp $\mathbf{M}^2_{loc}$), construída a partir de M (resp M^2) usando a Definição A.4.

O seguinte resultado proporciona uma caracterização simples de um submartingale (martingale)

PROPOSIÇÃO A.4. Uma condição necessária e suficiente para que o processo integrável $M = (M_t; t \geq 0)$ seja um submartingale (resp. martingale) é que para qualquer tempo de parada τ que assume dois valores, se verifique que $E(M_\tau) \geq E(M_0)$ (resp. $E(M_\tau) = E(M_0)$).

A.6. DECOMPOSIÇÃO DE DOOB-MEYER.

DEFINIÇÃO A.11. Um processo $X = (X_t; t \geq 0)$ é *de classe D* se

o conjunto de variáveis aleatórias $\{X_\tau : \tau < \infty$ *é tempo de parada*$\}$ é uniformemente integrável.

DEFINIÇÃO A.12. Um *processo crescente* $\mathbf{A} = (A_t; t \geq 0)$ é um processo adaptado, contínuo à direita com limites à esquerda com $A_0 = 0$ e cada trajetória A_t não decrescente.

O seguinte resultado é de fundamental importância na teoria de processos estocásticos.

TEOREMA A.2. (DECOMPOSIÇÃO DE DOOB-MEYER PARA SUBMARTINGALES)

Se X é um submartingale de classe D, existe um único (exceto equivalência estocástica) processo $\mathbf{A}$ crescente, integrável e predizível tal que $(X-\mathbf{A})$ é um martingale uniformemente integrável.

Se $M \in \mathbf{M}^2_{loc}$, a decomposição de Doob-Meyer garante que existe um único processo predizível crescente $\langle M \rangle = (\langle M \rangle_t; t \geq 0)$ tal que o processo $(M_t^2 - \langle M \rangle_t; t \geq 0)$ é um martingale local.

DEFINIÇÃO A.13. O processo $\langle M\rangle = (\langle M\rangle_t; t\geq 0)$ que aparece acima é chamado de *variação predizível* ou *processo variância* de M.

Um outro resultado de grande importância em integração estocástica está contido no seguinte teorema.

TEOREMA A.3. Seja $M = (M_t, \mathcal{F}_t; t\geq 0)$ um martingale e $H = (H_t; t\geq 0)$ um processo predizível limitado. Então o processo $L = (L_t, \mathcal{F}_t; t\geq 0)$ definido por

$$L_t = \int_0^t H_s dM_s$$

é um $\mathcal{F}_t$-martingale com variação predizível

$$\langle L\rangle_t = \int_0^t H_s^2 d\langle M\rangle_s .$$

A.7. RESULTADOS DETERMINÍSTICOS.

Os dois resultados a seguir foram utilizados no Capítulo 2.

TEOREMA A.4. (FÓRMULA PRODUTO).

Sejam $f(t)$ e $g(t)$ duas funções de variação limitada

sobre intervalos finitos, contínuas à direita com limites à esquerda. Então

$$f(t)g(t) = f(0)g(0) + \int_0^t f(s)\,dg(s) + \int_0^t g(s-)\,df(s)$$

TEOREMA A.5. (FÓRMULA EXPONENCIAL).

Seja $a(t)$ uma função crescente e contínua à direita com $a(0) = 0$ e seja $u(t)$ tal que

$$\int_0^t |u(s)|\,da(s) < \infty, \quad t \geq 0.$$

Então a equação

$$x(t) = x(0) + \int_0^t x(s-)\,u(s)\,da(s)$$

admite uma única solução localmente limitada dada por

$$x(t) = x(0) \prod_{0<s\leq t} [1 + u(s)\Delta a(s)] \exp\left(\int_0^t u(s)\,da^c(s)\right),$$

onde $\Delta a(t) = a(t) - a(t-)$ e $a^c(t)$ é a parte contínua de $a(t)$:

$$a^c(t) = a(t) - \sum_{s\leq t} \Delta a(s).$$

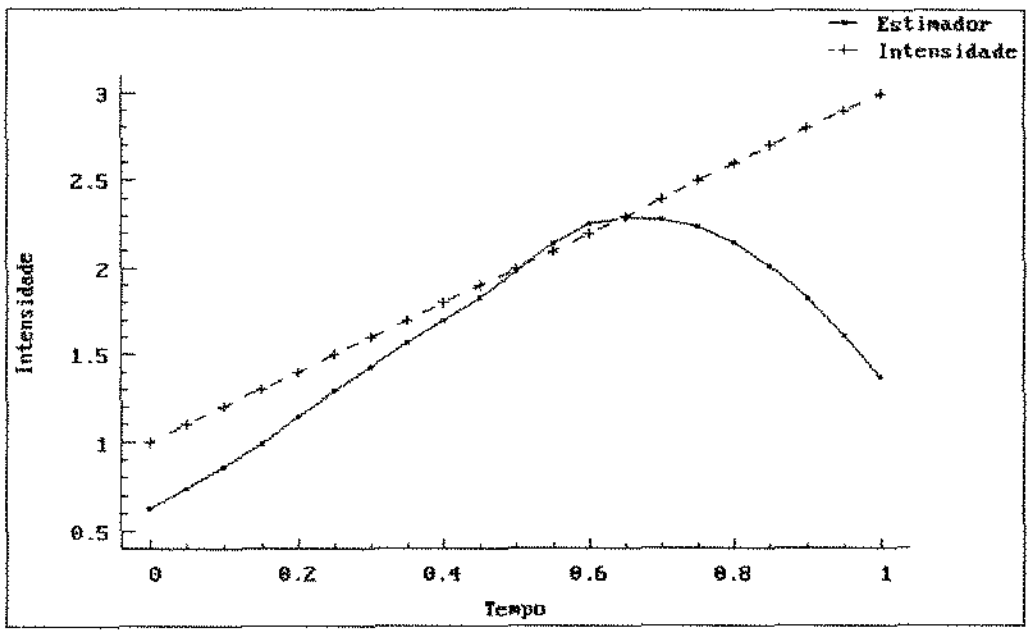
Apêndice B - Gráficos

Nos gráficos a seguir $\alpha(t)$ corresponde à intensidade real do processo pontual, h à janela obtida por validação cruzada, e SDQ à "soma dos desvios quadráticos", calculada sobre os pontos do "grid".

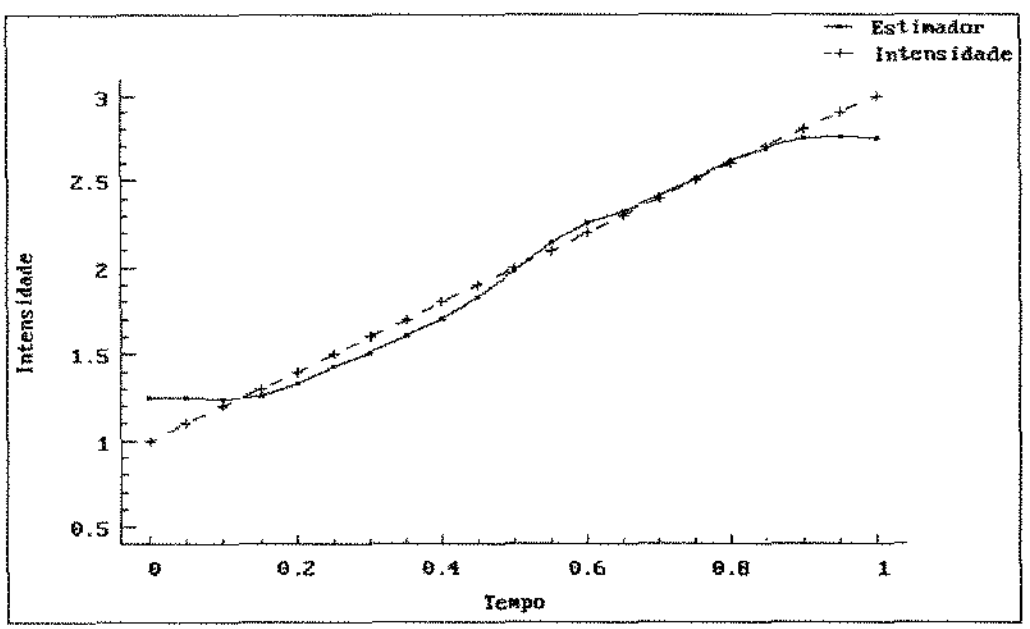
GRÁFICO 1.

$\alpha(t) = 2t+1$

$N = 50$ *trajetórias*, $h = 0.19$



a) Estimador sem correção, SDQ = 6.66.

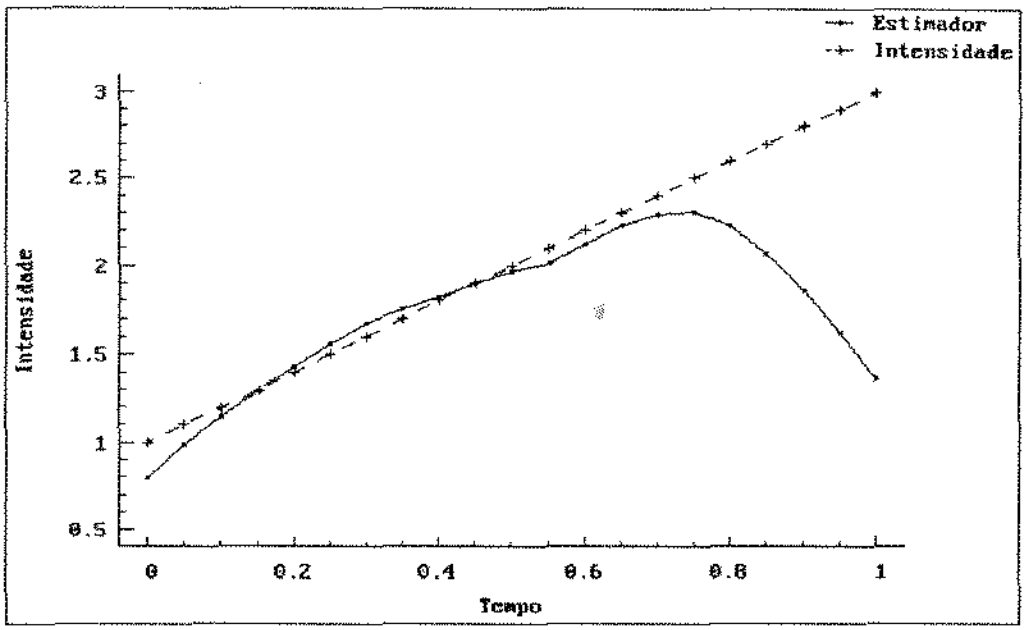


b) Estimador com correção, SDQ = 0.22.

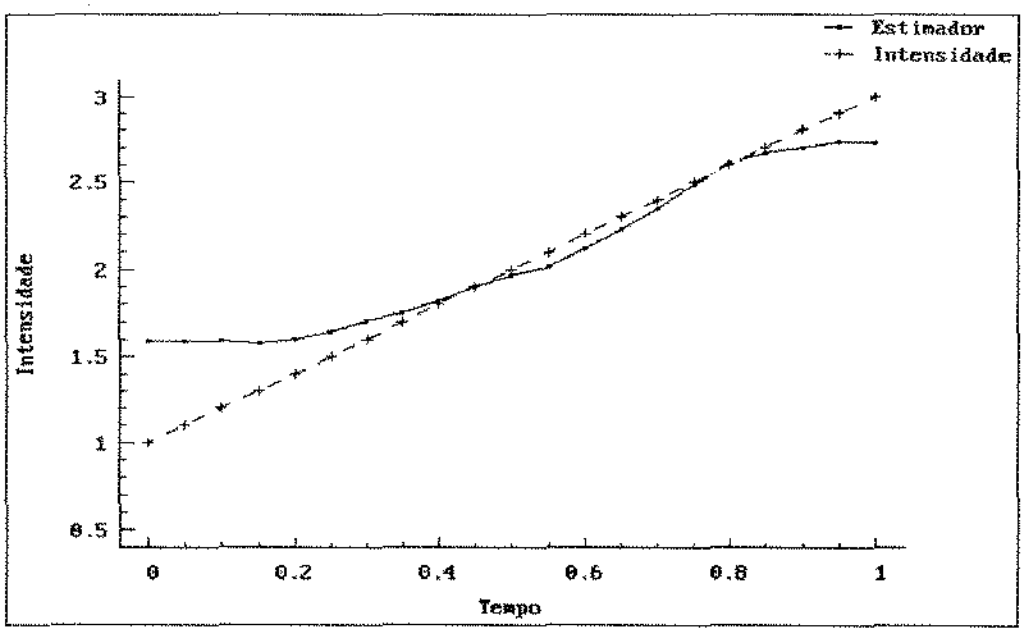
GRÁFICO 2.

$\alpha(t) = 2t+1$

$N = 100$ *trajetórias,* $h = 0.16$



a) Estimador sem correção, SDQ = 5.84.

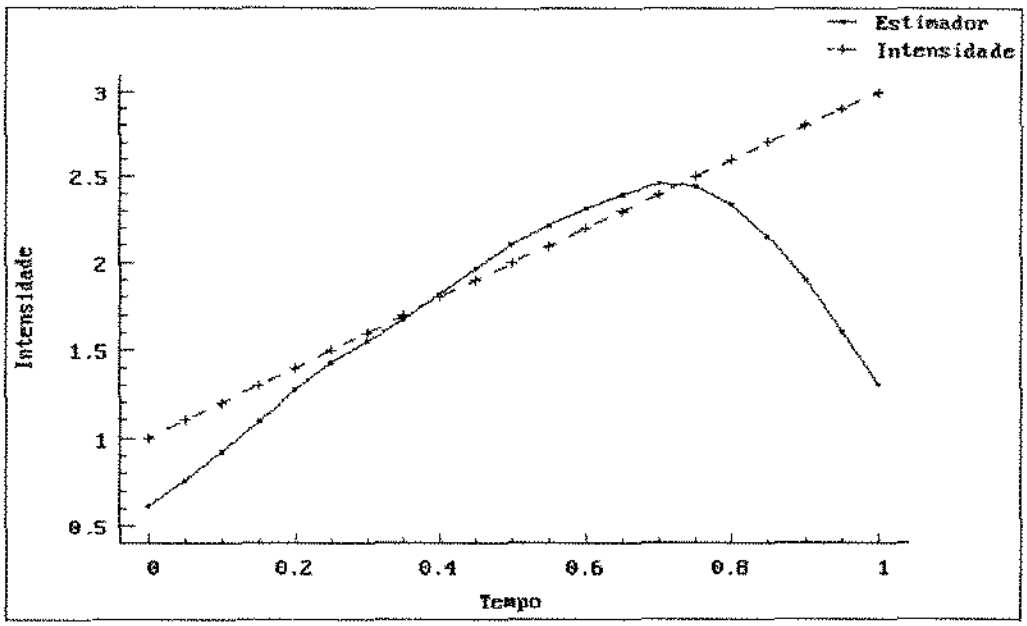


b) Estimador com correção, SDQ = 1.02.

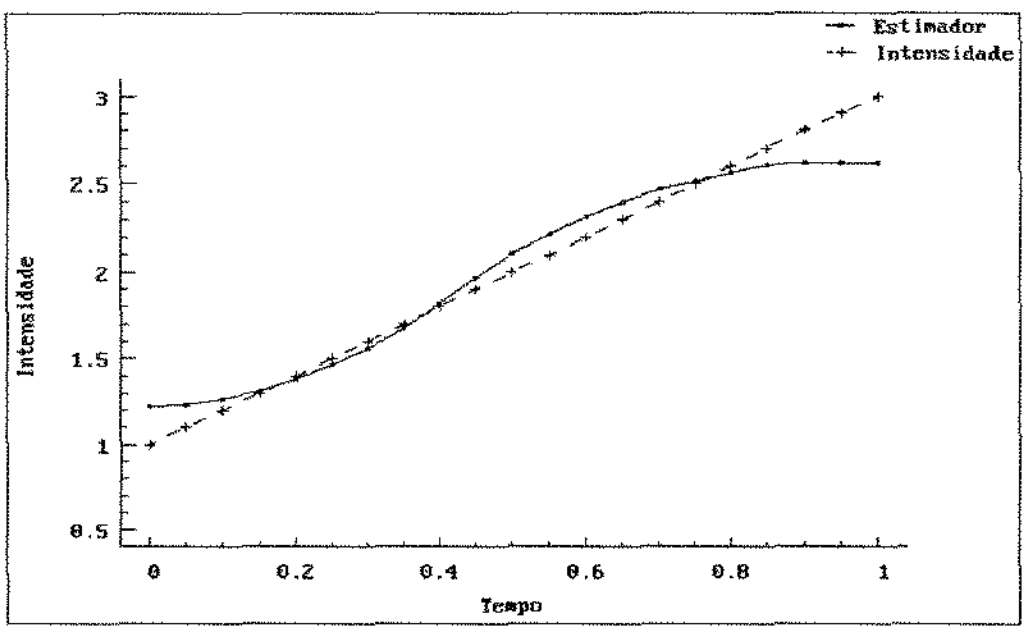
GRÁFICO 3.

$\alpha(t) = 2t+1$

$N = 200$ ***trajetórias,*** $h = 0.14$



a) Estimador sem correção, SDQ = 6.17.

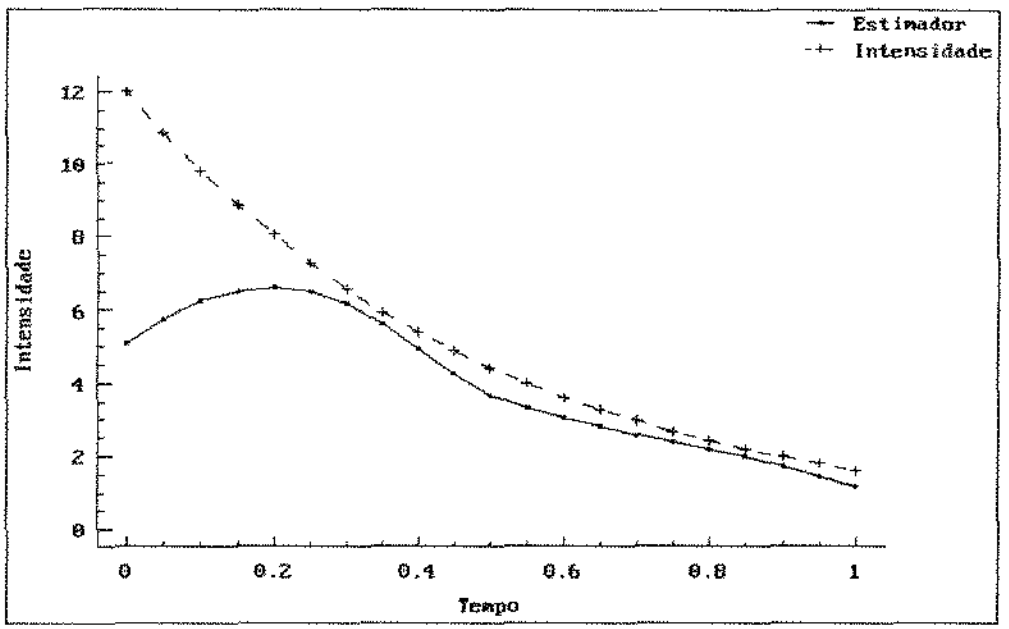


b) Estimador com correção, SDQ = 0.41.

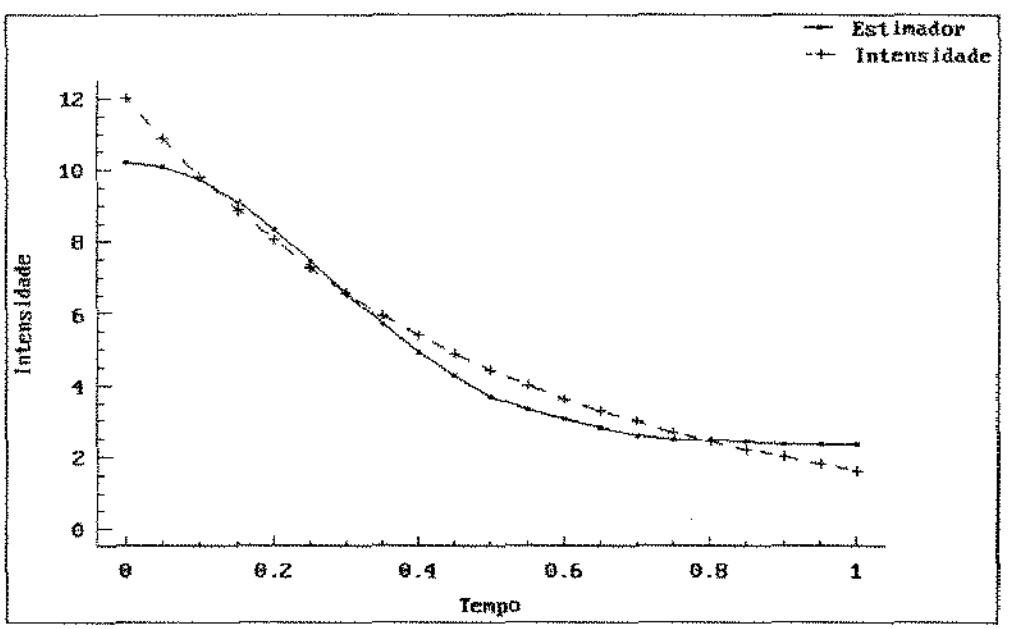
GRÁFICO **4.**

$\alpha(t) = 12e^{-2t}$

$N = 10$ *trajetórias*, $h = 0.17$



a) Estimador sem correção, SDQ = 97.84.

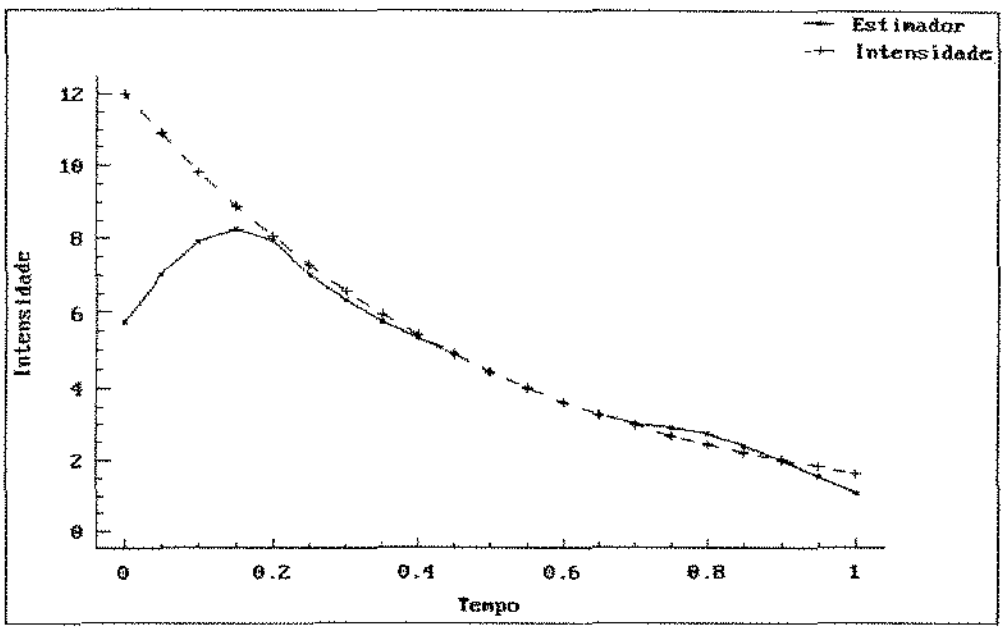


b) Estimador com correção, SDQ = 7.34.

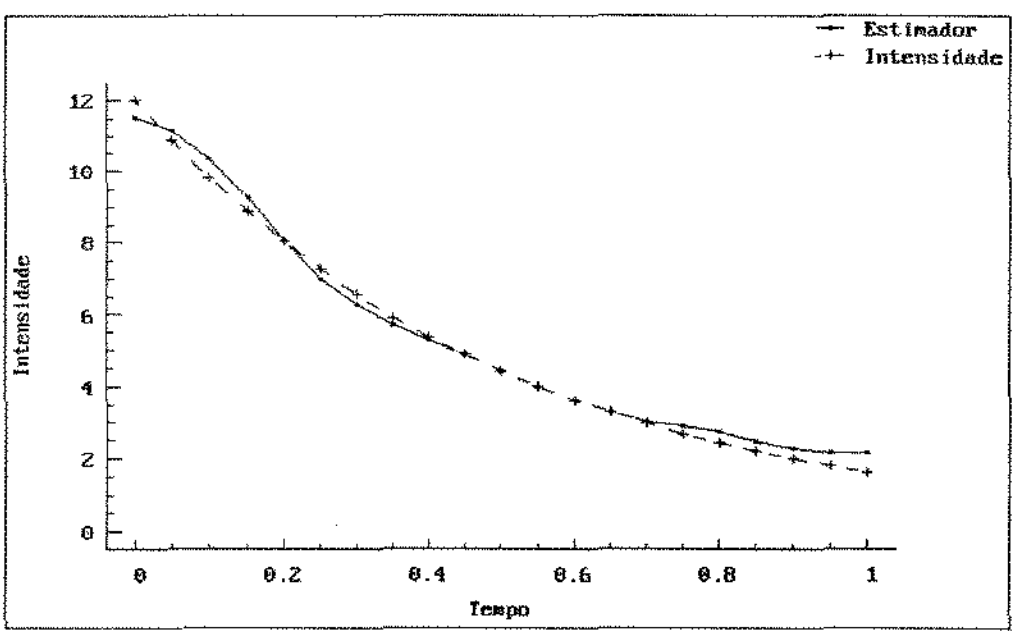
GRÁFICO 5.

$\alpha(t) = 12e^{-2t}$

$N = 50$ *trajetórias*, $h = 0.10$



a) Estimador sem correção, SDQ = 59.06.

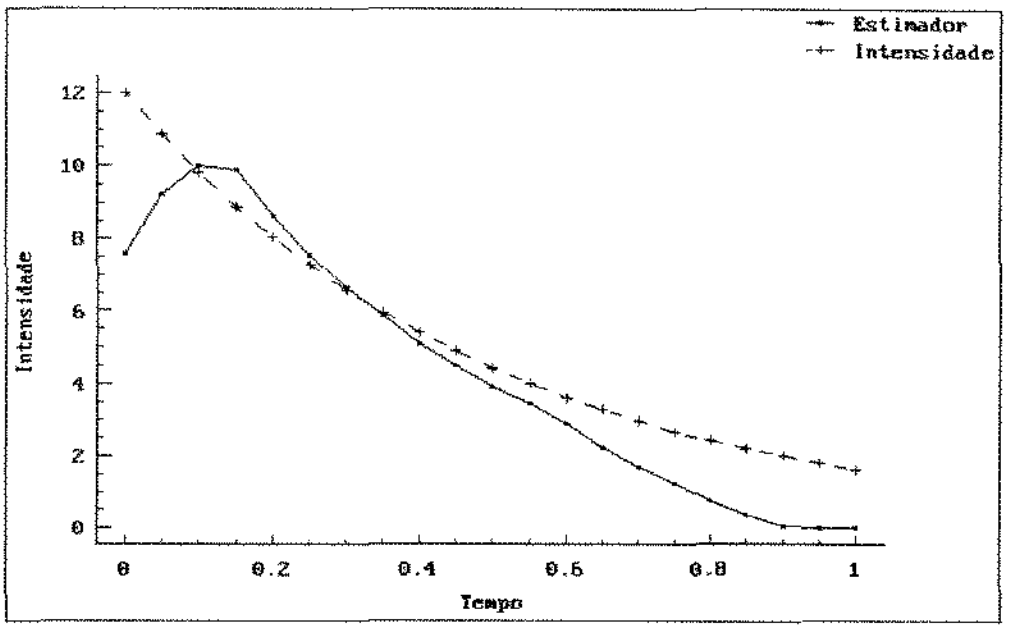


b) Estimador com correção, SDQ = 1.62.

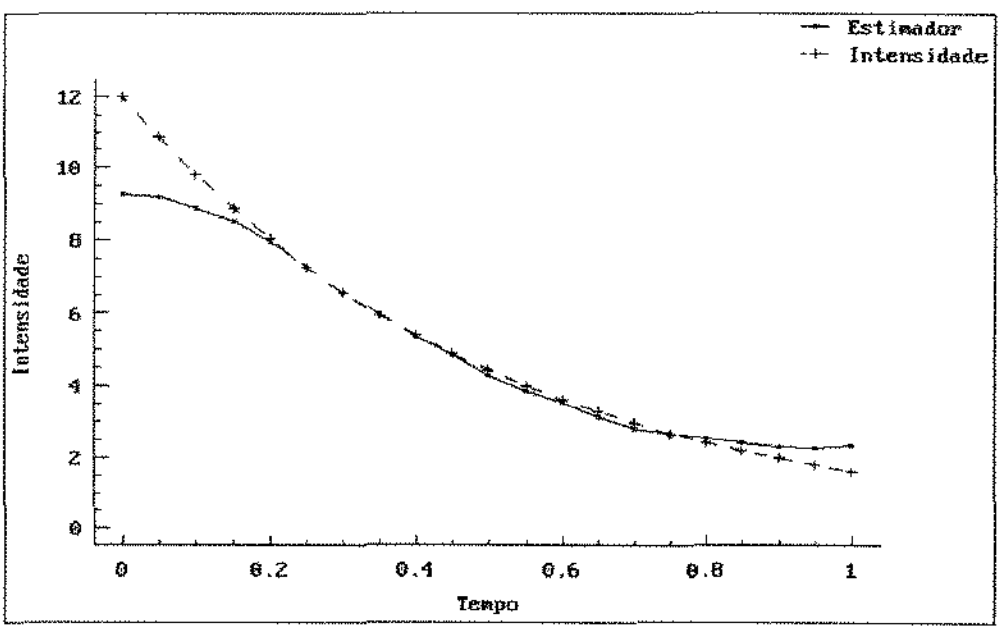
GRÁFICO 6.

$\alpha(t) = 12e^{-2t}$

N = 100 trajetórias, h = 0.09



a) Estimador sem correção, SDQ = 45.59.

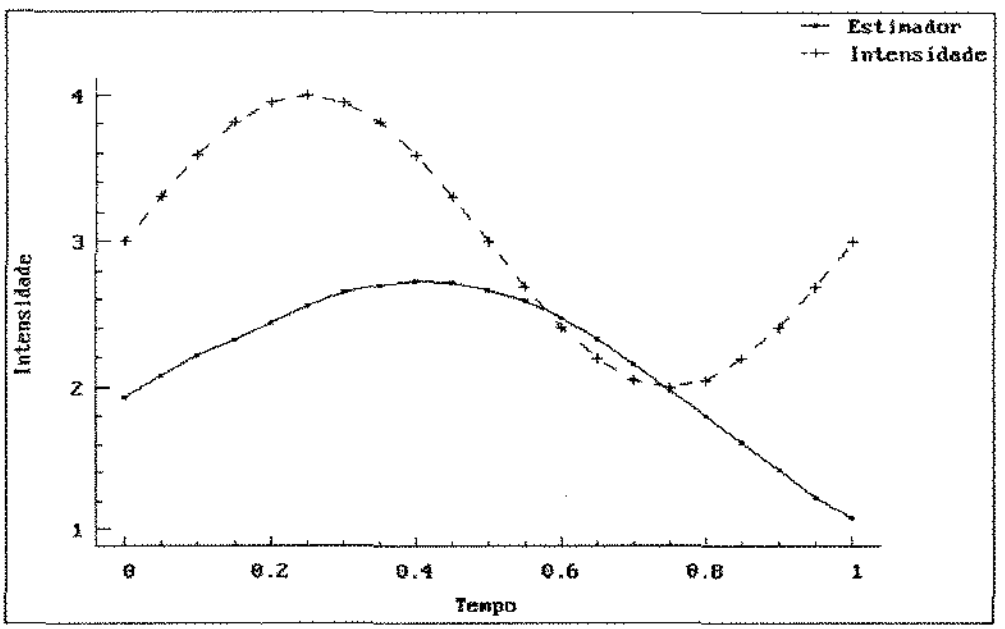


b) Estimador com correção, SDQ = 12.33.

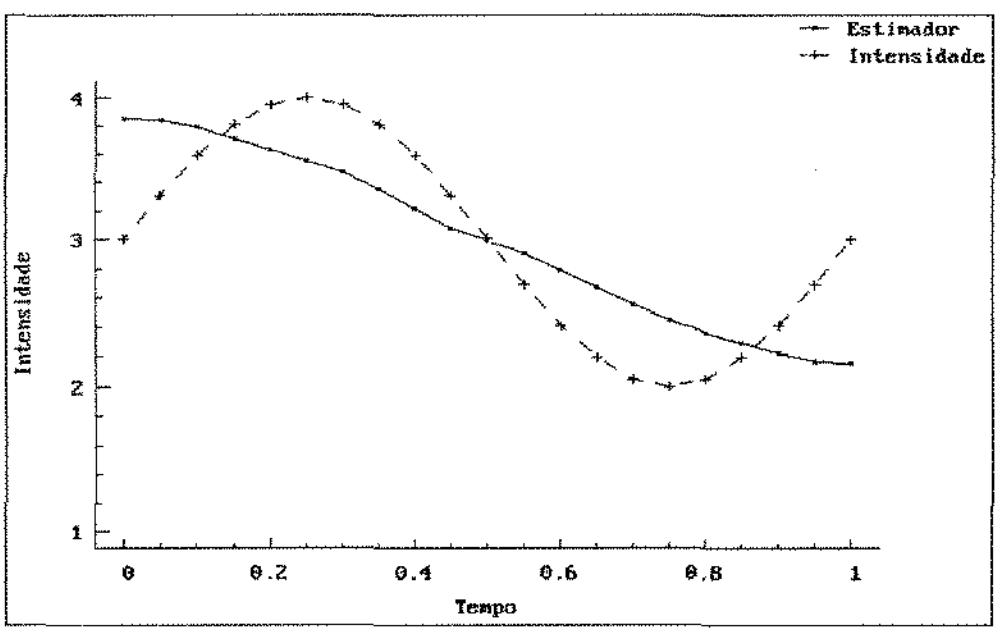
GRÁFICO 7.

$\alpha(t) = 3 + sen(2\pi t)$

$N = 10$ *trajetórias*, $h = 0.31$



a) Estimador sem correção, SDQ = 22.53.

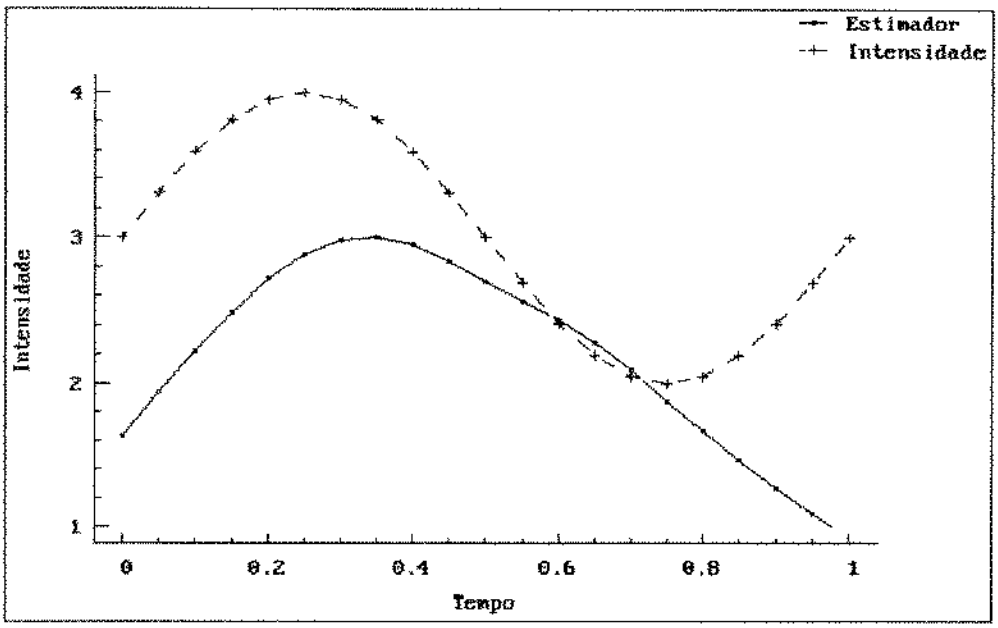


b) Estimador com correção, SDQ = 4.00.

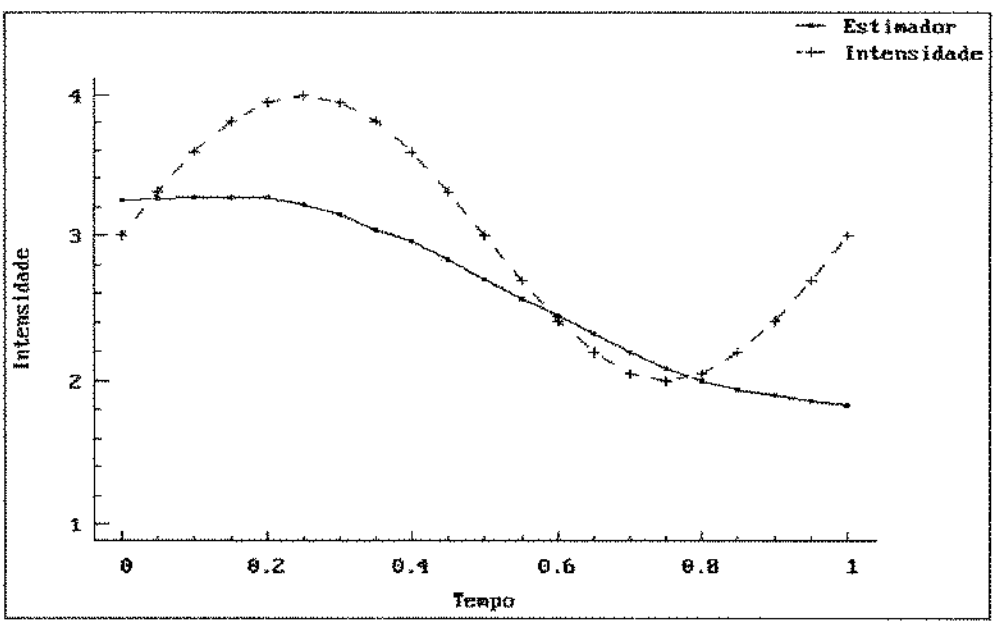
GRÁFICO 8.

$\alpha(t) = 3 + sen(2\pi t)$

$N = 50$ *trajetórias*, $h = 0.19$



a) Estimador sem correção, SDQ = 21.41.

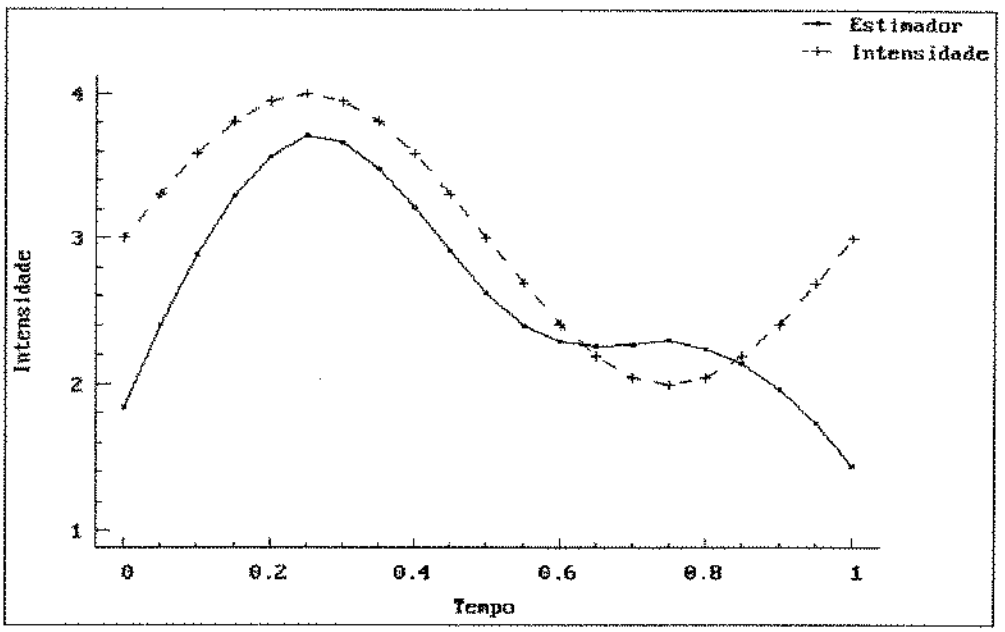


b) Estimador com correção, SDQ = 5.92.

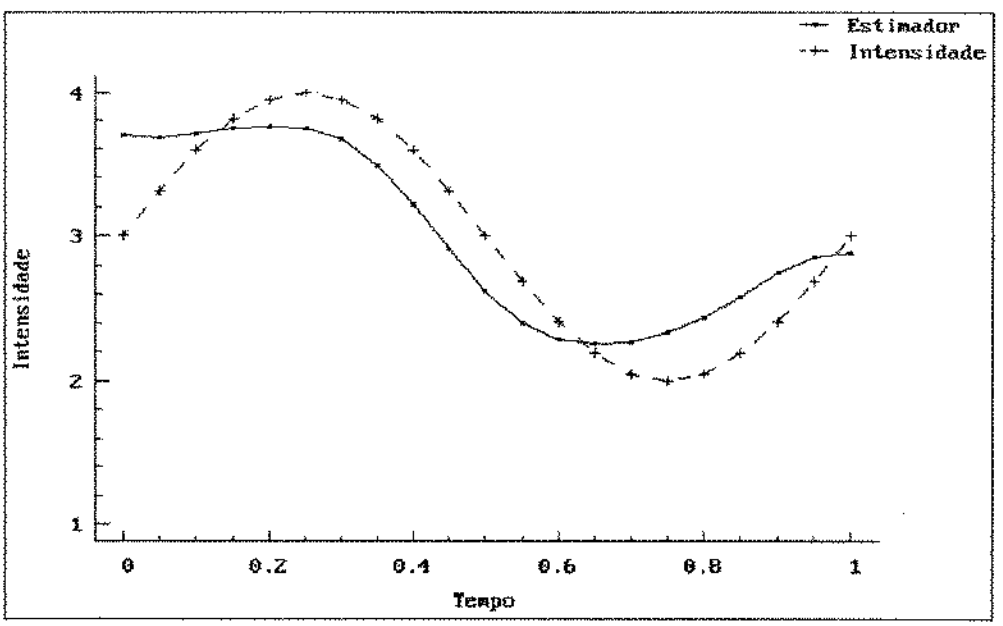
GRÁFICO 9.

$\alpha(t) = 3 + sen(2\pi t)$

$N = 70$ *trajetórias*, $h = 0.13$



a) Estimador sem correção, SDQ = 7.56.

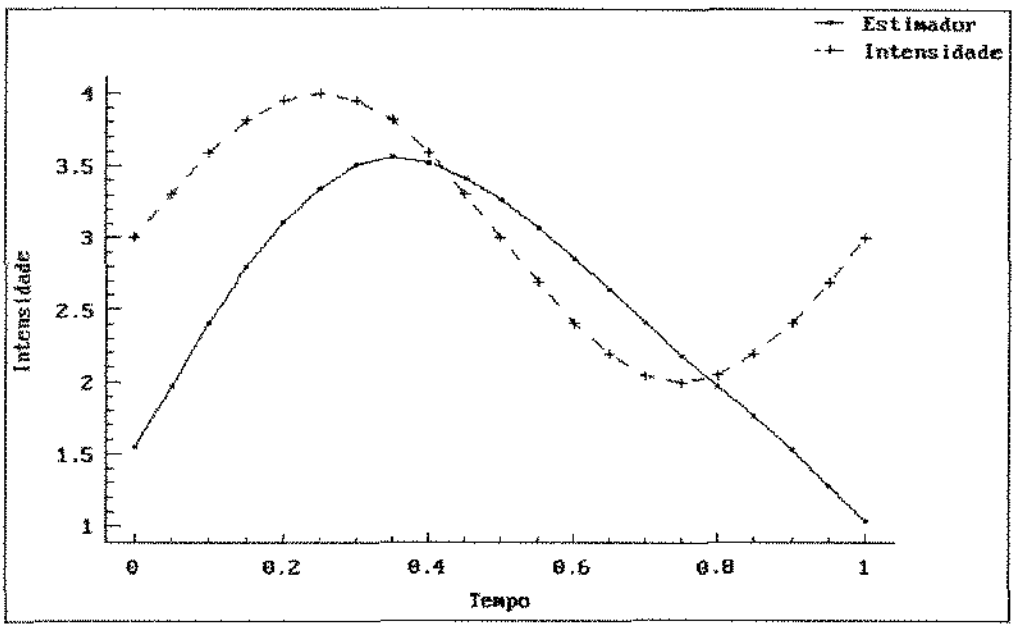


b) Estimador com correção, SDQ = 2.07.

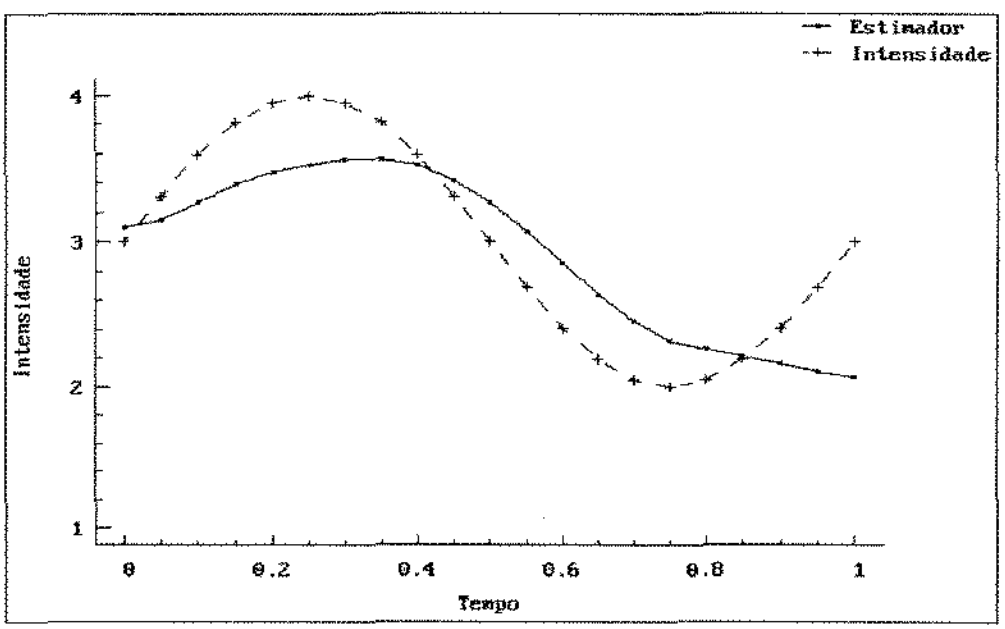
GRÁFICO 10.

$a(t) = 3 + sen(2\pi t)$

$N = 100$ *trajetórias*, $h = 0.16$



a) Estimador sem correção, SDQ = 15.34.



b) Estimador com correção, SDQ = 3.19.

Apêndice C – Programas

Programa 1

```
      PROGRAM EPAN1
C#####################################################################

C VALIDACAO CRUZADA DE M.Q. PARA O KERNEL DE EPANECHNIKOV
C SEM CORRECAO NA FRONTEIRA.
C#####################################################################

      REAL DT,ODT,LINF,LISUP,INCR,K1
      INTEGER N,NDT
      DIMENSION P(200,25),DT(500),ODT(500),EST(21),ALFA(21)
      OPEN(1,FILE="mexps.dat")
C*********************************************************************
C     GERAR TRAJETORIAS DE UM PROCESSO POISSON NAO-HOMOGENEO
C     (CASO LAMBDA INVERSIVEL)
C*********************************************************************
C     N=NUMERO DE TRAJETORIAS
C     I=NUMERO DA TRAJETORIA
10    WRITE(*,*)'NUMERO DE TRAJETORIAS ='
      READ(*,*) N
      WRITE(*,*)N
C     DATA LINF/.01/,LISUP/2./,INCR/.01/
      WRITE(*,*)'LIMITE INFERIOR DO INTERVALO DE BUSCA DA JANELA ='

      READ(*,*)LINF
      WRITE(*,15)LINF
      WRITE(*,*)'LIMITE SUPERIOR DO INTERVALO DE BUSCA DA JANELA ='

      READ(*,*)LISUP
      WRITE(*,15)LISUP
      WRITE(*,*)'INCREMENTO='
      READ(*,*)INCR
      WRITE(*,15)INCR
15    FORMAT(' ',F6.3)
      NDT=0
      DO 50 I=1,N
      S=0
      NS=1
20    U=RN1()
      Y=-LOG(1-U)
      S=S+Y
      P(I,NS)=RLAMBI(S)
      IF (P(I,NS).GE.1.) GOTO 30
      NDT=NDT+1
      DT(NDT)=P(I,NS)
      NS=NS+1
      GOTO 20
30    CONTINUE
```

```
      IF (NS.EQ.1) GOTO 45
      WRITE(*,32)
32    FORMAT('0',' TRAJET.  TPO.SALTO')
      DO 40 J=1,NS-1
      WRITE(*,35)P(I,J)
35    FORMAT(F7.5)
40    CONTINUE
      GOTO 50
45    WRITE(*,*)'NAO EXISTEM SALTOS PARA A TRAJETORIA',I
50    CONTINUE
C
      CALL EPANSCOR(DT,NDT,LINF,LISUP,INCR,SCORE,SP)
      WRITE(*,54)
      WRITE (*,55)SCORE,SP
54    FORMAT('0',63('*'))
55    FORMAT(' SCORE MINIMO =',F11.3,14X,' JANELA OTIMAL =',F8.4)

      WRITE(*,54)
      PAUSE
C  CALCULO DO ESTIMADOR SEM CORRECAO NA FRONTEIRA
       CALL SORTEP(NDT,DT,ODT)
       DO 65 V=0,1.05,.05
       SUM=0.
       J=20.*V+1
        DO 60 I=1,NDT
         A1=(V-ODT(I))/SP
         K1=SKER(A1)
         EST(J)=SUM+K1/(N*SP)
         SUM=EST(J)
60      CONTINUE
65     CONTINUE
C CALCULO DA INTENSIDADE E DA SDQ NA GRID
      SDQ=0.
      DO 70 L=1,21
      W=.05*(L-1)
      ALFA(L)=RALFA(W)
      SQI=(EST(L)-ALFA(L))**2
      SDQ=SDQ+SQI
      WRITE(*,66)W,EST(L),ALFA(L)
      WRITE(1,66)W,EST(L),ALFA(L)
66    FORMAT('0',F4.2,4X,F7.4,4X,F7.4)
70    CONTINUE
      WRITE(*,75)NDT
75    FORMAT('0','*NDT=',I3)
      WRITE(*,*)' SOMA DOS DESVIOS QUADRATICOS =',SDQ
      WRITE(*,*)'PARA OUTRA EXECUCAO DIGITE 1'
      READ(*,*)L
      IF (L.EQ.1) GOTO 10
      STOP
      END
C**************************************
C GERACAO DO NUMERO ALEATORIO EM (0,1)
```

```
C**************************************
      FUNCTION RN1()
      SAVE IX,IY,IZ
      DATA IX/11/,IY/111/,IZ/4715/
      IX=171*MOD(IX,177)-2*(IX/177)
      IY=172*MOD(IY,176)-35*(IY/176)
      IZ=170*MOD(IZ,178)-63*(IZ/178)
      IF (IX.LT.0) IX=IX+30269
      IF (IY.LT.0) IY=IY+30307
      IF (IZ.LT.0) IZ=IZ+30323

RN1=MOD(FLOAT(IX)/30269.0+FLOAT(IY)/30307.0+FLOAT(IZ)/30323.0,1.0)
      RETURN
      END

C***************************
C DEFINICAO DE ALFA(T)
C***************************
      FUNCTION RALFA(T)

C1       RALFA=2*T+1

      RALFA=12.*EXP(-2*T)
      RETURN
      END

C**************************************
C CALCULO DA INVERSA DE LAMBDA
C**************************************
      FUNCTION RLAMBI(T)

C1       RLAMBI=SQRT(T+0.25)-0.5

      IF(T .LT. 6) THEN
      RLAMBI=(-1/2.)*LOG(1-T/6.)
      ELSE
      RLAMBI=1.
      ENDIF
      RETURN
      END


C*************************************************************
C     VALIDACAO CRUZADA USANDO DIFERENCAS ABSOLUTAS ORDENADAS
C*************************************************************

      SUBROUTINE EPANSCOR(DT,NDT,LSP,USP,STEP,SCORE,SP)
C     Encontra janela que minimiza o score M0
C     A janela SP varia sobre o intervalo (LSP,USP)
C     INPUT:
C     DT = vetor de dados
C     NDT = numero de observacoes (<= 300)
```

```
C      STEP = incremento da particao de (LSP,USP)
C      OUTPUT:
C      SCORE = minimo valor do score M0
C      SP = valor da janela otimal
C
       INTEGER NDT,COUNT
       REAL H,LSP,USP,STEP,SCORE,SP,SQ5
       REAL DT,DIST,DIS2,DIS3,DIS5
       PARAMETER (N=250000)
       DIMENSION DT(500),DIST(N),DIS2(N),DIS3(N),DIS5(N)
C
C Formar vetor de diferencas X(I)-X(J)
C
       CALL MAKEDIS (NDT,DT,DIS2,COUNT)
C
C Ordenar o vetor diferenca
C
C       CALL SORTEP(NDT,DIS2,DIST)

       CALL HSORT(COUNT,DIS2)
       DO 26 I=1,COUNT
       DIST(I)=DIS2(I)
26     CONTINUE
C Efetuar soma do vetor diferenca ordenado
C
       CALL SUMDIS(COUNT,DIST,2.,DIS2)
       CALL SUMDIS(COUNT,DIST,3.,DIS3)
       CALL SUMDIS(COUNT,DIST,5.,DIS5)

C      N1 = maior valor tal que DN1 < SQ5*H
C      N2 = maior valor tal que DN2 < 2*SQ5*H
       N1=1
       N2=1
       SQ5=SQRT(5.)
       SCORE=999.
       DO 10 H=LSP,USP,STEP
11     IF((DIST(N1) .LT. SQ5*H) .AND. (N1 .LE. COUNT)) THEN
         N1=N1+1
         GOTO 11
       ENDIF
12     IF((DIST(N2) .LT. 2*SQ5*H) .AND. (N2 .LE. COUNT)) THEN
         N2=N2+1
         GOTO12
       ENDIF

SUM1=-DIS5(N2)/(100.*H**5.)+DIS3(N2)/H**3.-2.*SQ5*DIS2(N2)/H**2
     * +4.*SQ5*(2.*(N2-1.)+NDT/H)
       SUM2=N1-1.-DIS2(N1)/(5.*H**2.)
       SUM=3.*(SUM1/(100.*H)-SUM2/(SQ5*H))
       IF(SUM .LT. SCORE) THEN
         SCORE=SUM
         SP=H
```

```
      ENDIF
10    CONTINUE
      RETURN
      END

C*************************************************
      SUBROUTINE MAKEDIS(N,X,D,K)
C
C Forma vetor de diferencas absolutas X(I)-X(J)
C     INPUT:
C     N = Numero de observacoes
C     X = Vetor real (N) contendo os dados
C     OUTPUT :
C     D = Vetor de diferencas absolutas
C     K = Tamanho do vetor D (N*(N-1)/2)
      DIMENSION X(*),D(*)
      K=0
      DO 10 I=1,N
        DO 10 J=I+1,N
          K=K+1
          D(K)=ABS(X(I)-X(J))
10    CONTINUE
      RETURN
      END

C*************************************************
      SUBROUTINE SUMDIS(N,X,P,D)
C
C Efetua a soma das diferencas absolutas ordenadas
C     INPUT:
C     N = Tamanho do vetor X
C     X = Vetor real (N) contendo as diferencas absolutas ordenadas
C     P = P-esima potencia
C     OUTPUT:
C       D = Soma das diferencas absolutas ordenadas com P-esima
potencia
C     ( D(I) = X(1)**P+...+x(I-1)**P))
C
      DIMENSION X(*),D(*)

      D(1)=0.
      D(2)=X(1)**P
      DO 10 I=3,N+1
        D(I)=D(I-1)+X(I-1)**P
10    CONTINUE
      RETURN
      END

C*************************************************
      SUBROUTINE SORTEP(N,X,OX)
C Ordena uma lista de reais em ordem ascendente
C     INPUT:
```

```
C      N = Tamanho do vetor X
C      X = Vetor de reais a ordenar
C      OX = Vetor ordenado
       DIMENSION X(*),OX(*)
       INTEGER I,COUNT,J,INTVAL
       REAL TEMP
       INTVAL=N/2
       DO 20 COUNT=1,N
         IF (INTVAL .LE. 0) GOTO 25
         DO 30 I=INTVAL,N-1
           DO 40 J=I-INTVAL+1,1,-INTVAL
             IF(X(J) .GT. X(J+INTVAL)) THEN
               TEMP=X(J)
               X(J)=X(J+INTVAL)
               X(J+INTVAL)=TEMP
             ENDIF
40         CONTINUE
30       CONTINUE
       DO 50 K=1,N
        OX(K)=X(K)
50     CONTINUE
       INTVAL=INTVAL/2
20     CONTINUE
       RETURN
25     END

C****************************************
C CALCULO DO KERNEL NOS PONTOS DA GRID
C****************************************
       FUNCTION SKER(T)
       IF(ABS(T) .LT. SQRT(5.)) THEN
       SKER=(1-(ABS(T)**2.)/5.)*3./(4.*SQRT(5.))
       ELSE
       SKER=0.
       ENDIF
       RETURN
       END

C****************************************
       SUBROUTINE HSORT(N,RA)
C Ordena uma lista de reais em ordem ascendente por Heapsort

     DIMENSION RA(N)
       L=N/2+1
       IR=N
10     CONTINUE
           IF (L .GT. 1)THEN
                L=L-1
                RRA=RA(L)
           ELSE
                RRA=RA(IR)
                RA(IR)=RA(1)
```

```
                IR=IR-1
                IF (IR .EQ. 1) THEN
                    RA(1)=RRA
                    RETURN
                ENDIF
            ENDIF
            I=L
            J=L+L
20          IF(J .LE. IR)THEN
                IF(J .LT. IR) THEN
                    IF(RA(J) .LT. RA(J+1))J=J+1
                ENDIF
                IF(RRA .LT. RA(J)) THEN
                    RA(I)=RA(J)
                    I=J
                    J=J+J
                ELSE
                    J=IR+1
                ENDIF
            GOTO 20
            ENDIF
            RA(I)=RRA
        GOTO 10
        END
```

Programa 2

```
      PROGRAM EPANCR
C#####################################################################
C VALIDACAO CRUZADA DE M.Q. PARA O KERNEL DE EPANECHNIKOV
C COM CORRECAO NA FRONTEIRA.
C#####################################################################

      REAL DT,LINF,LISUP,INCR,K1,K2,K3
      INTEGER N,NDT
      DIMENSION P(200,25),DT(500),EST(21),ALFA(21)
      OPEN(1,FILE="mretc.dat")
C      OPEN(1,FILE="mexpc.dat")
C*********************************************************************
C     GERAR TRAJETORIAS DE UM PROCESSO POISSON NAO-HOMOGENEO
C     (CASO LAMBDA INVERSIVEL)
C*********************************************************************
C     N=NUMERO DE TRAJETORIAS
C     I=NUMERO DA TRAJETORIA
10    WRITE(*,*)'NUMERO DE TRAJETORIAS ='
      READ(*,*) N
      WRITE(*,*)N
C     DATA LINF/0.01/,LISUP/2./,INCR/0.01/
      WRITE(*,*)'LIMITE INFERIOR DO INTERVALO DE BUSCA DA JANELA ='

      READ(*,*)LINF
      WRITE(*,15)LINF
      WRITE(*,*)'LIMITE SUPERIOR DO INTERVALO DE BUSCA DA JANELA ='

      READ(*,*)LISUP
      WRITE(*,15)LISUP
      WRITE(*,*)'INCREMENTO='
      READ(*,*)INCR
      WRITE(*,15)INCR
15    FORMAT(' ',F6.3)
      NDT=0
      DO 50 I=1,N
      S=0
      NS=1
20    U=RN2()
      Y=-LOG(1-U)
      S=S+Y
      P(I,NS)=RLAMBI(S)
      IF (P(I,NS).GE.1.) GOTO 30
      NDT=NDT+1
      DT(NDT)=P(I,NS)
      NS=NS+1
      GOTO 20
30    CONTINUE
      IF (NS.EQ.1) GOTO 45
```

```
      WRITE(*,32)
32    FORMAT('0',' TRAJET.  TPO.SALTO')
      DO 40 J=1,NS-1
      WRITE(*,35)P(I,J)
35    FORMAT(F7.5)
40    CONTINUE
      GOTO 50
45    WRITE(*,*)'NAO EXISTEM SALTOS PARA A TRAJETORIA',I
50    CONTINUE
C
      CALL EPANSCOR(DT,NDT,LINF,LISUP,INCR,SCORE,SP)
      WRITE(*,54)
      WRITE (*,55)SCORE,SP
54    FORMAT('0',63('*'))
55    FORMAT(' SCORE MINIMO =',F12.5,14X,' JANELA OTIMAL =',F8.4)

      WRITE(*,54)
      PAUSE
C  CALCULO DO ESTIMADOR COM CORRECAO NA FRONTEIRA
      DO 65 V=0,1.05,.05
       J=20*V+1
       SUM=0.
        DO 60 I=1,NDT
         A1=(V-DT(I))/SP
         A2=(V+DT(I))/SP
         A3=(2-V-DT(I))/SP
         K1=SKER(A1)
         K2=SKER(A2)
         K3=SKER(A3)
         EST(J)=SUM+(K1+K2+K3)/(N*SP)
         SUM=EST(J)
60      CONTINUE
65    CONTINUE
C CALCULO DA INTENSIDADE E DA SDQ NA GRID
      SDQ=0.
      DO 70 L=1,21
      W=.05*(L-1)
      ALFA(L)=RALFA(W)
      SQI=(EST(L)-ALFA(L))**2
      SDQ=SDQ+SQI
      WRITE(*,66)W,EST(L),ALFA(L)
      WRITE(1,66)W,EST(L),ALFA(L)
66    FORMAT('0', F4.2,4X,F7.4,4X,F7.4)
70    CONTINUE
      WRITE(*,75)NDT
75    FORMAT('0','*NDT=',I3)
      WRITE(*,*)' SOMA DOS DESVIOS QUADRATICOS =',SDQ
      WRITE(*,*)'PARA OUTRA EXECUCAO DIGITE 1'
      READ(*,*)L
      IF (L.EQ.1) GOTO 10
      STOP
      END
```

```
C*************************************
C GERACAO DO NUMERO ALEATORIO EM (0,1)
C*************************************
      FUNCTION RN2()
      SAVE IX,IY,IZ
      DATA IX/11/,IY/111/,IZ/4715/
      IX=171*MOD(IX,177)-2*(IX/177)
      IY=172*MOD(IY,176)-35*(IY/176)
      IZ=170*MOD(IZ,178)-63*(IZ/178)
      IF (IX.LT.0) IX=IX+30269
      IF (IY.LT.0) IY=IY+30307
      IF (IZ.LT.0) IZ=IZ+30323

RN2=MOD(FLOAT(IX)/30269.0+FLOAT(IY)/30307.0+FLOAT(IZ)/30323.0,1.0)
      RETURN
      END

C***************************
C DEFINICAO DE ALFA(T)
C***************************
      FUNCTION RALFA(T)

      RALFA=2.*T+1

C1      RALFA=12*EXP(-2*T)

      RETURN
      END

C*************************************
C CALCULO DA INVERSA DE LAMBDA
C*************************************
      FUNCTION RLAMBI(T)

      RLAMBI=SQRT(T+0.25)-0.5

C1      IF (T .LT. 6) THEN
C1      RLAMBI=(-1/2.)*LOG(1-T/6.)
C1      ELSE
C1      RLAMBI=1.
C1      ENDIF

      RETURN
      END


C*************************************************************
C     VALIDACAO CRUZADA USANDO DIFERENCAS ABSOLUTAS ORDENADAS
C*************************************************************

      SUBROUTINE EPANSCOR(DT,NDT,LSP,USP,STEP,SCORE,SP)
C     Encontra janela que minimiza o score M0
```

```
C      A janela SP varia sobre o intervalo (LSP,USP)
C      INPUT:
C      DT = vetor de dados
C      NDT = numero de observacoes (<= 300)
C      STEP = incremento da particao de (LSP,USP)
C      OUTPUT:
C      SCORE = minimo valor do score M0
C      SP = valor da janela otimal
C
       INTEGER NDT,COUNT
       REAL H,LSP,USP,STEP,SCORE,SP,SQ5
       REAL DT,DIST,DIS2,DIS3,DIS5
       PARAMETER (N=250000)
       DIMENSION DT(500),DIST(N),DIS2(N),DIS3(N),DIS5(N)

C Formar vetor de diferencas absolutas X(I)-X(J)
C
       CALL MAKEDIS (NDT,DT,DIS2,COUNT)
C
C Ordenar o vetor diferenca
        CALL HSORT(COUNT,DIS2)
       DO 26 I=1,COUNT
       DIST(I)=DIS2(I)
26     CONTINUE

C Efetuar soma do vetor diferenca ordenado
C
       CALL SUMDIS(COUNT,DIST,2.,DIS2)
       CALL SUMDIS(COUNT,DIST,3.,DIS3)
       CALL SUMDIS(COUNT,DIST,5.,DIS5)

C      N1 = maior valor tal que DN1 < SQ5*H
C      N2 = maior valor tal que DN2 < 2*SQ5*H
       N1=1
       N2=1
       SQ5=SQRT(5.)
       SCORE=999.
       DO 10 H=LSP,USP,STEP
11     IF((DIST(N1) .LT. SQ5*H) .AND. (N1 .LE. COUNT)) THEN
         N1=N1+1
         GOTO 11
       ENDIF
12     IF((DIST(N2) .LT. 2*SQ5*H) .AND. (N2 .LE. COUNT)) THEN
         N2=N2+1
         GOTO12
       ENDIF

SUM1=-DIS5(N2)/(100.*H**5.)+DIS3(N2)/H**3.-2.*SQ5*DIS2(N2)/H**2
     * +4.*SQ5*(2.*(N2-1.)+NDT/H)
       SUM2=N1-1.-DIS2(N1)/(5.*H**2.)
       SUM=3.*(SUM1/(100.*H)-SUM2/(SQ5*H))
       IF(SUM .LT. SCORE) THEN
```

```
          SCORE=SUM
          SP=H
        ENDIF
10      CONTINUE
        RETURN
        END

C*************************************************
        SUBROUTINE MAKEDIS(N,X,D,K)
C
C Forma vetor de diferencas absolutas X(I)-X(J)
C       INPUT:
C       N = Numero de observacoes
C       X = Vetor real (N) contendo os dados
C       OUTPUT :
C       D = Vetor de diferencas absolutas
C       K = Tamanho do vetor D (N*(N-1)/2)
        DIMENSION X(*),D(*)
        K=0
        DO 10 I=1,N
          DO 10 J=I+1,N
            K=K+1
            D(K)=ABS(X(I)-X(J))
10      CONTINUE
        RETURN
        END

C*************************************************
        SUBROUTINE SUMDIS(N,X,P,D)
C
C Efetua a soma das diferencas absolutas ordenadas
C       INPUT:
C       N = Tamanho do vetor X
C      X = Vetor real (N) contendo as diferencas absolutas ordenadas
C       P = P-esima potencia
C       OUTPUT:
C         D = Soma das diferencas absolutas ordenadas com P-esima
potencia
C       ( D(I) = X(1)**P+...+x(I-1)**P))
C
        DIMENSION X(*),D(*)

        D(1)=0.
        D(2)=X(1)**P
        DO 10 I=3,N+1
          D(I)=D(I-1)+X(I-1)**P
10      CONTINUE
        RETURN
        END

C*************************************************
        SUBROUTINE SORTEP(N,X,OX)
```

```
C Ordena uma lista de reais em ordem ascendente
C      INPUT:
C      N = Tamanho do vetor X
C      X = Vetor de reais a ordenar
C      OX = Vetor ordenado
       DIMENSION X(*),OX(*)
       INTEGER I,LCOUNT,J,INTVAL
       REAL TEMP
       INTVAL=N/2
       DO 20 LCOUNT=1,N
         IF (INTVAL .LE. 0) GOTO 25
         DO 30 I=INTVAL,N-1
           DO 40 J=I-INTVAL+1,1,-INTVAL
             IF(X(J) .GT. X(J+INTVAL)) THEN
               TEMP=X(J)
               X(J)=X(J+INTVAL)
               X(J+INTVAL)=TEMP
             ENDIF
40         CONTINUE
30       CONTINUE
       DO 50 K=1,N
        OX(K)=X(K)
50     CONTINUE
       INTVAL=INTVAL/2
20     CONTINUE
       RETURN
25     END

C**************************************
C CALCULO DO KERNEL NOS PONTOS DA GRID
C**************************************
       FUNCTION SKER(T)
       IF(ABS(T) .LT. SQRT(5.)) THEN
       SKER=(1-(ABS(T)**2.)/5.)*3./(4.*SQRT(5.))
       ELSE
       SKER=0.
       ENDIF
       RETURN
       END


C*****************************************
       SUBROUTINE HSORT(N,RA)
C Ordena reais por Heapsort

      DIMENSION RA(N)
       L=N/2+1
       IR=N
10     CONTINUE
           IF (L .GT. 1)THEN
                L=L-1
                RRA=RA(L)
```

```
            ELSE
                 RRA=RA(IR)
                 RA(IR)=RA(1)
                 IR=IR-1
                 IF (IR .EQ. 1) THEN
                     RA(1)=RRA
                     RETURN
                 ENDIF
             ENDIF
             I=L
             J=L+L
20           IF(J .LE. IR)THEN
                IF(J .LT. IR) THEN
                     IF(RA(J) .LT. RA(J+1))J=J+1
                ENDIF
                IF(RRA .LT. RA(J)) THEN
                    RA(I)=RA(J)
                    I=J
                    J=J+J
                ELSE
                    J=IR+1
                ENDIF
             GOTO 20
             ENDIF
             RA(I)=RRA
         GOTO 10
         END
```

Programa 3

```
      PROGRAM EPANIC
C#####################################################################
C ESTIMACAO POR KERNEL,CASO INTENSIDADE NAO-INVERSIVEL COM CORRECAO

C NA FRONTEIRA.
C (VALIDACAO CRUZADA DE M.Q. PARA O KERNEL DE EPANECHNIKOV)
C
C#####################################################################
      REAL DT1,ODT1,LINF,LISUP,INCR,K1,K2,K3,LAMBES,U2
      INTEGER N,NDT
      DIMENSION P1(200,25),DT1(500),ODT1(500),EST1(21),ALFA(21)
C*********************************************************************
C     GERAR TRAJETORIAS DE UM PROCESSO POISSON NAO-HOMOGENEO
C     (CASO LAMBDA INVERSIVEL)
C*********************************************************************
C     N=NUMERO DE TRAJETORIAS
C     I=NUMERO DA TRAJETORIA
10    WRITE(*,*)'NUMERO DE TRAJETORIAS ='
      READ(*,*) N
      WRITE(*,*)N
      OPEN(1,FILE="sen.dat")
C     DATA LINF/0.01/,LISUP/2./,INCR/0.01/
      WRITE(*,*)'LIMITE INFERIOR DO INTERVALO DE BUSCA DA JANELA ='

      READ(*,*)LINF
      WRITE(*,15)LINF
      WRITE(*,*)'LIMITE SUPERIOR DO INTERVALO DE BUSCA DA JANELA ='

      READ(*,*)LISUP
      WRITE(*,15)LISUP
      WRITE(*,*)'INCREMENTO='
      READ(*,*)INCR
      WRITE(*,15)INCR
15    FORMAT(' ',F6.3)
      NDT=0
      DO 50 I=1,N
      S=0
      NS=1
20    U=RN1()
      Y=-LOG(1-U)
      S=S+Y
      P1(I,NS)=RLAMB1(S)
      IF (P1(I,NS).GE.1.) GOTO 30
      NDT=NDT+1
      DT1(NDT)=P1(I,NS)
      NS=NS+1
      GOTO 20
30    CONTINUE
```

```
      IF (NS.EQ.1) GOTO 45
      GOTO 50
45    WRITE(*,*)'NAO EXISTEM SALTOS PARA A TRAJETORIA',I
50    CONTINUE
      CALL SORTEP(NDT,DT1,ODT1)
      WRITE(*,*)' VALORES ORDENADOS:'
      DO 40 J=1,NDT
      WRITE(*,35)ODT1(J)
35    FORMAT('0',F7.5)
40    CONTINUE

C GERACAO DOS PONTOS DO P.P.N.H. COM LAMBDA NAO-INVERSIVEL
      LAMBES=4.
C LAMBES=MAXIMO DE RLAMBD(T)
      KONT=0
      DO 100 J=1,NDT
      U2=RN1()
      IF (U2 .GT. RLAMBD(ODT1(J))/LAMBES) GOTO 100
      KONT=KONT+1
      ODT1(KONT)=ODT1(J)
100   CONTINUE
      WRITE(*,*)
      WRITE(*,*)' PONTOS NAO DELETADOS:'
      DO 105 J=1,KONT
      WRITE(*,35) ODT1(J)
105   CONTINUE
C
      CALL EPANSCOR(ODT1,KONT,LINF,LISUP,INCR,SCORE,SP)
      WRITE(*,54)
      WRITE (*,55)SCORE,SP
54    FORMAT('0',63('*'))
55    FORMAT(' SCORE MINIMO =',F11.3,14X,' JANELA OTIMAL =',F8.4)

      WRITE(*,54)
      PAUSE
C  CALCULO DO ESTIMADOR COM CORRECAO NA FRONTEIRA
      DO 65 V=0,1.05,.05
       SUM=0.
       J=20*V+1
        DO 60 I=1,KONT
         A1=(V-ODT1(I))/SP
         A2=(V+ODT1(I))/SP
         A3=(2-V-ODT1(I))/SP
         K1=SKER(A1)
         K2=SKER(A2)
         K3=SKER(A3)
         EST1(J)=SUM+(K1+K2+K3)/(N*SP)
         SUM=EST1(J)
60       CONTINUE
65    CONTINUE
C      CALCULO DA INTENSIDADE E DA SDQ NA GRID
      SDQ=0.
```

```
      DO 70 L=1,21
      W=.05*(L-1)
      ALFA(L)=RLAMBD(W)
      SQI=ABS(EST1(L)-ALFA(L))**2
      SDQ=SDQ+SQI
      WRITE(*,66)W,EST1(L),ALFA(L)
      WRITE(1,66)W,EST1(L),ALFA(L)
66    FORMAT('0',F4.2,4X,F7.4,4X,F7.4)
70    CONTINUE
      WRITE(*,75)NDT,KONT
75    FORMAT('0','*NDT=',I3,3X,'NDEL=',I3)
      WRITE(*,*)' SOMA DOS DESVIOS QUADRATICOS =',SDQ
      WRITE(*,*)'PARA OUTRA EXECUCAO DIGITE 1'
      READ(*,*)L
      IF (L.EQ.1) GOTO 10
      STOP
      END
C**************************************
C GERACAO DO NUMERO ALEATORIO EM (0,1)
C**************************************
      FUNCTION RN1()
      SAVE IX,IY,IZ
      DATA IX/11/,IY/111/,IZ/4715/
      IX=171*MOD(IX,177)-2*(IX/177)
      IY=172*MOD(IY,176)-35*(IY/176)
      IZ=170*MOD(IZ,178)-63*(IZ/178)
      IF (IX.LT.0) IX=IX+30269
      IF (IY.LT.0) IY=IY+30307
      IF (IZ.LT.0) IZ=IZ+30323

RN1=MOD(FLOAT(IX)/30269.0+FLOAT(IY)/30307.0+FLOAT(IZ)/30323.0,1.0)
      RETURN
      END

C****************************
C DEFINICAO DE LAMBDA(T)
C****************************
      FUNCTION RLAMBD(T)

      RLAMBD=3+SIN(2*3.1415926*T)
      RETURN
      END

C**************************************
C CALCULO DA INVERSA DE LAMBDA
C**************************************
      FUNCTION RLAMB1(T)

      RLAMB1=T/4.
      RETURN
      END
```

```
C*************************************************************
C     VALIDACAO CRUZADA USANDO DIFERENCAS ABSOLUTAS ORDENADAS
C*************************************************************

      SUBROUTINE EPANSCOR(DT,NDT,LSP,USP,STEP,SCORE,SP)
C     Encontra janela que minimiza o score M0
C     A janela SP varia sobre o intervalo (LSP,USP)
C     INPUT:
C     DT = vetor de dados
C     NDT = numero de observacoes (<= 300)
C     STEP = incremento da particao de (LSP,USP)
C     OUTPUT:
C     SCORE = minimo valor do score M0
C     SP = valor da janela otimal
C
      INTEGER NDT,COUNT
      REAL H,LSP,USP,STEP,SCORE,SP,SQ5
      REAL DT,DIST,DIS2,DIS3,DIS5
      PARAMETER (N=250000)
      DIMENSION DT(500),DIST(N),DIS2(N),DIS3(N),DIS5(N)
C
C Formar vetor de diferencas X(I)-X(J)
C
      CALL MAKEDIS (NDT,DT,DIS2,COUNT)
C
C Ordenar o vetor diferenca
C
       CALL HSORT(COUNT,DIS2)
      DO 26 I=1,COUNT
      DIST(I)=DIS2(I)
26    CONTINUE
C
C Efetuar soma do vetor diferenca ordenado
C
      CALL SUMDIS(COUNT,DIST,2.,DIS2)
      CALL SUMDIS(COUNT,DIST,3.,DIS3)
      CALL SUMDIS(COUNT,DIST,5.,DIS5)

C     N1 = maior valor tal que DN1 < SQ5*H
C     N2 = maior valor tal que DN2 < 2*SQ5*H
      N1=1
      N2=1
      SQ5=SQRT(5.)
      SCORE=999.
      DO 10 H=LSP,USP,STEP
11    IF((DIST(N1) .LT. SQ5*H) .AND. (N1 .LE. COUNT)) THEN
        N1=N1+1
        GOTO 11
      ENDIF
12    IF((DIST(N2) .LT. 2*SQ5*H) .AND. (N2 .LE. COUNT)) THEN
        N2=N2+1
        GOTO12
```

```
      ENDIF

SUM1=-DIS5(N2)/(100.*H**5.)+DIS3(N2)/H**3.-2.*SQ5*DIS2(N2)/H**2
     * +4.*SQ5*(2.*(N2-1.)+NDT/H)
      SUM2=N1-1.-DIS2(N1)/(5.*H**2.)
      SUM=3.*(SUM1/(100.*H)-SUM2/(SQ5*H))
      IF(SUM .LT. SCORE) THEN
        SCORE=SUM
        SP=H
      ENDIF
10    CONTINUE
      RETURN
      END

C*************************************************
      SUBROUTINE MAKEDIS(N,X,D,K)
C
C Forma vetor de diferencas absolutas X(I)-X(J)
C     INPUT:
C     N = Numero de observacoes
C     X = Vetor real (N) contendo os dados
C     OUTPUT :
C     D = Vetor de diferencas absolutas
C     K = Tamanho do vetor D (N*(N-1)/2)
      DIMENSION X(*),D(*)
      K=0
      DO 10 I=1,N
        DO 10 J=I+1,N
          K=K+1
          D(K)=ABS(X(I)-X(J))
10    CONTINUE
      RETURN
      END

C*************************************************
      SUBROUTINE SUMDIS(N,X,P,D)
C
C Efetua a soma das diferencas absolutas ordenadas
C     INPUT:
C     N = Tamanho do vetor X
C     X = Vetor real (N) contendo as diferencas absolutas ordenadas
C     P = P-esima potencia
C     OUTPUT:
C       D = Soma das diferencas absolutas ordenadas com P-esima
potencia
C     ( D(I) = X(1)**P+...+x(I-1)**P))
C
      DIMENSION X(*),D(*)

      D(1)=0.
      D(2)=X(1)**P
      DO 10 I=3,N+1
```

```
        D(I)=D(I-1)+X(I-1)**P
10     CONTINUE
       RETURN
       END

C*********************************************
       SUBROUTINE SORTEP(N,X,OX)
C Ordena uma lista de reais em ordem ascendente
C      INPUT:
C      N = Tamanho do vetor X
C      X = Vetor de reais a ordenar
C      OX = Vetor ordenado
       DIMENSION X(*),OX(*)
       INTEGER I,COUNT,J,INTVAL
       REAL TEMP
       INTVAL=N/2
       DO 20 COUNT=1,N
         IF (INTVAL .LE. 0) GOTO 25
         DO 30 I=INTVAL,N-1
           DO 40 J=I-INTVAL+1,1,-INTVAL
             IF(X(J) .GT. X(J+INTVAL)) THEN
               TEMP=X(J)
               X(J)=X(J+INTVAL)
               X(J+INTVAL)=TEMP
             ENDIF
40         CONTINUE
30       CONTINUE
       DO 50 K=1,N
        OX(K)=X(K)
50     CONTINUE
       INTVAL=INTVAL/2
20     CONTINUE
       RETURN
25     END

C*************************************
C CALCULO DO KERNEL NOS PONTOS DA GRID
C*************************************
       FUNCTION SKER(T)
       IF(ABS(T) .LT. SQRT(5.)) THEN
       SKER=(1-(ABS(T)**2.)/5.)*3./(4.*SQRT(5.))
       ELSE
       SKER=0.
       ENDIF
       RETURN
       END


C*************************************
       SUBROUTINE HSORT(N,RA)
C Ordena reais por Heapsort
```

```
      DIMENSION RA(N)
       L=N/2+1
       IR=N
10     CONTINUE
            IF (L .GT. 1)THEN
                 L=L-1
                 RRA=RA(L)
            ELSE
                 RRA=RA(IR)
                 RA(IR)=RA(1)
                 IR=IR-1
                 IF (IR .EQ. 1) THEN
                     RA(1)=RRA
                     RETURN
                 ENDIF
            ENDIF
            I=L
            J=L+L
20          IF(J .LE. IR)THEN
               IF(J .LT. IR) THEN
                    IF(RA(J) .LT. RA(J+1))J=J+1
               ENDIF
               IF(RRA .LT. RA(J)) THEN
                   RA(I)=RA(J)
                   I=J
                   J=J+J
               ELSE
                   J=IR+1
               ENDIF
            GOTO 20
            ENDIF
            RA(I)=RRA
       GOTO 10
       END
```

Bibliografia.

Aalen, O.O. (1975). *Statistical Inference for a Family of Counting Processes.* Ph. D. dissertation, The University of California, Berkeley.

Aalen, O.O. (1978). Nonparametric inference for a family of counting processes. *The Annals of Statistics*, 6, 701-726.

Andersem, P.K. and Borgan, O. (1985). Counting process models for life history data: a review. *Scandinavian Journal of Statistics*, 12, 97-158.

Anderson, G.D. (1969). *A Comparison of Methods for Estimating a Probability Density Function.* Ph. D. dissertation, University of Washington.

Bean, S.J. and Tsokos, C.P. (1980). Developments in non-parametric density estimation. *International Statistical Review*, 48, 267-287.

Bowman, Adrian W. (1984). An alternative method of cross-validation for the smoothing of density estimates. *Biometrika*, 71, 363-370.

Brémaud, P. (1981). *Point Processes and Queues: Martingale Dynamics.* Springer-Verlag, New York.

Brooks, M.M. (1991). *Bandwidth Selection Methods for Kernel Estimators of the Intensity Funtions of a Nonhomogeneous Poisson Process.* Ph. D. dissertation, The University of North Carolina at Chapel Hill.

Burkholder, D.L. (1973). Distribution function inequalities for martingales. *The Annals of Probability*, 1, 19-24.

Crestana, C. (1991). *Um Método para a Estimação da Função Intensidade do Processo de Poisson não homogêneo.* Tese de Mestrado, Unicamp-Imecc.

Daley, D.J. and Vere-Jones, D. (1988). *An Introduction to the Theory of Point Processes.* Springer-Verlag, Berlin.

Diggle, P. (1985). A kernel method for smoothing point process data. *Applied Statistics*, 34, 138-147.

Diggle, P. and Marron, J.S. (1988). Equivalence of smoothing parameter selectors in density and intensity estimation. *Journal of the American Statistical Association*, 83, 793-800.

Duin, R.P.W. (1976). On the choice of smoothing parameters for Parzen estimators of probability density functions. *IEEE Transactions on Computers*, C-25, 1175-1179.

Elliot, R.J. (1982). *Stochastic Calculus and Applications*. Springer-Verlag, New York.

Epanechnikov, V.A. (1969). Nonparametric estimators of a multivariate probability density. *Theory of Probability and Its Applications*, 14, 153-158.

Gill, R.D. (1980). Censoring and stochastic integrals. *Mathematical Centre Tracts*, 124, Mathematical Centre, Amsterdam.

Gomes, A.E. (1991). *Estimação da Função Intensidade de Processos de Poisson Não Homogêneos pelo Método de Granander*. Tese de Mestrado, Unicamp-IMECC.

Hall, P. (1983). Large sample optimality of least squares cross-validation in Density Estimation. *The Annals of Statistics*, 11, 1156-1174.

Härdle, W., Marron, J. S. and Wand, M. P. (1990). Bandwidth choice for density derivates. *Journal of the Royal Statistical Society*, Series B, 52, 223-232.

Jacobsen, M. (1982). Statistical analysis of counting processes. *Lecture Notes in Statistics*, 12, Springer-Verlag.

Jacod J. and Shiryayev, A.N. (1986). *Limit Theorems for Stochastic Processes*. Springer-Verlag, New York.

Karr, A.F. (1986). *Point Processes and their Statistical Inference*. Marcel Dekker, New York.

Lee, B.G. and Kim, B.C. (1991). An efficient algorithm for the least-squares cross-validation with symmetric and polynomial kernels. *Communications in Statistics-Simulations*, 19, 1523-1526.

Lewis, P.A.W. and Shedler, G.S. (1977). Simulation of nonhomogeneous Poisson processes by thinning. *Naval Researchs Logistics Quart*, 28, 403-413.

Liptser, R.S. and Shiryayev, A.N. (1977). *Statistics of Random Processes*, vol. I. Springer-Verlag, Berlin.

Liptser, R.S. and Shiryayev, A.N. (1978). *Statistics of Random Processes*, vol. II. Springer-Verlag, Berlin.

Liptser, R.S. and Shiryayev, A.N. (1980). A functional central limit theorem for semimartingales. *Theory of Probability and Its Applications*, 25, 667-688.

Marron, J.S. (1985). An asymptotically efficient solution to the bandwidth problem of kernel density estimation. *The Annals of Statistics*, 13, 1011-1023.

Marron, J.S. (1987). A comparison of croos-validation techniques in density estimation. *The Annals of Statistics*, 15, 152-162.

Marron, J.S. and Brooks, M.M. (1991). Asymptotic optimality of the least-squares cross-validation bandwidth for kernel estimates of intensity functions. *Stochastic Processes and their Applications*, 38, 157-165.

Meyer, P.-A. (1976). Un cours sur les intégrales stochastiques. *Lecture Notes in Mathematics*, 511, 140-165, Springer-Verlag.

Nelson, W. (1969). Hazard plotting for incomplete failure data. *Journal of Quality Technology*, 1, 27-52.

Nelson, W. (1972). Theory and application of hazard plotting for censored failure data. *Technometrics*, 14, 945-966.

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

Ramlau-Hansen, H. (1983 a). Smoothing counting processes intensities by means of kernel functions. *The Annals of Statistics*, 11, 453-466.

Ramlau-Hansen, H. (1983 b). The choice of a kernel function in the graduation of counting process intensities. *Scandinavian Actuarial Journal*, 165-182.

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

Rosenblatt, M. (1971). Curve estimates. *The Annals of Mathematical Statistics*, 42, 1815-1842.

Rudemo, M. (1982). Empirical choice of histograms and kernel density estimators. *Scandinavian Journal of Statistics*, 9, 65-78.

Schuster, E. (1985). Incorporating support constraints into nonparametric estimators of densities. *Communications in Statistics*, Series A, 14, 1123-1136.

Silverman, B.W. (1978). Choosing the window width when estimating a density. *Biometrika*, 65, 1-11.

Silverman, B.W. (1982). Kernel density estimation using the fast Fourier transform. Statistical Algorithm AS 176. *Applied Statistics*, 31, 93-97.

Silverman, B.W. (1986). *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, New York.

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

Wichmann, B.A. and Hill, I.D. (1982). An efficient and portable pseudo-random number generator. Statistical Algorithm AS 183. *Applied Statistics*, 31, 188-190.

Woodroofe, M. (1970). On chosing a delta sequence. *The Annals of Mathematical Statistics*, 41, 1665-1671.