



Universidade de Brasília
IE - Instituto de Exatas
Departamento de Estatística

Processos de Hawkes

Uma Modelagem dos Books de Ofertas no Mercado Acionário
Brasileiro

Yuri Sampaio Maluf

Brasília
2013

Yuri Sampaio Maluf

Processos de Hawkes

Uma Modelagem dos Books de Ofertas no Mercado Acionário
Brasileiro

Monografia apresentada ao Departamento
de Estatística da Universidade de Brasília,
como requisito parcial para a obtenção do
título de Bacharel em Estatística.

Orientadora: Dra. Cira Etheowalda Gue-
vara Otiniano.

Brasília
2013

Maluf, Yuri S.

Processos de Hawkes: Uma Modelagem dos Books de Ofertas no Mercado Acionário Brasileiro

74 páginas

Monografia - Instituto de Exatas da Universidade de Brasília. Departamento de Estatística.

1. Processo Pontuais
2. Processo de Hawkes
3. Book Ofertas

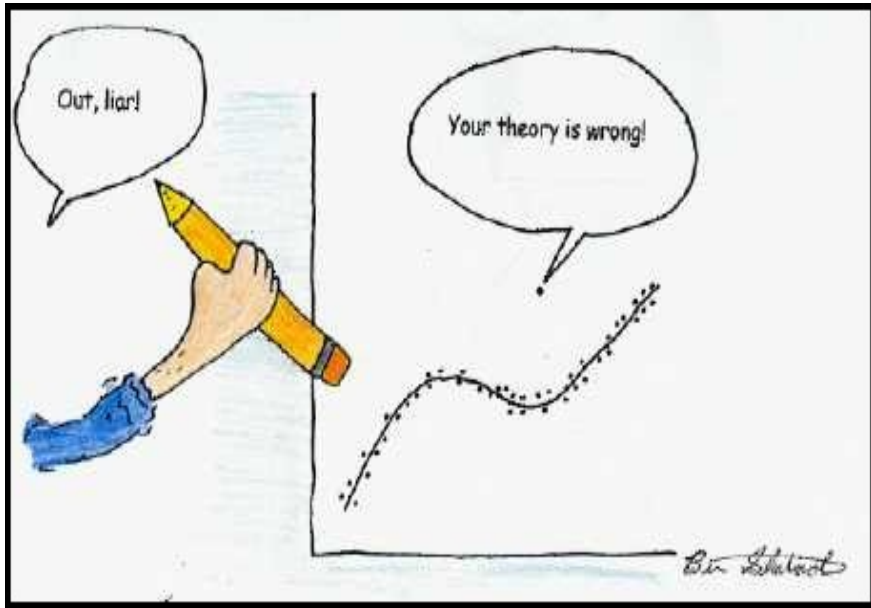
I. Universidade de Brasília. Instituto de Exatas. Departamento de Estatística.

Comissão Julgadora:

Prof. Dr. Jhames Matos Sampaio

Prof. Dr. Lucas Moreira

Prof. Dra. Cira Etheowalda Guevara Otiniano.
Orientador



Agradecimentos

Agradeço a minha professora e orientadora, Cira Etheowalda, pela atenção e pela imensa ajuda e conhecimentos transmitidos, sem os quais não seria possível a construção deste trabalho. A toda equipe de professores da estatística, além das secretárias do departamento, Kaliane e Tathyanna. Adicionalmente aos meus amigos, à toda minha família, Ana Helena, minha mentora na Terra, Fares, Ana Júlia, Ana Cristhina, Berti, Andréa, Bandit, Sivuca, e ao meu grande amor, Denise. E finalmente ao meu mentor espiritual.

Resumo

Neste trabalho abordamos o processo de Hawkes na modelagem do Book de ofertas do fundo de índice ETF *iShare Ibovespa*. O estudo tem como objetivo investigar a dinâmica das influências das ofertas em relação às ordens passadas, além da interação das taxas de ordens dos lados opostos do Book. Outro ponto abordado é a verificação da estratégia de operação que capte a dinâmica estudada. No primeiro momento, foi realizada a abordagem univariada e no segundo a multivariada do processo de Hawkes. Os resultados mostram que em ambos os casos os dados se ajustaram bem ao processo de Hawkes. Os parâmetros estimados indicam que os agentes possuem comportamentos similares quando atuam como compradores ou como vendedores de ativos. Quanto a estratégia, não foi possível estabelecer ganhos sobressalentes a partir das mudanças nas taxas de ofertas. No entanto, estes resultados se devem mais ao *bid-ask spread* do que a capacidade preditiva das funções intensidades.

Palavras-chave: Processos Pontuais, Processo de Hawkes, Book de Ofertas.

Lista de Figuras

1.1	Exemplo: Book de Ofertas	4
3.1	Processo Pontual: Eventos, processo de contagem e as <i>durations</i>	22
4.1	Diagrama dos tempos de ocorrências t_i dos eventos na reta	44
5.1	Série preços do BOVA11	53
5.2	QQ-Plot <i>Durations</i> das Melhores Ofertas de Compra e Vendas do ETF .	54
5.3	Série das <i>Durations</i> das Ofertas de Compra e Vendas do ETF	54
5.4	Funções de Autocorrelações da Série das <i>Durations</i> das Ofertas de Com- pra e Vendas do ETF	55
5.5	Gráfico QQ-Plot da Série das <i>Durations</i> das Ofertas de Compra do ETF	56
5.6	Gráfico QQ-Plot da Série das <i>Durations</i> das Ofertas de Venda do ETF .	57
5.7	Gráfico da interpolação $\phi(t)$ - ETF BOVA11	57
5.8	Intensidade $\lambda(t)$ de Compra (Azul) e Venda (Lilás) do ETF BOVA11 . .	58
5.9	Intensidade $\lambda(t)$ de Compra (Azul) e Venda (Lilás) do ETF BOVA11 . .	60
5.10	Gráfico QQ-Plot da Série das <i>Durations</i> das Ofertas de Compra do ETF	61
5.11	Gráfico QQ-Plot da Série das <i>Durations</i> das Ofertas de Venda do ETF .	61
5.12	Gráfico do Nível de Assertividade	64

Lista de Tabelas

5.1	Estatísticas Descritivas das Durations - ETF BOVA11	53
5.2	Estimativa Função Intensidade - ETF BOVA11	56
5.3	Intensidade Média por Minuto - ETF BOVA11	57
5.4	Estimativa Função Intensidade com Ajuste Sazonal- ETF BOVA11 . . .	58
5.5	Estimativa Função Intensidade - ETF BOVA11	59
5.6	Estimativa Função Intensidade - ETF BOVA11	62
5.7	Resultado da Assertividade Máxima	64

Sumário

1	Introdução	1
1.1	Book de Ofertas	3
1.2	Objetivos	4
1.2.1	Objetivos Específicos	5
2	Conceitos Básicos	7
2.1	Modelos Determinísticos e Estocásticos	7
2.2	Probabilidade e Processos Estocásticos	9
2.3	Esperança Condicional	16
3	Processos de Hawkes	19
3.1	Processos Pontuais	19
3.2	Processos de Poisson	23
3.3	Função Intensidade	26
3.4	Processo de Hawkes	32
3.4.1	Processo de Hawkes Univariado	32
3.4.2	Processo de Hawkes Multivariado	35
3.5	Estratégia de Operação	39
4	Estimação do Processo de Hawkes	43
4.1	Estimação Univariada	43

4.2	Estimação Multivariada	47
5	Resultados	51
5.1	Descrição dos Dados	51
5.2	Resultados: Preliminares	53
5.3	Resultados: Processo Hawkes Univariado	55
5.4	Resultados: Processo Hawkes Multivariado	59
5.5	Simulação: Estratégia de Operação	63
6	Considerações Finais	67
	Referências Bibliográficas	69
A	Primeiro apêndice	72
B	Segundo apêndice	74

Capítulo 1

Introdução

No sistema capitalista, majoritariamente vigente no mundo, os agentes econômicos transacionam bens livremente. A intermediação de alguns destes ativos são viabilizados nos mercados financeiros. Incorporado ao mercado financeiro o chamado mercado de capitais detêm papel importante no que tange o crescimento econômico dos países. Ele representa o elo entre os compradores e vendedores de títulos representativos do capital das empresas. Os valores dos ativos são regidos pelo equilíbrio entre a demanda e oferta do ativo neste mercado ([Assaf Neto, 2009](#)).

Nos estudos hegemônicos sobre o mercado de capitais, surge a chamada moderna teoria de finanças calcada na hipótese dos mercados eficientes. Esta hipótese, alega que os agentes estão diante de uma concorrência perfeita e têm livre acesso aos mercados e as informações. Desta maneira, os preços dos ativos reagem plenamente as informações na medida em tornam-se disponíveis publicamente convergindo para seu valor fundamental. Sem o acréscimo de novas informações os preços flutuam aleatoriamente em torno do valor fundamental ocorrendo o deslocamento apenas com a chegada de novos fatos relevantes. Este comportamento vai ao encontro do modelo *Random Walk*, dos movimentos dos ativos, o que torna os preços imprevisíveis e independentes de eventos pretéritos. Em virtude desta hipótese, há a eliminação da possibilidade de arbitragem

ao encontrar ativos super ou subvalorizados sistematicamente.

Um dos principais proponentes desta hipótese [Fama \(1970\)](#) divide a eficiência dos mercados em três níveis: no tipo fraca, os preços dos ativos já compreendem todos os valores precedentes. No segundo nível, semi-forte, os preços dos ativos contêm tanto as informações dos valores históricos, bem como descontam todas as informações públicas, como demonstrativos financeiros das empresas e notícias veiculadas na mídia. Por último, na eficiência forte, implica na inserção de informações tanto públicas quanto privadas na composição dos preços dos ativos.

As ofertas de compra e venda de ativos com seus respectivos valores são registradas no *Book* de ofertas e ficam disponíveis em sistemas eletrônicos de negociação. As transações são executadas sempre que os valores de compra e venda se equivalerem. As ordens podem ser divididas basicamente em duas: As chamadas *limit orders*, o agente dispõe a quantidade e o valor de interesse da negociação o qual fica registrado no *book* de oferta aguardando outro agente inserir o mesmo valor na posição contrária. Em contra partida, a ordem a mercado (*market orders*) é executada no melhor preço disponível no momento ([Jondeau et al., 2007](#)).

A diferença entre o melhor preço de compra e o de venda é denominado *bid-ask spread*. Ele reflete o custo de negociação imediata, ou seja, o valor no qual o agente deve ceder para executar uma ordem a mercado ([Schwartz, 2010](#)). O *bid-ask spread* também indica o grau de liquidez do ativo no mercado ([Gro-Klumann e Hautsch, 2011](#)).

Dentro deste contexto, um questionamento plausível é se existe alguma medida direta para capturar o fluxo de informações nos mercados. Alguns trabalhos como [Clark \(1973\)](#), [Lamoureux e Lastrapes \(1990\)](#) e [Jones et al. \(1994\)](#) propuseram utilizar o volume de transações. Entretanto, a ideia de usar o volume como uma medida *proxy* para captar as chegadas de informações não se verificou empiricamente. Uma variante é utilizar uma medida do nível de atividade de negociações. Para um dado um intervalo de tempo, quanto maior a chegada de informações maior será o número de transações

e por consequência, menor será o intervalo (*duration*) entre elas. Diante disto, alguns estudos surgiram para examinar a dinâmica das *durations* como Engle e Russel (1998) conhecido como modelo ACD, Darolles et al. (2001) conhecido como modelo composto autoregressivo (CAR) e Gourioux e Jasiak (2003) com modelos gama autoregressivos. Uma alternativa a estes modelos é o emprego de processos pontuais que incorporem as características dos intervalos entre as transações de ativos.

De maneira geral, podemos entender um processo estocástico pontual como um processo utilizado para modelagem de ocorrências de um determinado evento em um ponto específico no tempo (Daley e Vere-Jones, 2002). Assim, é possível descrever uma realização deste processo como uma sequência de pontos, ou intervalos deles (*durations*) sobre a reta.

Um caso especial de processos pontuais é o chamado processo de auto excitação de Hawkes. Este processo foi inicialmente descrito por Hawkes (1971) para modelar as chegadas de abalos sísmicos. Nele, a taxa de chegada dos eventos é modelada inserindo uma componente de dependência temporal entre os tempos das ocorrências. Em finanças, o processo de Hawkes é de grande interesse, pois dados empíricos mostram que em muitos mercados os intervalos entre as chegadas apresentam aglomerações em determinados períodos seguidos por outros com maiores espaçamentos (Hewlett, 2006).

Em suma, estudar os fluxos de ordens permite jogar luz sobre um maior entendimento de como ocorre a relação entre as negociações ao longo do tempo. Esta importância se dá sobre tudo porque não há estudos que abordam este tipo de processo no mercado acionário brasileiro.

1.1 Book de Ofertas

No mercado acionário as ações são negociadas nos moldes de um leilão. Nele, os agentes fazem as ofertas de quanto tem de interesse em negociar e por sua vez ficam registrado

no chamado Book de Ofertas. O Book de ofertas, portanto, é uma lista ordenada pelos valores de oferta de cada agente com suas respectivas quantidades de ações desejadas. A lista é dividida em duas partes onde se situam as ordens de compra e as de venda. Um exemplo de Book pode ser visto na figura abaixo.

Corretora	Qtd	Ofc	Ofv	Qtd	Corretora
102 Banif Primus	200	29,13	29,15	300	Banif Primus 102
039 Agora	500	29,12	29,15	200	Prosper 150
003 XP Investimentos	800	29,11	29,15	400	Prosper 150
058 Socopa	100	29,11	29,18	500	Codepe 234
174 Itaú	100	29,11	29,18	2.5K	Agora 039
077 Intra	20.5K	29,10	29,19	100	Bradesco 072
077 Intra	1K	29,10	29,19	1.1K	Codepe 234
077 Intra	10K	29,10	29,20	200	Gradual 227
082 Tov	2K	29,10	29,20	1.4K	Bradesco 072
070 HSBC	400	29,10	29,20	100	Itaú 114

Figura 1.1: Exemplo: Book de Ofertas

Este quadro fica disponível aos investidores via internet por meio do *home broker* no momento em que vão inserir suas ordens. Neste exemplo, o lado azul (esquerdo) representa as ofertas de compra e o lado verde (direito) as de venda. A primeira coluna para cada tipo de oferta retrata a corretora de origem. A segunda coluna mostra a quantidade de lote de ações ofertadas, cada uma contendo 100 ações. A última coluna indica o valor da oferta para cada lote de ações. As transações somente são liquidadas quando dois agentes em posições contrárias inserem ordens com o mesmo valor de oferta.

1.2 Objetivos

O objetivo geral da monografia é verificar se o comportamento das séries dos tempos entre chegadas das ofertas de compra e venda das cotas do fundo ETF *iShare Ibovespa* podem ser bem descritas por um processo pontual Hawkes.

1.2.1 Objetivos Específicos

A fim de modelar a série de chegadas de ofertas do *Book* por meio dos processos de Hawkes é necessário, portanto, subdividir em outros pontos.

- Investigar o grau de influência das chegadas de ordens passadas na taxa de chegadas de ofertas.
- Verificar o grau de interação das ordens de uma posição com relação a posição contrária.
- Testar a eficiência de estratégias que incorporem a dinâmica das relações das taxas de ofertas no *book* ao longo do tempo.

Capítulo 2

Conceitos Básicos

Neste capítulo apresentamos alguns dos elementos que alicerçam o desenvolvimento deste trabalho. Dentre eles, o espaço amostral, σ -álgebra, filtração, funções \mathcal{F} -mensuráveis e esperança condicional. Estes conceitos serão úteis tanto para especificar os processos de Hawkes, exposto no Capítulo 3, quanto para caracterizar a série de dados do book de ofertas mediante este processo.

2.1 Modelos Determinísticos e Estocásticos

Em muitos casos ao estudarmos algum fenômeno específico, formulamos um modelo matemático para nos auxiliar a entender seu comportamento. Os modelos em que as condições sobre o experimento determinam o resultado, são denominados de determinísticos. No contexto de finanças, podemos considerar o seguinte exemplo,

Exemplo 1 *Suponha que o experimento, ε , consista em observar os ganhos de capital, em moeda nacional de um agente no tempo, t , quando aplicado em um determinado ativo de renda fixa. Podemos formular um modelo no qual o ganho de capital seja proporcional ao montante, p , investido no ativo. Este modelo pode ser escrito de acordo*

com a seguinte equação diferencial,

$$\frac{dp}{dt} = pf(p). \quad (2.1)$$

Sendo a f uma função que represente alíquotas de imposto de renda que varia conforme o valor de p .

Por meio da solução desta equação diferencial podemos prever com exatidão o ganho de capital de um agente no tempo, t , bastando para isto determinar as condições do investimento inicial, p_0 , e a função f .

No entanto, há outros casos práticos que não podemos atribuir uma causa determinística direta. Nestas circunstâncias, temos um experimento do tipo aleatório. Esta característica impede a predição exata dos resultados. Os fenômenos ligados a este tipo de comportamento requerem um modelo matemático não determinístico ou estocástico. Tomando como base o experimento do exemplo 1 podemos formular um modelo estocástico.

Exemplo 2 *Supondo a mesma situação do exemplo 1, podemos formular um outro modelo no qual o ganho de capital é medido em moeda estrangeira. Neste caso o ganho não é exclusivamente proporcional ao montante p e sim sujeito as oscilações cambiais. Ele pode ser escrito de acordo com a seguinte equação diferencial,*

$$\frac{dp}{dt} = pf(p) + u(p). \quad (2.2)$$

Onde $u(p)$ representa uma componente de incerteza visto que se realizarmos outro experimento as condições do câmbio mudam e conseqüentemente o resultado.

No caso do exemplo 2, ao solucionarmos a equação 2.2 não poderemos prever o ganho de capital a partir das condições do experimento, diferentemente do exemplo 1. Para isto, necessitamos de trabalhar com modelos probabilísticos. No presente trabalho

estamos diante destes modelos. Os elementos mais fundamentais para lidar com estes tipos de modelos são abordados na próxima seção.

2.2 Probabilidade e Processos Estocásticos

Ao executar um experimento do tipo aleatório, temos algumas características em comum. Primeiro, apesar da impossibilidade de prever o resultado com exatidão é possível descrever o conjunto de todos os possíveis resultados. Segundo, o experimento pode ser realizado sob as mesmas condições. Por último, conforme ocorre a repetição do experimento emerge padrões na frequência dos resultados.

Ao conjunto de todos os possíveis resultados denominamos de espaço amostral. Formalmente, estabelecemos o espaço amostral da seguinte forma.

Definição 1 *Seja ε um experimento aleatório. O conjunto de todos os resultados possíveis de ε é conhecido como espaço amostral e é denotado por Ω . A cada elemento deste conjunto denominamos ponto amostral.*

Ao longo de todo o texto vamos sempre considerar exemplos envolvendo o caso de um book de ofertas a fim de introduzir os conceitos dentro da perspectiva de finanças. De forma ilustrativa, considere o seguinte caso,

Exemplo 3 *Um experimento consiste em observar o número de ofertas de compra e venda encaminhada ao mercado a cada hora. Considere também que este mercado fique em funcionamento durante sete horas por dia. Todos os possíveis números de ofertas realizadas nas sete horas compõem o espaço amostral de nosso experimento. Este conjunto é descrito por,*

$$\Omega = \{\omega = (S_1, S_2, \dots, S_7) : S_j \in \mathbb{N}, 1 \leq j \leq 7\},$$

onde ω é um ponto amostral e S_j representa o número de ofertas encaminhadas ao

book na j -ésima hora, $1 \leq j \leq 7$.

Com base neste conceito podemos definir um evento.

Definição 2 *Um evento A relativo a um espaço amostral Ω associado a um experimento ε é todo e qualquer subconjunto do espaço amostral Ω .*

Retornando o Exemplo 3, à medida que o tempo passa são obtidas informações a respeito do possível ponto amostral de nosso experimento. Se considerarmos que no tempo $t = 1$, o número de ofertas encaminhadas ao book é $S_1 = 10$, então alguns pontos amostrais em Ω não poderão ocorrer, o que reduz os possíveis resultados a um evento, tal que,

$$A = \{\omega : (S_1 = 10, S_2 = s_2, \dots, S_7 = s_7)\}.$$

Portanto, no tempo $t = 1$, tem-se o conhecimento que o futuro ponto amostral está em A e não em $\Omega \setminus A = A^C$. Para elucidar melhor a inserção destas informações, consideremos o seguinte exemplo.

Exemplo 4 *Seja S_t a observação de chegadas de ordens ao book de um ativo de baixa liquidez no tempo $t = 1, 2$, tal que $S_t = \alpha$ quando há ofertas, $S_t = \beta$ quando não há ofertas e $t = 1$ e $t = 2$ representa o turno matutino e vespertino respectivamente. Desta forma, o conjunto dos possíveis resultados são,*

$$\Omega = \{\omega_1 = (\alpha, \alpha), \omega_2 = (\alpha, \beta), \omega_3 = (\beta, \beta), \omega_4 = (\beta, \alpha)\}.$$

Antes de começar o dia não temos conhecimento algum sobre a quantidade de ofertas, logo, o conjunto de possíveis observações futuras é $\mathcal{F}_0 = \{\Omega, \emptyset\}$. Supondo agora que possuímos a informação de que houve oferta no período matutino, $S_1 = \alpha$, então temos o seguinte evento

$$A = \{\omega_1 = (\alpha, \alpha), \omega_2 = (\alpha, \beta)\}.$$

Assim, sabemos que o resultado está em A e não em seu complemento. Portanto, nosso conhecimento no tempo $t = 1$ da série de ofertas é,

$$\mathcal{F}_1 = \{\Omega, \emptyset, A, \bar{A}\}.$$

Com a mudança do tempo $t = 0$ para $t = 1$ o conhecimento da série é ampliado. Em outras palavras tem-se que $\mathcal{F}_0 \subset \mathcal{F}_1$, pois no tempo $t = 1$ ainda é armazenada a informação no período anterior, $t - 1$.

Genericamente o conjunto, \mathcal{F}_t de subconjuntos é uma coleção de eventos do espaço amostral. Este conjunto representa a informação que temos em relação ao ponto amostral até o instante de tempo t . Dizer que o evento A ocorre significa que algum resultado do experimento $\omega \in A$ ocorreu. No Exemplo 4 o conjunto das informações até o tempo $t = 2$ é,

$$\begin{aligned} \mathcal{F}_2 = \{ & \Omega, \emptyset, \{\omega_1\}, \{\omega_2\}, \{\omega_3\}, \{\omega_4\}, \{\omega_1, \omega_2\}, \{\omega_1, \omega_3\}, \{\omega_1, \omega_4\}, \{\omega_2, \omega_3\}, \\ & \{\omega_2, \omega_4\}, \{\omega_3, \omega_4\}, \{\omega_1, \omega_2, \omega_3\}, \{\omega_1, \omega_2, \omega_4\}, \{\omega_2, \omega_3, \omega_4\}, \{\omega_1, \omega_3, \omega_4\} \} \end{aligned} \quad (2.3)$$

A esta coleção de eventos denota-se por σ -álgebra, o qual é definida abaixo.

Definição 3 Um conjunto não vazio \mathcal{F} de coleções de subconjuntos de Ω denomina-se σ -álgebra se possui as seguintes propriedades (Hoel et al., 2004),

i) Se $A \in \mathcal{F}$ então $\bar{A} \in \mathcal{F}$

ii) Se $A_n \in \mathcal{F}$, $n = 1, 2, \dots$ então $\bigcup_{n=1}^{\infty} A_n$ e $\bigcap_{n=1}^{\infty} A_n \in \mathcal{F}$

A partir dos conceitos de espaço amostral e σ -álgebra podemos estabelecer um espaço mensurável.

Definição 4 Dado um espaço amostral Ω , não vazio e uma σ -álgebra, \mathcal{F} , dizemos que o par (Ω, \mathcal{F}) forma um espaço mensurável.

A medida que o tempo passa a quantidade de informação em que vai sendo inserida aos agentes no mercado amplia-se. Podemos perceber isto por meio do Exemplo 4. Conforme o tempo passa, a coleção de subconjuntos aumenta. Isto pode ser verificado ao observar que $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \mathcal{F}_2$. Para uma situação em que coletamos informações num período de tempo maior tem-se que,

$$\mathcal{F}_1 \subset \mathcal{F}_2 \subset \mathcal{F}_3, \dots, \subset \mathcal{F}_t \subset \mathcal{F}_{t+1}.$$

A maior σ -álgebra gerada por todos os subconjuntos $A \subseteq \Omega$ será denotada por 2^Ω . No Exemplo 4, a maior σ -álgebra, expressa da equação 2.3 é $\mathcal{F}_2 = 2^\Omega$. Ao conjunto formado pelas coleções de σ -álgebras,

$$\mathbf{F} = \{\mathcal{F}_1, \mathcal{F}_1, \mathcal{F}_2, \dots, \mathcal{F}_t, \dots, \mathcal{F}_T\},$$

e denotamos como *filtração*, o qual segue-se sua definição,

Definição 5 Seja (Ω, \mathcal{F}) um espaço mensurável e $T > 0$ tal que assuma para cada $t \in [0, T]$ uma σ -álgebra \mathcal{F}_t . Quando para quaisquer s e t onde $T > s > t > 0$ tem-se que $\mathcal{F}_t \subset \mathcal{F}_s$, então a família $\{\mathcal{F}_t\}_{t \in T}$ de σ -álgebra é chamada de *filtração*.

Do ponto de vista de finanças, a filtração mostra a ordenação da informação que inserida ao longo do tempo para os agentes no mercado financeiro. No contexto do Exemplo 4, a medida que o tempo transcorre, o agente possui mais informações acerca da série de chegadas de ofertas.

Outro ponto igualmente importante é o conceito de variável aleatória (v.a.).

Definição 6 Seja o espaço mensurável (Ω, \mathcal{F}) a função $X : \Omega \rightarrow \mathbb{R}$ é chamada de

\mathcal{F} -mensurável ou de variável aleatória se

$$X^{-1}(I) = \{\omega \in \Omega : X(\omega) \in I\} \subset \mathcal{F}$$

para todo intervalo $I \subset \mathbb{R}$.

Em outras palavras, se temos a informação de qual elemento de \mathcal{F} ocorreu, então sabemos o valor de X . Portanto, se $\mathcal{F} = 2^\Omega$ então qualquer função em Ω é uma variável aleatória (Klebaner, 2004). Para ilustrar melhor segue-se um exemplo.

Exemplo 5 Considere novamente o contexto do Exemplo 4, em que observamos em dois períodos se houve ou não ofertas. Tomando como evento $A = \{\omega_1, \omega_2\}$, em que há chegadas de ofertas no período matutino e a σ -álgebra $\mathcal{F}_1 = \{\Omega, \emptyset, A, \bar{A}\}$ podemos definir a seguinte variável aleatória,

$$X(\omega) = \begin{cases} 1, & \text{se } \omega \in A \\ 0, & \text{se } \omega \in \bar{A}. \end{cases}$$

A variável aleatória X é \mathcal{F}_1 -mensurável em Ω pois o conjunto $X^{-1}(1) = A = \{\omega_1, \omega_2\} \in \mathcal{F}_1$ bem como $X^{-1}(0) = A^C = \{\omega_3, \omega_4\} \in \mathcal{F}_1$. Como podemos ver, para cada valor da variável aleatória X tem-se um evento em \mathcal{F}_1 correspondente.

No mesmo sentido, podemos formular um contra exemplo de uma variável aleatória que não é \mathcal{F}_t -mensurável necessitando, assim, definir outra σ -álgebra para torná-la mensurável. O exemplo a seguir têm como base os Exemplos 4 e 5.

Exemplo 6 Considere agora a seguinte v.a.

$$Y(\omega) = \begin{cases} 0, & \text{se } \omega \in \{\omega_3\} \\ 1, & \text{se } \omega \in \{\omega_2, \omega_4\} \\ 2, & \text{se } \omega \in \{\omega_1\} \end{cases}$$

em que computa o número de ofertas nos 2 períodos. Pode-se comprovar que esta função não é \mathcal{F}_1 -mensurável, pois o conjunto,

$$Y^{-1}(1) = B_1 = \{\omega_2, \omega_4\} \notin \mathcal{F}_1.$$

Entretanto, se considerarmos uma σ -álgebra definida como \mathcal{F}_2 podemos ver que $Y(\omega)$ é \mathcal{F}_2 -mensurável.

Uma vez estabelecido o conceito de variável aleatória podemos definir o conceito de processo estocástico.

Definição 7 Um processo estocástico é uma coleção de v.a.'s,

$$\{X_t\}_{t \in T}$$

definido sobre o espaço mensurável (Ω, \mathcal{F}) e assumindo valores em \mathbb{R} .

Quando T é enumerável, temos um processo estocástico a tempo discreto. Caso contrário, temos um processo a tempo contínuo. Além disso, se o conjunto de valores do processo for não enumerável, então o processo tem espaço de estado geral. Por outro lado, se o conjunto de valores for enumerável, o processo tem espaço de estados discreto.

Com este conceito em mente podemos definir agora um processo adaptado a filtração.

Definição 8 Um processo $\{X_t\}_{t \in T}$ é dito adaptado à filtração $\{\mathcal{F}_t\}$ se para todo t , X_t é \mathcal{F}_t -mensurável.

Para tornar mais claro o conceito de processo adaptado a filtração segue um exemplo abaixo,

Exemplo 7 Tomando o caso do exemplo 4, podemos formular o processo estocástico,

$$\{X_t\}_{t \in \{1,2\}},$$

em que X_t representa o total de ofertas encaminhadas ao book até o tempo t . É possível perceber que este processo é adaptado pela seguinte filtração,

$$\mathbb{F} = \{\mathcal{F}_1, \mathcal{F}_2\}.$$

Pois em $t = 1$, a v.a. X_1 é \mathcal{F}_1 -mensurável enquanto que em $t = 2$ a v.a. X_2 é \mathcal{F}_2 -mensurável.

Outro conceito importante é o de σ -álgebra gerada por uma v.a. No caso discreto, uma σ -álgebra é dita gerada por uma variável aleatória quando o conjunto de eventos

$$A_i = \{\omega \in \Omega : X(\omega) = x_i\}, i = 1, \dots, k$$

que forma uma partição de Ω , isto é,

$$\begin{aligned} \Omega &= \bigcup_{i=1}^k A_i \\ &= \bigcup_{i=1}^k \{X = x_i\}, \end{aligned}$$

gera uma σ -álgebra. A menor σ -álgebra que contenha todos os eventos $A_i = \{X = x_i\}$ é denotado por \mathcal{F}_X ou $\sigma(X)$.

Exemplo 8 Utilizando o caso do Exemplo 5, a variável aleatória S_1 , que representa se houve ou não oferta no período matutino do mercado, forma a seguinte partição de Ω ,

$$A = \{\omega \in \Omega : S_1 = 1\} = \{\omega_1, \omega_2\} \text{ e } \bar{A} = \{\omega \in \Omega : S_1 = 0\} = \{\omega_3, \omega_4\}$$

e conseqüentemente gera a seguinte σ -álgebra,

$$\mathcal{F}_{S_1} = \{\Omega, \emptyset, A, \bar{A}\}.$$

Finalmente definimos um espaço de probabilidade como,

Definição 9 *Seja (Ω, \mathcal{F}) , um espaço mensurável. A medida de probabilidade \mathcal{P} é uma função que leva cada conjunto $A \in \mathcal{F}$ a um número no intervalo $[0,1]$, chamada de probabilidade de A e é denotada por \mathcal{P} . A medida de probabilidade tem as seguintes propriedades,*

a) $\mathcal{P}(\Omega) = 1$

b) *Para qualquer seqüência enumerável disjunta $A_1, A_2, \dots \in \mathcal{F}$,*

$$\mathcal{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathcal{P}(A_i).$$

O trinômio $(\Omega, \mathcal{F}, \mathcal{P})$ defini o espaço de probabilidade.

2.3 Esperança Condicional

Agora iremos tratar do conceito de esperança condicional de uma variável aleatória. Quando não temos informação a respeito de uma determinada variável aleatória, X , uma possível previsão de seu valor é através da obtenção de sua esperança matemática, $\mathbb{E}[X]$. Diferentemente da esperança incondicional, a condicional, $\mathbb{E}[X|Y]$ provê algum conhecimento sobre a variável aleatória, X , por meio da variável Y , mas não completamente.

Definição 10 *Seja uma variável aleatória,*

$$S(\omega) : \Omega \longrightarrow \{s_1, s_2, \dots, s_p\}$$

e a σ -álgebra, \mathcal{F} , gerada pelos $A_i = \{S(\omega) = s_i\}$ a esperança condicional da variável aleatória $S(\omega)$ dado \mathcal{F} é calculada como,

$$\mathbb{E}[S|\mathcal{F}] = \sum_{i=1}^p s_i P(A_i|\mathcal{F}).$$

Podemos observar que $\mathbb{E}[S|\mathcal{F}]$ é uma combinação linear de variáveis aleatórias logo ela também é uma variável aleatória. A esperança condicional possui duas propriedades básicas (Lawler, 2006):

1. O valor da variável $\mathbb{E}[S|\mathcal{F}]$ depende apenas dos valores da variável aleatória, Y que é \mathcal{F} -mensurável, logo a esperança condicional é uma função de Y , $\mathbb{E}[S|Y] = \phi(Y)$.
2. Sendo A um evento que depende de Y e I_A definindo uma variável indicadora do evento A então,

$$\mathbb{E}[SI_A] = \mathbb{E}[\mathbb{E}(S|Y) I_A].$$

A partir destas duas propriedades são decorridas outras. A primeira delas ocorre se definirmos a função indicadora I_Ω desta última propriedade, então a esperança da esperança condicional será,

$$\mathbb{E}[\mathbb{E}(S|\mathcal{F})] = \mathbb{E}[S].$$

Entretanto, pela definição de mensurabilidade se a variável S for \mathcal{F} -mensurável teremos como consequência,

$$\mathbb{E}[S|\mathcal{F}] = S,$$

uma vez que a S é determinada pelo conhecimento gerado pela σ -álgebra. Agora, ao considerar a filtração em que $\mathcal{F}_m \subset \mathcal{F}_n$ então computando a esperança,

$$\mathbb{E}[\mathbb{E}[S|\mathcal{F}_n]|\mathcal{F}_m] = \mathbb{E}[S|\mathcal{F}_m],$$

ou seja, prevalece o conteúdo de menor informação. Sua demonstração se encontra no apêndice do trabalho. No caso em que o produto SY , o qual Y é \mathcal{F} -mensurável, então sua esperança será,

$$\mathbb{E}[SY|\mathcal{F}] = Y\mathbb{E}[S|\mathcal{F}].$$

Finalmente, levando em conta o caso em que a informação da σ -álgebra \mathcal{F} não é útil para minimizar a incerteza de S , então sua esperança condicional é igual a incondicional, $\mathbb{E}[S|\mathcal{F}] = \mathbb{E}[S]$.

Capítulo 3

Processos de Hawkes

No presente capítulo continuaremos o desenvolvimento da metodologia utilizando os conceitos básicos expostos no capítulo anterior. Ao longo desta seção, serão delineados também alguns conceitos adjacentes aos processos pontuais tais como: processos de contagem, duration, função intensidade, F -Compensador. Em especial, trataremos uma classe especial de função intensidade, chamada de *Linear Self-Exciting*. Dependendo da estrutura desta função intensidade, o processo é conhecido como de Hawkes. Este processo é o foco principal de nosso estudo devido a sua aplicação na modelagem do Book de ofertas de ativos no mercado financeiro.

3.1 Processos Pontuais

Um processo pontual na reta pode ser visto como uma coleção de pontos ou eventos ocorridos em determinado tempo, $\{t_i\}_{i=1}^K$. Estes pontos podem representar as ocorrências de um fenômeno específico ao longo do tempo. Os processos pontuais são utilizados em várias áreas como: finanças, física quântica, engenharia elétrica, pesquisa operacional, engenharia de tráfego etc (Cox e Ishan, 1980). Em particular, na modelagem do Book o conjunto de pontos servem para representar os tempos específicos em que as ordens

de compradores e vendedores chegam ao Book ou quando as ordens são efetivadas.

Existem quatro maneiras equivalentes de descrever uma amostra de processo pontual [Daley e Vere-Jones \(2002\)](#):

- i) Medida de contagem;
- ii) Função escada;
- iii) Sequência de intervalos;
- iv) Sequência de pontos.

As três últimas formas são usadas essencialmente para descrever um processo pontual na reta. Entretanto, a primeira é utilizada para o caso mais geral, no \mathbb{R}^n . Para ter uma noção do processo pontual a partir da medida de contagem podemos tomar qualquer subconjunto A da reta real. Sendo $N(A)$ uma medida de ocorrências do processo em A logo,

$$N(A) = \# \{i : t_i \in A\}.$$

No âmbito de finanças podemos considerar o seguinte exemplo,

Exemplo 9 *Tomando o conjunto I o intervalo de tempo correspondente ao horário de funcionamento do pregão e os pontos como uma chegada de oferta de compra, então o processo,*

$$N(I) = \text{Número de ofertas de compra durante todo o horário do pregão}$$

Quando A é expresso como uma união disjunta de intervalos A_1, \dots, A_r ou seja,

$$A = \bigcup_{i=1}^r A_i \text{ em que } A_i \cap A_j = \emptyset \text{ para } i \neq j$$

então temos que,

$$N\left(\bigcup_{i=1}^r A_i\right) = \sum_{i=1}^r N(A_i).$$

O processo é sempre $N(A) \geq 0$ e é um valor inteiro. Como queremos considerar contagens em intervalos e uniões de intervalos, então temos também que $N(A)$ é definido nos conjuntos dos Borelianos. Para excluir o "acúmulo intenso" de pontos temos a restrição de que $N(A) < \infty$ para todo conjunto A limitado. Por meio destas características será permitido estender $r = \infty$ e fazer $N(\cdot)$ com uma medida de contagem sobre a σ -álgebra $\mathcal{B}_{\mathbb{R}}$. Para mais detalhes vide [Daley e Vere-Jones \(2002\)](#).

Outro conceito importante é o de estacionariedade.

Definição 11 *Um processo pontual é estacionário quando para todo $r = 1, 2, \dots$ e todo o conjunto limitado de Borel A_1, A_2, \dots, A_r sobre a linha real, a função de distribuição conjunta de,*

$$\{N(A_1 + t), N(A_2 + t), \dots, N(A_r + t)\},$$

não depende do valor de t ($-\infty < t < \infty$).

Para distinguir $N(\cdot)$, como uma função de conjunto, de uma função real, vamos utilizar a notação $N(a, b]$ quando o conjunto A for um intervalo semi-aberto $(a, b]$. Assim, as seguintes notações são equivalentes:

$$N(t) = N(0, t] = N_t.$$

Agora estamos aptos a escrever o processo por meio da uma função escada descrita na segunda abordagem. Esta função também pode ser vista como uma soma da função Heaviside,

$$\theta(t) = \begin{cases} 0, & \text{se } t < 0 \\ 1, & \text{se } t \geq 0 \end{cases}$$

Assim, temos que,

$$N_t = \sum_{t_i < t} \theta(t - t_i),$$

em que o indexador i representa os pontos. Por convenção, tomamos como origem do processo, $t = 0$ logo $N(t = 0) = 0$, ou seja, não há a ocorrência de eventos antes e no instante zero do processo. Esta função é não decrescente e contínua pela direita com valores inteiros positivos, como pode ser visto no gráfico abaixo.

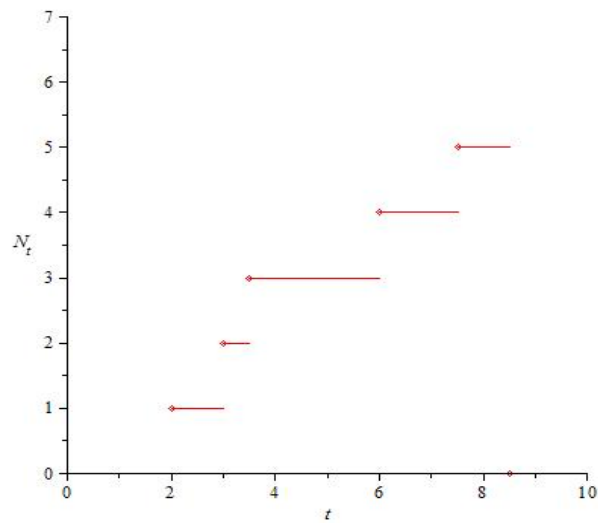


Figura 3.1: Processo Pontual: Eventos, processo de contagem e as *durations*

Os pontos de saltos representam o processo pontual $t_{i \in \mathbb{N}^*}$. O número de saltos ocorridos até o tempo t representa o processo de contagem e o comprimento de cada segmento de reta em vermelho descreve os intervalos. Neste trabalho, podemos interpretar os saltos do gráfico como as chegadas de ofertas tanto de compra quanto de venda e as retas, em vermelho, como o tempo em que cada ordem de transação fica disponível no Book de Ofertas.

A partir da terceira abordagem, dos intervalos, podemos introduzir o seguinte conceito

Definição 12 O processo dos intervalos $\{\tau_i\}_{i \in \mathbb{N}^*}$ definido por

$$\tau_i = t_i - t_{i-1}, \forall i \in \mathbb{N}^*.$$

é chamado de processo de duração ou duration associado a um processo pontual.

e de forma análoga,

Definição 13 Um processo pontual é estacionário nos intervalos quando para todo $r = 1, 2, \dots$ e para todo i_1, \dots, i_r a distribuição conjunta de

$$\{\tau_{i_1+k}, \dots, \tau_{i_r+k}\},$$

não depende do valor de k ($k = \pm 1, \dots$).

Para nosso propósito podemos agora definir o processo pontual da seguinte forma,

Definição 14 Uma sequência de valores, $\{t_i\}_{i \in \mathbb{N}^*} \in \mathbb{R}^+$, de variáveis aleatórias tais que $\forall i \in \mathbb{N}^*, t_i < t_{i+1}$ definido em um espaço de probabilidade (Ω, \mathcal{F}, P) é dito um processo pontual.

Até este ponto abordamos os processos pontuais sobre perspectivas mais amplas com $N(A)$, $N(t)$, τ_i e t_i . O processo mais básico $N(t)$ é o processo de Poisson o qual será descrito na próxima seção.

3.2 Processos de Poisson

Um processo de Poisson estacionário sobre a reta é definido pela seguinte função de probabilidade em que novamente utilizaremos a notação $N(a_i, b_i]$ para representar o número de ocorrências no intervalo $(a_i, b_i]$, sendo $a_i < b_i \leq a_{i+1}$,

$$P \{N(a_i, b_i] = n_i, i = 1, 2, \dots, k\} = \prod_{i=1}^k \frac{(\lambda(b_i - a_i))^{n_i}}{n_i!} e^{-\lambda(b_i - a_i)}. \quad (3.1)$$

A partir desta distribuição tem-se três características básicas

- i) O número de pontos em cada intervalo finito $(a_i, b_i]$ segue uma distribuição de Poisson.
- ii) O número de pontos em intervalos disjuntos são variáveis aleatórias independentes.
- iii) A distribuição é estacionária (veja na Definição 12)

A partir da função (3.1) podemos descrever algumas propriedades. A esperança matemática de $N((a,b])$ é,

$$\begin{aligned}
 \mathbb{E} [N((a,b])] &= \sum_{n=1}^{\infty} n \frac{(\lambda(b-a))^n}{n!} e^{-(b-a)} \\
 &= \sum_{n=1}^{\infty} \frac{\lambda(b-a) (\lambda(b-a))^{n-1}}{(n-1)!} e^{-(b-a)} \\
 &= (\lambda(b-a)) \sum_{n=1}^{\infty} \frac{(\lambda(b-a))^{n-1}}{(n-1)!} e^{-(b-a)} \\
 &= \lambda(b-a) 1
 \end{aligned} \tag{3.2}$$

O cálculo do segundo momento $\mathbb{E} [N((a,b])^2]$ é análogo a esperança. Logo a distribuição de Poisson tem a seguinte característica,

$$\mathbb{E} \{N(a,b)\} = \lambda(b-a) = \mathbb{V}ar \{N(a,b)\}.$$

A constante λ é interpretada como a taxa média de ocorrência ou a intensidade média de pontos do processo. Com a característica de que a média assim como a variância são proporcionais ao intervalo analisado, a taxa λ provê uma maneira útil de fazer um diagnóstico de estacionariedade do processo de Poisson. Para isto, podemos estimar $\mathbb{E} \{N(a,b)\}$ e a $\mathbb{V}ar \{N(a,b)\}$ para diferentes intervalos semi-abertos $(a,b]$ e plotar a seguinte razão,

$$\eta = \frac{\mathbb{V}ar \{N(a,b)\}}{(b-a)}$$

O valor η deve ser aproximadamente constante para processos de Poisson estacionários e igual a λ , ($\hat{\eta} \approx \lambda$). Qualquer alteração deste regime pode indicar uma não conformidade com o processo de Poisson ou estacionariedade (Daley e Vere-Jones, 2002).

A partir de (3.1), podemos extrair a seguinte função,

$$P\{N(0,\tau) = 0\} = e^{-\lambda\tau}. \quad (3.3)$$

Ela representa a probabilidade de não haver nenhum ponto sobre o intervalo $(0,\tau]$. Isto é, ela pode ser interpretada como a probabilidade de o primeiro ponto ocorrer em algum momento depois de τ . A função que descreve este comportamento é chamada de função sobrevivência. Como a função sobrevivência pode ser escrita como $1 - P(X \leq x)$ então temos de (3.3) que a variável aleatória τ , ou seja, a *duration* segue uma distribuição exponencial com parâmetro λ . Tomando a propriedade da falta de memória (*memory-less*) da função exponencial, podemos descrever a distribuição dos intervalos entre dois pontos consecutivos com a distribuição exponencial. Esta relação pode ser comprovada verificando a equivalência dos eventos,

$$\{t_k \geq x\} \Leftrightarrow \{N(0,x) < k\}.$$

Daí, segue-se que

$$P\{t_k > x\} = P\{N(0,x) < k\} = \sum_{n=0}^{k-1} \frac{(\lambda x)^n}{n!} e^{-\lambda x}. \quad (3.4)$$

Tomando a derivada de (3.4) com relação a x e obtemos a função densidade de Erlang,

$$f_k(x) = \frac{\lambda^k x^{k-1}}{(k-1)!} e^{-\lambda x} = \frac{\lambda^k}{\Gamma(k)} x^{k-1} e^{-\lambda x}. \quad (3.5)$$

maiores detalhes podem ser vistos no Apêndice B. A função (3.5) representa a distri-

buição da soma dos k intervalos. Sendo os intervalos independentes, este resultado já era esperado uma vez que a soma de variáveis aleatórias exponenciais segue a distribuição de Erlang. Estas características serão importantes para montar a função de verosimilhança de uma realização finita de um processo de Poisson.

3.3 Função Intensidade

A partir de um processo de contagem uma questão natural surge é quanto a frequência de aparecimento de pontos. Uma maneira intuitiva de medir a densidade de pontos é observar o número de ocorrências média. Se o processo pontual é estacionário então podemos estabelecer a seguinte medida,

$$m = \mathbb{E} \{ (0,1] \}.$$

Outra maneira de mensurar a taxa de ocorrências de pontos pode ser feita por,

$$\lim_{\Delta \rightarrow 0} \frac{P \{ N(0, \Delta] > 0 \}}{\Delta}. \quad (3.6)$$

O valor λ é chamado de intensidade de um processo pontual quando $\lambda < \infty$. Uma outra forma de reescrever o numerador da função (3.6) é por,

$$\begin{aligned} P \{ N(x, x + \Delta] > 0 \} &= P \{ \text{Ocorrer pelo menos um ponto no intervalo } (x, x + \Delta] \} \\ &= \lambda \Delta + o(\Delta) \end{aligned}$$

Ambas as medidas de taxa de ocorrências em um processo pontual estacionário coincidem quando o processo é dito simples.

Definição 15 *Um processo pontual é simples quando,*

$$P(N(\{t\}) = 0) + P(N(\{t\}) = 1) = 1,$$

para todo t .

No caso em que a taxa, $\lambda(t)$, muda ao longo do tempo, sua forma incondicional é definida como segue.

Definição 16 *Seja N_t um processo de contagem associado a um processo pontual $\{t_i\}_{i \in \mathbb{N}^*}$. Então, a função intensidade é definida por,*

$$\lambda(t) = \lim_{\Delta t \rightarrow 0} \frac{P(N_{t+\Delta t} - N_t > 0)}{\Delta t} \quad (3.7)$$

Observamos que a função intensidade especificada a cima não leva em conta as ocorrências passadas, entretanto, para a abordagem de nosso estudo não há garantia de que os eventos precedentes não tenham influência nas ocorrências futuras. A suposição de dependência é baseada na ideia de que os agentes, na medida em que as informações são inseridas no mercado, vão se posicionando e efetuando novas ordens. Logo, dependendo dos movimentos precedentes do mercado o agente é influenciado a inserir suas ofertas no Book. Assim, pode-se definir a função intensidade condicional, $\lambda(t|\mathcal{F}_t)$, de modo análogo ao caso incondicional.

Definição 17 *Seja N_t um processo de contagem associado a um processo pontual $\{t_i\}_{i \in \mathbb{N}^*}$, adaptado a uma filtração \mathcal{F}_t . Então, a função intensidade condicional é definida por,*

$$\lambda(t|\mathcal{F}_t) = \lim_{\Delta t \rightarrow 0} \frac{P(N_{t+\Delta t} - N_t > 0|\mathcal{F}_t)}{\Delta t}.$$

Podemos expressar a função intensidade condicional também em termos de sua

esperança matemática.

$$\mathbb{E}[N_{t+\Delta t} - N_t | \mathcal{F}_t] = 1P(N_{t+\Delta t} - N_t = 1 | \mathcal{F}_t) + \sum_{n=2}^{\infty} P(N_{t+\Delta t} - N_t \geq n | \mathcal{F}_t) \quad (3.8a)$$

$$\Leftrightarrow \frac{\Delta t}{\Delta t} \mathbb{E}[N_{t+\Delta t} - N_t | \mathcal{F}_t] = \lambda(t | \mathcal{F}_t) \Delta t + o(\Delta t) \quad (3.8b)$$

$$\Leftrightarrow \mathbb{E} \left[\frac{N_{t+\Delta t} - N_t}{\Delta t} | \mathcal{F}_t \right] = \lambda(t | \mathcal{F}_t) + \frac{o(\Delta t)}{\Delta t} \quad (3.8c)$$

O somatório da primeira linha, (3.8a), é igual a $o(\Delta t)$ da segunda linha, (3.8b), para ver isto note que cada parcela da série, (3.8a), tem a seguinte probabilidade,

$$P(N_{t+\Delta t} - N_t \geq n | \mathcal{F}_t) = o(\Delta t), \forall n \in \mathbb{N} \geq 2.$$

Como $o(\Delta t) + o(\Delta t) = o(\Delta t)$, então o somatório da linha (3.8a) fica igual a $o(\Delta t)$.

Tomando o limite em (3.8c) fica claro que

$$\lim_{\Delta t \rightarrow 0} \mathbb{E} \left[\frac{N_{t+\Delta t} - N_t}{\Delta t} | \mathcal{F}_t \right] = \lambda(t | \mathcal{F}_t). \quad (3.9)$$

Pode-se notar que a função intensidade depende da escolha da filtração, mas será adotada a convenção da filtração gerada pelas variáveis aleatórias $\{t_i\}_{t_i < t}$. Agora, aplicando o operador esperança na equação (3.8c) observamos que seu valor será não correlacionado aos valores pretéritos, ou seja, ela não terá dependência da filtração estabelecida, \mathcal{F}_t . Portanto, rearranjando a equação (3.8c) e tomando seu limite quando $\Delta t \rightarrow 0$, tem-se que,

$$\mathbb{E} \left\{ \lim_{\Delta t \rightarrow 0} \mathbb{E}[N_{t+\Delta t} - N_t | \mathcal{F}_t] - \lambda(t | \mathcal{F}_t) \Delta t \right\} = \mathbb{E} \left[\lim_{\Delta t \rightarrow 0} \frac{o(\Delta t)}{\Delta t} \right] \quad (3.10a)$$

$$\Leftrightarrow \mathbb{E} \left[\lim_{\Delta t \rightarrow 0} (N_{t+\Delta t} - N_t) - \lambda(t | \mathcal{F}_t) \Delta t \right] = 0. \quad (3.10b)$$

Pela propriedade de linearidade do operador esperança podemos obter a soma dos eventos entre os tempos s_0 e s_1 nos mesmos moldes da Equação (3.10b).

$$\mathbb{E} \left[\lim_{\Delta t \rightarrow 0} \sum_{k=1}^{\frac{s_1-s_0}{\Delta t}} (N_{s_0+k\Delta t} - N_{s_0+(k-1)\Delta t}) - \lambda(s_0 + k\Delta t | \mathcal{F}_t) \Delta t \right] = 0, \quad (3.11a)$$

$$\mathbb{E} \left[\lim_{\Delta t \rightarrow 0} (N_{s_1} - N_{s_0}) - \sum_{k=1}^{\frac{s_1-s_0}{\Delta t}} \lambda(s_0 + k\Delta t | \mathcal{F}_t) \Delta t \right] = 0, \quad (3.11b)$$

$$\mathbb{E} \left[\int_{s_0}^{s_1} \lambda(s_0 | \mathcal{F}_t) dt \right] = N_{s_1} - N_{s_0}. \quad (3.11c)$$

A integral da função intensidade de dois eventos quaisquer, s_i e s_{i+1} , é chamada de \mathcal{F}_t -Compensator e é denotada por,

$$\Lambda(s_i, s_{i+1}) = \int_{s_i}^{s_{i+1}} \lambda(t | \mathcal{F}_t) dt, \quad \forall i \in \mathbb{N}^*. \quad (3.12)$$

Note que a esperança matemática dos F -compensators de dois eventos consecutivos, $\mathbb{E}[\Lambda(s_i, s_{i+1})] = 1$, pois $N_{s_i} - N_{s_{i+1}} = 1$. Tomando a \mathcal{F}_t segue o teorema abaixo.

Teorema 1 *Seja N_t um processo pontual simples definido em \mathbb{R}_+ adaptado a uma filtração \mathcal{F}_t tal que a função intensidade $\int_0^\infty \lambda(t | \mathcal{F}_t) = \infty$ e \mathcal{F} -Compensators $\Lambda^*(t) = \int_0^t \lambda(u) du$. Então a transformação do tempo aleatório $t \mapsto \Lambda(0, t)$ torna o processo,*

$$\tilde{N}(t) = N(\Lambda^{-1}(0, t)),$$

em um processo de Poisson homogêneo com taxa $\lambda = 1$.

Por consequência desta transformação do tempo, temos que $\Lambda(t_i, t_{i+1}) \sim \text{Exp}(\lambda = 1)$.

Este resultado é importante para estabelecer um teste de ajustamento dos dados aos

parâmetros do modelo do processo de Hawkes. Um método conveniente e direto se baseia em plotar os quantis teóricos de uma distribuição exponencial ($\lambda = 1$) contra os quantis empíricos da distribuição dos *compensators* estimados. Caso haja um bom ajustamento do modelo, os pontos do gráfico conhecido também como QQ-Plot, devem situar sobre a reta bissetriz. Outro desdobramento importante do Teorema 1 é quanto a sua utilização para gerar simulações com estrutura de um processo pontual, com função intensidade $\lambda(t|\mathcal{F}_t)$.

As funções de intensidade do tipo *linear Self-Exciting* de um processo pontual têm uma estrutura genérica para captar as influências do passado. Segue-se abaixo a definição de sua estrutura.

Definição 18 *A forma geral de uma função intensidade do tipo linear Self-Exciting de um processo N_t é,*

$$\lambda(t) = \lambda_0(t) + \int_{-\infty}^t \nu(t-s) dN_s, \quad (3.13a)$$

$$= \lambda_0(t) + \sum_{t_k < t} \nu(t-t_k), \quad (3.13b)$$

onde a função $\lambda_0(t)$ é determinística e a função $\nu : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ (Cox e Ishan, 1980).

As influências pretéritas são computadas e ponderadas pela função ν em cada tempo. Pelas características desta função, cada ponto ocorrido no passado exerce influências positivas nos eventos futuros. A função $\lambda_0(t) = \lambda_0$ pode ser definida como uma constante e neste caso é interpretada com a função de intensidade de equilíbrio no longo prazo.

Assumindo a estacionariedade da função intensidade, $\lambda(t)$, tem-se a condição necessária de que $\mathbb{E}[\lambda(t)] = \mu$. Esta característica é observada no tipo *linear Self-Exciting* quando definimos a composição determinística, ($\lambda_0(t) = \lambda_0$). Logo, para calcular a

média da função intensidade aplicamos o operador esperança na Equação

$$\mu = \mathbb{E}[\lambda(t)], \quad (3.14a)$$

$$\mu = \mathbb{E} \left[\lambda_0(t) + \int_{-\infty}^t \nu(t-s) dN_s \right], \quad (3.14b)$$

$$\mu = \mathbb{E} \left[\lambda_0(t) + \int_{-\infty}^t \nu(t-s) \lambda(s) ds \right]. \quad (3.14c)$$

Como $\mathbb{E}[dN_t] = \mathbb{E}[\lambda(t)ds]$ ¹ e tomando $h = t - s$ temos,

$$\mu = \lambda_0 + \int_{-\infty}^t \nu(t-s) \mu ds, \quad (3.15a)$$

$$\mu = \lambda_0 + \mu + \int_0^{\infty} \nu(h) dh. \quad (3.15b)$$

Sendo a função contínua $\nu(t) : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ e a integral $\int_{-\infty}^t \nu(t) < \infty$, logo podemos inserir o operador esperança dentro da integral. Sendo a variável $h = t - s$ e multiplicando por -1 a Equação (3.15b) para inverter os limites de integração e isolando μ tem-se,

$$\mu = \frac{\lambda_0}{1 - \int_0^{\infty} \nu(h) dh}. \quad (3.16)$$

No caso em que $\lambda_0(t)$ não é constante podemos subtraí-lo da Equação (3.15b) para retirar a componente determinística.

¹Pela Equação (3.10b) tem-se que $\mathbb{E} \left[\lim_{\Delta t \rightarrow 0} \frac{N_{t+\Delta t} - N_t}{\Delta t} \right] = \mathbb{E}[\lambda(t|\mathcal{F}_t)]$

3.4 Processo de Hawkes

Nesta seção, apresentamos uma estrutura particular da função intensidade do tipo *linear Self-Exciting*. Este tipo de estrutura foi proposto por [Hawkes \(1971\)](#). As aplicações inicialmente foram feitas na modelagem de abalos sísmicos, em que a taxa de ocorrência de terremotos são alteradas por eventos precedentes. No mercado de financeiro temos o caso análogo em que os eventos são representados pela chegadas de ofertas de compra e venda.

3.4.1 Processo de Hawkes Univariado

[Hawkes \(1971\)](#) propôs o kernel de uma distribuição exponencial como a função $\nu(h)$ da Equação (3.13). Neste caso, as influências dos eventos passados tem um decaimento exponencial no cálculo da taxa instantânea de ocorrência. No processo de Hawkes unidimensional, a função ν adquire a seguinte expressão,

$$\nu(t) = \sum_{j=1}^P \alpha_j e^{-\beta_j t}, P \in \mathbb{N}. \quad (3.17)$$

O valor P remete a ordem do modelo. O caso mais simples do processo de Hawkes é quando $P = 1$. Esta configuração da função intensidade será usada nas aplicações deste trabalho. A condição de estacionaridade no processo univariado é que na Equação (3.17),

$$\sum_{j=1}^P \frac{\alpha_j}{\beta_j} < 1.$$

Portanto, a função intensidade média incondicional do processo unidimensional é calculada por,

$$\mu = \frac{\lambda_0}{1 - \sum_{j=1}^P \frac{\alpha_j}{\beta_j}}. \quad (3.18)$$

Inserindo o kernel (3.17) na expressão (3.13) temos a seguinte função intensidade,

$$\lambda(t) = \lambda_0(t) + \int_0^t \sum_{j=1}^P \alpha_j e^{-\beta_j(t-s)} dN_s, \quad (3.19a)$$

$$= \lambda_0(t) + \sum_{j=1}^P \alpha_j \sum_{k=0}^{N_t-1} e^{-\beta_j(t-t_k)}, \quad (3.19b)$$

$$= \lambda_0(t) + \sum_{j=1}^P \alpha_j B_j(N_t). \quad (3.19c)$$

Onde a função $B_j(i)$ é dada por,

$$B_j(i) = \sum_{t_k < t_i} e^{-\beta_j(t_i - t_{i-1} + t_{i+1} - t_k)}. \quad (3.20)$$

Note que,

$$B_j(i) = e^{-\beta_j(t_i - t_{i-1})} \sum_{k=0}^{i-1} e^{-\beta_j(t_{i-1} - t_k)}, \quad (3.21a)$$

$$= e^{-\beta_j(t_i - t_{i-1})} \left(1 + \sum_{k=0}^{i-2} e^{-\beta_j(t_{i-1} - t_k)} \right), \quad (3.21b)$$

$$= e^{-\beta_j(t_i - t_{i-1})} (1 + B_j(i-1)). \quad (3.21c)$$

Esta expressão (3.21) pode ser usada para calcular $B_j(i)$ recursivamente. A função $\lambda_0(t)$ é chamada de função intensidade determinística base. A Equação (3.21) também é utilizada para calcular a função verossimilhança. Os estimadores dos parâmetros via máxima verossimilhança serão apresentados no capítulo seguinte.

Considerando dois eventos consecutivos, o *compensator* (3.12) de um processo de Hawkes, é expresso substituindo a função $\lambda(t)$ da Equação (3.12) pela função (3.19c).

Logo temos que,

$$\Lambda(t_{i-1}, t_i) = \int_{t_{i-1}}^{t_i} \left(\lambda_0(t) + \sum_{j=1}^P \alpha_j B_j(N_t) \right) dt, \quad (3.22a)$$

$$= \int_{t_{i-1}}^{t_i} \lambda_0(t) + \sum_{j=1}^P \alpha_j \sum_{k=0}^{i-1} \int_{t_{i-1}}^{t_i} e^{-\beta_j(t-t_k)} dt, \quad (3.22b)$$

$$= \int_{t_{i-1}}^{t_i} \lambda_0(t) + \sum_{k=0}^{i-1} \sum_{j=1}^P \frac{\alpha_j}{\beta_j} \left(e^{-\beta_j(t_{i-1}-t_k)} - e^{-\beta_j(t_i-t_k)} \right), \quad (3.22c)$$

$$= \int_{t_{i-1}}^{t_i} \lambda_0(t) + \sum_{j=1}^P \frac{\alpha_j}{\beta_j} \left(1 - e^{-\beta_j(t_i-t_{i-1})} \right) A_j(i-1). \quad (3.22d)$$

Na última linha, (3.22d), procedemos da mesma forma que as Equações (3.20). Isto é, os valores da função $A_j(i)$ são calculados recursivamente por,

$$A_j(i) = \sum_{t_k \leq t_i} e^{-\beta_j(t_i+t_{i-1}-t_{i-1}-t_k)}, \quad (3.23a)$$

$$= 1 + e^{-\beta_j(t_i-t_{i-1})} A_j(i-1), \quad (3.23b)$$

sendo $A_j(0) = 0$. Note que diferentemente da função (3.20) t_k vai até t_i . Tanto a função intensidade como os *F-Compensators* são calculados com o auxílio de informações pretéritas, o que facilita na construção de uma algoritmo para computa-los.

Em um primeiro momento, vamos considerar a função intensidade determinística base como uma constante, $\lambda_0(t) = \mu$. No entanto, é bem conhecido que a intensidade base, ao longo do pregão, sofre uma sazonalidade. Nas primeiras horas de negociação há uma maior intensidade nas taxas de chegadas de ofertas, reduzindo paulatinamente até o período do meio dia em que é conhecido como efeito almoço. As taxas voltam a aumentar apenas nas últimas horas do pregão. O comportamento, portanto, ao longo do dia tem o formato de um *U* (*U shape*). Para lidar com este caráter sazonal posteriormente utilizaremos uma interpolação por meio de um *spline* cúbico, $\phi(t_i)$, com

14 nós espaçados a cada 30 minutos. Cada ponto foi calculado por

$$P_j = \frac{1}{N_{(zj)} - N_{(z(j-1))}} \sum_{i=N_{(z(j-1))}}^{N_{(z(j))}} \frac{1}{t_i - t_{i-1}}, \quad j = 1, 2, \dots, 14, \quad (3.24)$$

em que z é o número de segundo em 30 minutos. Por meio da interpolação temos uma aproximação da função intensidade base determinística, $\phi(t) \approx \lambda_0(t)$.

3.4.2 Processo de Hawkes Multivariado

Na modelagem do book de ofertas, é imprescindível analisar a dinâmica mútua entre as partes dos compradores e vendedores. Para compor esta dinâmica, faz-se necessário inserir uma estrutura multivariada. Neste caso, podemos verificar como o comportamento de um processo de chegada é afetado por outro. O processo pontual multivariado é descrito a seguir.

Definição 19 *Uma sequência, $\{(t_i^m)\}_{m=1,2,\dots,M} \in \mathbb{R}^+$, M -dimensional de variáveis aleatórias em que $M \in \mathbb{N}^*$.*

Por consequência, o processo de contagem associado é definido por:

Definição 20 *Sendo $\{(t_i^m)\}_{m=1,2,\dots,M}$ um processo pontual, o processo de contagem associado é,*

$$\mathbf{N}_t = (N_t^1, N_t^2, \dots, N_t^M),$$

em que cada entrada é a soma de funções indicadoras I_{t_i} ,

$$(N_t^m)_{m=1,2,\dots,M} = \sum_{i \in \mathbb{N}^*} I_{t_i}^m.$$

A função intensidade de um processo de Hawkes multivariado M -dimensional é dada por,

$$\lambda_m(t) = \mu_m + \sum_{n=1}^M \int_0^t \sum_{j=1}^P \alpha_{j,m,n} e^{-\beta_{j,m,n}(t-s)} dN_s^n, \quad (3.25a)$$

$$= \mu_m + \sum_{n=1}^M \sum_{j=1}^P \alpha_{j,m,n} \sum_{t_k < t} e^{-\beta_{j,m,n}(t-t_k^n)}, \quad (3.25b)$$

$$= \mu_m + \sum_{n=1}^M \sum_{j=1}^P \alpha_{j,m,n} R_j^{m,n}(N_t^n). \quad (3.25c)$$

onde $m = 1, 2, \dots, M$. No presente trabalho, utilizaremos $P = 1$. Na situação multivariada, a função recursiva,

$$R_j^{m,n}(i) = \sum_{t_k^n < t_i^m} e^{-\beta_{j,m,n}(t_i^m - t_k^n)},$$

se modifica em relação a univariada para,

$$R_j^{m,n}(i) = \begin{cases} e^{-\beta_{j,m,n}(t_i^m - t_{i-1}^m)} R_j^{m,n}(i-1) + \sum_{t_i^m \leq t_k^n < t_i^m} e^{-\beta_{j,m,n}(t_i^m - t_k^n)} & \text{se } m \neq n, \\ e^{-\beta_{j,m,n}(t_i^m - t_{i-1}^m)} (1 + R_j^{m,n}(i-1)) & \text{se } m = n, \end{cases} \quad (3.26)$$

em que a condição inicial é $R_j^{m,n}(0) = 0$. Uma alternativa mais sintética a expressão (3.25), com $P = 1$, é obtida por meio da notação vetorial,

$$\boldsymbol{\lambda}(t)_{mn} = \boldsymbol{\lambda}_0 + \int_0^t \mathbf{v}(t-s) dN_s, \quad (3.27)$$

onde on vetor,

$$\mathbf{v}(t) = \left(\alpha_{mn} e^{-\beta_{mn}(t)} \right)_{m,n=1,2,\dots,M}.$$

A intensidade média, considerando o processo estacionário, isto é $\mathbb{E}[\boldsymbol{\lambda}(t)] = \boldsymbol{\mu}$, é dada por,

$$\boldsymbol{\mu} = \left(\mathbf{I} - \int_0^\infty \mathbf{v}(u) du \right)^{-1} \boldsymbol{\lambda}_0, \quad (3.28)$$

onde \mathbf{I} é a matriz identidade.

O F -compensator da m -ésima coordenada do processo é computado por,

$$\Lambda_m(t_{i-1}^m, t_i^m) = \int_{t_{i-1}^m}^{t_i^m} \mu_m + \sum_{n=1}^M \sum_{j=1}^P \sum_{t_k^n < t_i^m} \alpha_{j,m,n} e^{-\beta_{j,m,n}(s-t_k^n)} ds \quad (3.29a)$$

$$= \eta_m + \sum_{n=1}^M \sum_{t_k^n < t_{i-1}^m} \sum_{j=1}^P \frac{\alpha_{j,m,n}}{\beta_{j,m,n}} \left(e^{-\beta_{j,m,n}(t_{i-1}^m - t_k^n)} - e^{-\beta_{j,m,n}(t_i^m - t_k^n)} \right) \quad (3.29b)$$

$$+ \sum_{n=1}^M \sum_{t_{i-1}^m \leq t_k^n < t_i^m} \sum_{j=1}^P \frac{\alpha_{j,m,n}}{\beta_{j,m,n}} \left(1 - e^{-\beta_{j,m,n}(t_i^m - t_k^n)} \right),$$

onde $\eta_m = \int_{t_{i-1}^m}^{t_i^m} \mu_m ds$. Neste caso, não é possível criar uma função recursiva completa como no caso univariado. Entretanto, com a separação do somatório que computa os $t_k^n < t_i^m$ em duas partes, nos permite utilizar a função recursiva, $R_j^{m,n}(i)$. Desta forma, conseguimos economizar recursos computacionais para os valores referentes a $t_k^n < t_{i-1}^m$, restando o cálculo dos valores restantes, $t_{i-1}^m \leq t_k^n < t_i^m$. Empregando a função (3.26) podemos reescrever (3.29) da seguinte forma,

$$\Lambda_m(t_{i-1}^m, t_i^m) = \kappa_m + \int_{t_{i-1}^m}^{t_i^m} \sum_{n=1}^M \sum_{j=1}^P \sum_{t_k^n < t_i^m} \alpha_{j,m,n} e^{-\beta_{j,m,n}(s-t_k^n)} ds \quad (3.30a)$$

$$= \kappa_m + \sum_{n=1}^M \sum_{j=1}^P \frac{\alpha_{j,m,n}}{\beta_{j,m,n}} \left[\left(1 - e^{-\beta_{j,m,n}(t_i^m - t_{i-1}^m)} \right) R_j^{m,n}(i-1) \right] \quad (3.30b)$$

$$+ \sum_{t_{i-1}^m \leq t_k^n < t_i^m} \left(1 - e^{-\beta_{j,m,n}(t_i^m - t_k^n)} \right).$$

Estes valores são importantes sobre tudo para realizar uma avaliação do ajuste da série, como enunciado no Teorema 1. Caso a realização do processo siga um processo de Hawkes os $\Lambda_m(t_{i-1}^m, t_i^m)$ devem possuir uma distribuição exponencial ($\lambda = 1$).

Neste trabalho, abordamos os casos uni e bi-variado do processo de Hawkes. No primeiro caso, teremos uma estimativa para os compradores e outros para vendedores. No segundo caso, a abordagem será mútua entre compradores e vendedores. Desta maneira, é possível verificar o grau de influência das chegadas de ordens de compra passadas (venda) e a influência cruzada de ordens de venda (compra) sobre a taxa de chegada de ofertas de compra (venda) no book. Esta dinâmica pode ser expressa a partir da Equação (3.25), com $M = 2$, como segue abaixo.

$$\lambda_C(t) = \mu_c + \int_0^t v_{cc}(t-s)dN_c(s) + \int_0^t v_{cv}(t-s)dN_v(s), \quad (3.31a)$$

$$\lambda_V(t) = \mu_v + \int_0^t v_{vc}(t-s)dN_c(s) + \int_0^t v_{vv}(t-s)dN_v(s). \quad (3.31b)$$

No sistema de equações (3.31) os índices c e v dizem respeito a ordem de compra e venda respectivamente. As constantes μ_c e μ_v representam a intensidade de chegada de ordens a longo prazo. As funções $v_{cc}(t)$ e $v_{vv}(t)$ expressam as influências passadas das ordens de compra e vendas respectivamente. Já as funções $v_{cv}(t)$ e $v_{vc}(t)$ mostram as interações entre as chegadas de ofertas de compra e venda e vice versa. No caso de Hawkes podemos reescrever (3.31) como,

$$\lambda_C(t) = \mu_c + \sum_{t_i < t} \alpha_{cc} e^{-\beta_{cc}(t-t_i)} + \sum_{t_j < t} \alpha_{cv} e^{-\beta_{cv}(t-t_j)}, \quad (3.32a)$$

$$\lambda_V(t) = \mu_v + \sum_{t_i < t} \alpha_{vc} e^{-\beta_{vc}(t-t_i)} + \sum_{t_j < t} \alpha_{vv} e^{-\beta_{vv}(t-t_j)}. \quad (3.32b)$$

Como é possível observar, as influências tem um decaimento exponencial em relação ao tempo. Os parâmetros α_{cc} , α_{cv} , α_{vc} e α_{vv} mostram o impacto instantâneo na função intensidade quando ocorre a chegada de uma oferta de acordo com a natureza da ordem.

Já os parâmetros β_{cc} , β_{cv} , β_{vc} e β_{vv} indicam como ocorre o decaimento da influência de chegada de oferta no tempo para cada tipo de ordem. Um análise de interesse é verificar se $\beta_{cv} = \beta_{vc}$, isto é, se os impactos passados de ordens de compra (venda) afetam as ordens de venda (compra) da mesma forma. Naturalmente outros pontos relevantes surgem como examinar se $\alpha_{cc} = \alpha_{vv}$ e $\alpha_{vc} = \alpha_{cv}$ ocorrem. Caso estes pontos se verifiquem, os agentes tanto compradores como vendedores agem de modo similar quanto aos impactos instantâneos com as chegadas de ordens ao book.

3.5 Estratégia de Operação

A configuração do Book demonstra a dinâmica dos desbalanços entre a demanda e oferta do mercado. Conseqüentemente, o Book contém informações sobre a reação dos investidores com as mudanças dos preços (Easley et al., 2008). A partir das influências das taxas de chegadas de ordens é possível obter informações sobre as direções que guiam os preços a curto prazo (Shek, 2011). Portanto, para testar esta capacidade preditiva aplicamos uma estratégia ativa, \mathcal{E} , de compra e venda de cotas do fundo ETF, do principal índice do mercado brasileiro. Em seguida, comparamos com uma estratégia passiva, \mathcal{P} , de investimento, isto é, comprar e manter o ativo em carteira.

A estratégia ativa de operação depende da relação entre as intensidades de ofertas de compra e venda. A função estratégia, $\mathcal{E}(t)$, é expressa da seguinte forma,

$$\mathcal{E}(t) = \begin{cases} 1, & \text{se } \pi(t) > \pi^b, \\ -1, & \text{se } \pi(t) < \pi^s, \\ 0, & \text{se } \pi^s \leq \pi(t) \leq \pi^b \end{cases} \quad (3.33)$$

enquanto que a função estratégia $\mathcal{P}(t)$ é dada por,

$$\mathcal{P}(t) = \begin{cases} 1, & \text{se } 0 \leq t \leq T, \\ 0, & \text{se } t > T. \end{cases} \quad (3.34)$$

Os valores, 1 e -1 , representam uma posição comprada (*long position*) e vendida (*short position*) respectivamente. As operações *short position* são também conhecidas como aluguéis de ações. O valor 0 ocorre quando o ativo não estiver em carteira e T é o tamanho da série. A função $\pi(t) = \frac{\lambda_C(t)}{\lambda_V(t)}$ indica a relação entre as taxas de compra e venda. Caso ela cresça (decreça) significa que as chegadas de ofertas de compra no book estão maiores (menores) do que a de venda o que consequentemente tende a gerar pressões no sentido de aumento (diminuição) dos preços. Os parâmetros π^b e π^s são os limites em que o agente exerce a transação. O retorno de cada transação da estratégia ativa é dado por,

$$g(\mathcal{E}(t)) = R_{t+1}^T \mathcal{E}(t) + (1 - \mathcal{E}(t)), \quad (3.35)$$

onde R_t^T denota o retorno com base no melhor preço² no momento da transação, t , ou seja, o retorno teórico e não os retornos passados com base nos preços já realizados. Por meio da função (3.35) temos que a estratégia só produz ganhos quando $\mathcal{E}(t) = 1$, caso contrário, $R_t = 0$. Como consequência da função estratégia, $\mathcal{P}(t)$, a série de retornos da estratégia passiva são os próprios retornos efetivos, R_t^3 , no período analisado.

O retorno total é calculado pela estratégia e é dado por,

$$G(T) = \prod_{t=1}^T g(\mathcal{E}(t)) = \prod_{t=1}^T R_{t+1}^T \mathcal{E}(t) + (1 - \mathcal{E}(t)), \quad (3.36)$$

em que T é o tamanho da série e t é o indexador de tempo.

²No caso de sinal $\mathcal{E}(t) = 1$ adquire o ativo pelo melhor preço de venda em t , e no momento da venda $\mathcal{E}(t) = 0$ adquire pelo melhor preço de compra disponível no book.

³ $R_t = \frac{P_t - P_{t-1}}{P_{t-1}}$, em que P_t é o preço do ativo no tempo t .

Para estimar os valores ótimos de π^b e π^s , que maximizem o retorno total, será realizada uma simulação. Com o intuito de operacionalizar a geração destes resultados, o processo de simulação obedece os seguintes passos:

1. Estabelecer um valor aleatório para π^b e π^s .
2. Gerar os valores da função estratégia.
3. Extrair os valores dos retornos totais.

A partir da definição dos parâmetros, π^s , π^b é possível auferir os ganhos da estratégia ativa e compará-lo com a estratégia passiva, conhecida como *Buy and Hold*. Entretanto, é necessário ajustar a estratégia ativa em relação ao risco. Isto é, os ganhos da estratégia ativa, acima da passiva, podem ser gerados meramente em razão do aumento de risco e não em função das mudanças de interesses dos agentes captados pela estratégia.

Uma possível forma de comparar as duas estratégias é por meio da seguinte regressão,

$$g(\mathcal{E}(t)) = \alpha + \beta R_t + \varepsilon_t, \quad (3.37)$$

em que ε_t é o termo de erro estocástico e β e α são os parâmetros a serem estimados. O parâmetro β fornece uma medida de risco da estratégia, $\mathcal{E}(t)$. O parâmetro α é conhecido como alfa de Jensen, caso ele seja estatisticamente significativo a estratégia ativa trás ganhos consistentes de fato, caso contrário os retornos excedentes produzidos pela estratégia ativa são frutos do aumento de risco. Além disso, a estratégia está sujeita ao risco de liquidez, traduzida pelo *Bid-Ask Spread*. Em seguida, é possível executar este mesmo procedimento com a inserção de custos operacionais.

Outra possibilidade mais simplista da estratégia é considerar apenas as operações de *long positions* e desconsiderar as de *short positions*. Isto se deve ao fato de que as transações de *short position* dependem de fatores externos, pois eles possuem um Book específico de oferta de aluguel de ações. Neste caso, teremos a seguinte função

estratégia,

$$\mathcal{E}(t) = \begin{cases} 1, & \text{se } \pi(t-1) > \pi^b, \\ 1, & \text{se } \pi(t-1) > \pi^a \text{ e } \mathcal{E}(t-2) = 1 \\ 0, & \text{caso contrário.} \end{cases} \quad (3.38)$$

Com esta estratégia há o exercício de compra, pelo melhor preço de venda, quando a razão das intensidades é maior que π^b e mantém o ativo em carteira até que a razão, $\pi(t)$, torna-se menor que π^a , vendendo pelo melhor preço de compra.

Capítulo 4

Estimação do Processo de Hawkes

No capítulo anterior foi apresentado os processos de Hawkes univariado e multivariado, e como ele pode ser aplicado em finanças na modelagem do book de ofertas. Outro ponto ressaltado foi quanto ao significado que alguns de seus parâmetros trazem para a análise da relação entre os agentes no book. Neste capítulo, abordaremos a estimação dos parâmetros do processo de Hawkes tanto para o caso univariado como o multivariado. Em ambos os casos utilizaremos os estimadores de máxima verossimilhança.

4.1 Estimação Univariada

A função de verossimilhança de uma realização finita de pontos pode ser caracterizada pela probabilidade de obter um certo número de pontos no período analisado, vezes a função de densidade conjunta condicional das posições dos pontos dado seu número total, isto é,

$$L(N(A); t_1, t_2, \dots, t_{N(A)}) = P \{t_1, t_2, \dots, t_{N(A)} | N(A)\} P \{N(A)\}. \quad (4.1)$$

A função de verossimilhança será denota por L . Em muitos processos pontuais, a aplicação direta desta metodologia é intratável. Uma forma alternativa é utilizar

a função intensidade condicional para avaliar a verossimilhança (Daley e Vere-Jones, 2002).

Antes de prosseguir tomemos o caso mais simples em que a função intensidade é uma constante λ . Supondo o caso em que temos N observações no intervalo $(0, T]$ em que os pontos são t_1, t_2, \dots, t_N , a função de verossimilhança está embasada nas posições dos pontos t_i e no número total de pontos, $N((0, T))$.

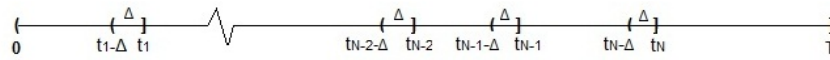


Figura 4.1: Diagrama dos tempos de ocorrências t_i dos eventos na reta

Tomando a função (3.1), podemos escrever a probabilidade de haver um único evento em cada um dos intervalos $(t_i - \Delta, t_i]$, $i = 1, 2, \dots, N$ e nenhum ponto nos intervalos restantes de $(0, T]$. Isto é dado por,

$$e^{-\lambda T} \prod_{j=1}^N \lambda \Delta. \tag{4.2}$$

Dividindo (4.2) por Δ^N e tomar $\Delta \rightarrow 0$, obtemos a densidade. A função de verossimilhança é gerada nos mesmos moldes de (4.1),

$$L_T(N; t_1, t_2, \dots, t_N) = \lambda^N e^{-\lambda T}.$$

Como a probabilidade de obter N pontos no intervalo $(0, T]$ é $\left[\frac{(\lambda T)^N}{N!} \right] e^{-\lambda T}$ o que implica dizer que $P(t_1, t_2, \dots, t_N | N((0, T))) = \frac{N!}{T^N}$, ou seja, dado certo número de pontos no intervalo, seus valores são uniformemente distribuídos em todo este intervalo. No caso de um processo de Poisson não homogêneo em que a função intensidade varia com o

tempo, $\lambda(t)$, o termo $\lambda(b_i - a_i)$ da funo (3.1) e substituída pelo *compensator*, $\Lambda(a_i, b_i]$.

$$\begin{aligned} L_T(N, t_1, t_2, \dots, t_N) &= e^{-\Lambda(0, T]} \prod_{i=1}^N \lambda(t_i) \\ &= \exp\left(-\int_0^T \lambda(t) dt + \int_0^T \log \lambda(t) dN_t\right) \end{aligned} \quad (4.3)$$

Em geral, a funo verossimilhana de um processo pontual é expressa por meio das funoes intensidades. A classe de processos pontuais em que a funo intensidade condicional pode ser expressa em termos das ocorrncias passadas, a verossimilhana torna-se análogo à (4.3) bastando substituir $\lambda(t)$ por $\lambda(t|\mathcal{F}_t)$. Muitos so os processos pontuais incluindo os estacionário e no estacionários que possuem esta característica como o processo de Poisson, renovao, processo de Wold e o de Hawkes (Daley e Vere-Jones, 2002).

Proposio 1 *Seja N um processo pontual regular em $(0, T]$ para algum $0 < T < \infty$ e sendo $t_1, t_2, \dots, t_{N(T)}$ a realizao do processo sobre o intervalo $(0, T]$ ento a funo de verossimilhana L deste processo é expressa da seguinte forma,*

$$L = \left[\prod_{i=1}^{N_T} \lambda(t_i | \mathcal{F}_t) \right] \exp\left(-\int_0^T \lambda(u | \mathcal{F}_u) du\right),$$

em que $\lambda(t|\mathcal{F}_t)$ é a funo intensidade condicional.

Para maiores informaoes veja em Daley e Vere-Jones (2002)

A partir da Proposio 1 podemos montar a funo de verossimilhana do processo de Hawkes. Tendo em mente que o método de máxima verossimilhana busca otimizar a funo verossimilhana com relao ao conjunto de parâmetros, $\Theta = \{\mu, \alpha, \beta\}$, com $P = 1$. Denotaremos a funo verossimilhana por $L_T(\Theta | N_T, t_1, t_2, \dots, t_{N(T)})$.

Para tratar o procedimento de maximizao aplicamos o log na funo verossimilhana. Esta transformao torna o método de maximizao mais fácil e no implica prejuízo pois a funo log é monótona. Denotaremos a funo $\log \{L_T(\Theta | N_T, t_1, t_2, \dots, t_{N(T)})\} =$

$\mathcal{L}_T(\Theta|N_T, t_1, \dots, t_{N_T})$. Assim, a função log verossimilhança univariada do processo de Hawkes é expressa por,

$$\mathcal{L}_T(\Theta|N_T, t_1, \dots, t_{N_T}) = -\Lambda(0, T] + \sum_{i=1}^{N_T} \ln(\lambda(t_i|\mathcal{F}_t)), \quad (4.4a)$$

$$= -\Lambda(0, T] + \sum_{i=1}^{N_T} \ln \left(\lambda_0(t) + \sum_{j=1}^P \sum_{k=1}^{i-1} \alpha_j e^{-\beta_j(t_i - t_k)} \right). \quad (4.4b)$$

Abrindo a expressão do *compensator*, $\Lambda(0, T]$, tem-se que,

$$\begin{aligned} \mathcal{L}_T(\Theta|N_T, t_1, \dots, t_{N_T}) &= - \int_0^T \lambda_0(u) du - \sum_{i=1}^{N_T} \sum_{j=1}^P \frac{\alpha_j}{\beta_j} (1 - e^{-\beta_j(t_{N_T} - t_i)}) \\ &\quad + \sum_{i=1}^{N(T)} \ln \left(\lambda_0(t) + \sum_{j=1}^P \alpha_j R_j(i) \right). \end{aligned} \quad (4.5)$$

Onde a função $R_j(i)$ é gerada recursivamente e tem a mesma forma de (3.20), dada por,

$$R_j(i) = \sum_{k=1}^{i-1} e^{-\beta(t_i - t_{i-1} + t_{i-1} - t_k)} \quad (4.6a)$$

$$= e^{-\beta_j(t_i - t_{i-1})} (1 + R_j(i-1)). \quad (4.6b)$$

Em nosso trabalho, $T = t_{N(T)}$ e quando a função $\lambda_0(t) = \mu$ for tratada como uma constante, logo a integral,

$$\int_0^T \lambda_0(u) du = \mu T.$$

No caso univariado, o esforço computacional é reduzido com as relações de recorrência, no entanto, no caso multivariado não é possível estabelecer uma função recursiva que dependa exclusivamente de um único valor passado. Na próxima seção abordaremos a estimação para o caso multivariado.

4.2 Estimaco Multivariada

Para investigarmos no somente a relao de auto excitao que um evento passado provoca nas taxas de chegadas de eventos futuros, mas tambm a excitao cruzada, onde um salto de um processo pode elevar a intensidade de um outro processo, necessitamos abordar o processo do ponto de vista multivariado. Como explicitado no Captulo 3, o processo de Hawkes multivariado de dimenso M est associado a um processo pontual multivariado com suas respectivas taxas de chegadas $\lambda_m(t|\mathcal{F}_t)$ em que $m = 1, 2, \dots, M$.

Assim como no caso univariado a log verossimilhana est estabelecida a partir das funoes intensidades. Desta forma, tomando a Proposio 1 e as funoes descritas na seo 3.4.2 do Captulo 3 a funo log verossimilhana do processo de Hawkes multivariado est expressa por,

$$\mathcal{L}_T(\Theta | \mathbf{N}_T, \mathbf{t}^1, \mathbf{t}^2, \dots, \mathbf{t}^m) = \sum_{m=1}^M \Lambda_m(0, T) + \sum_{i=1}^{N_T^m} \ln \left(\mu_m + \sum_{n=1}^M \sum_{j=1}^P \sum_{t_k^n < t_i^m} \alpha_{j,m,n} e^{-\beta_{j,m,n}(t_i^m - t_k^n)} \right). \quad (4.7)$$

O valor $\mathbf{t}^m = \{t_1^m, \dots, t_{N_T^m}^m\}_{m=1,2,\dots,M}$ representa a realizao do processo da m -sima dimenso. O vetor $\Theta = \{\Theta_1, \Theta_2, \dots, \Theta_M\}$ sintetiza os M espaos paramtrico do processo em que cada um deles so $\Theta_m = \{\mu_m, \alpha_{j,m,n}, \beta_{j,m,n}\}_{n=1,2,\dots,M}$ e $j=1,2,\dots,P$.

Proposio 2 *A funo log verossimilhana de um processo de Hawkes multivariado pode ser expressa com a soma de M funoes log verossimilhana, uma para cada conjunto de parmetros Θ_m ,*

$$\mathcal{L}_T(\Theta | \mathbf{N}_T, \mathbf{t}^1, \mathbf{t}^2, \dots, \mathbf{t}^m) = \sum_{m=1}^M \mathcal{L}_T^m(\Theta_m | \mathbf{N}_T, \mathbf{t}^1, \mathbf{t}^2, \dots, \mathbf{t}^m).$$

Por meio desta formulao est possvel otimizar cada funo individualmente [Ogata \(1978\)](#).

A partir da função (3.30) temos que,

$$\Lambda_m(0, T) = \int_0^T \lambda_m(t) dt = \sum_{i=1}^{N_T^m} \Lambda(t_{i-1}^m, t_i^m).$$

Portanto, podemos reescrever (4.7) como,

$$\begin{aligned} \mathcal{L}_T(\Theta | N_T, \mathbf{t}^1, \mathbf{t}^2, \dots, \mathbf{t}^m) &= \sum_{m=1}^M T \mu_m - \sum_{i=1}^{N_T^m} \sum_{n=1}^M \sum_{j=1}^P \frac{\alpha_{j,m,n}}{\beta_{j,m,n}} \left[\left(1 - e^{-\beta_{j,m,n}(t_i^m - t_{i-1}^m)} \right) R_j^{m,n}(i-1) \right. \\ &\quad \left. + \sum_{t_{i-1}^m \leq t_k^n < t_i^m} \left(1 - e^{-\beta_{j,m,n}(t_i^m - t_k^n)} \right) + \sum_{i=1}^{N_T^m} \ln \left(\mu_m + \sum_{n=1}^M \sum_{j=1}^P \alpha_{j,m,n} R_j^{m,n}(i) \right) \right]. \end{aligned} \quad (4.8)$$

O valor T representa o tempo total de observação da série com a função $\lambda_0(t) = \mu$ representada por uma constante. O N_T^m é o total de saltos do tipo m realizados no período T e P representa a ordem do modelo. Em nosso estudo estamos diante do caso $M = 2$ sendo $m = 1$ referente as ofertas de compras e $m = 2$ as de vendas. Como explicitado na capítulo anterior, temos um processo de Hawkes bivariado com a estrutura dada por (3.31). Neste caso, a função log verossimilhança é desmembrada em duas,

$$\mathcal{L}_T(\Theta_C, \Theta_V | N_T, \mathbf{t}^c, \mathbf{t}^v) = \mathcal{L}_T(\Theta_C | N_T, \mathbf{t}^c, \mathbf{t}^v) + \mathcal{L}_T(\Theta_V | N_T, \mathbf{t}^c, \mathbf{t}^v),$$

uma para cada conjunto de parâmetros $\Theta_C = \{\mu_c, \alpha_{cc}, \beta_{cc}, \alpha_{cv}, \beta_{cv}\}$ e $\Theta_V = \{\mu_v, \alpha_{vv}, \beta_{vv}, \alpha_{vc}, \beta_{vc}\}$.

As duas funções, portanto, são expressas respectivamente por,

$$\begin{aligned} \mathcal{L}_T(\Theta_C | N_T, \mathbf{t}^c, \mathbf{t}^v) &= - \int_{t_0^c}^{T^c} \lambda(t | \mathcal{F}_t) + \sum_{i=1}^{N_T^c} \ln(\lambda(t_i | \mathcal{F}_t)) \\ &= -\mu_c T^c - \frac{\alpha_{cc}}{\beta_{cc}} \sum_{t_k < T^c} \left(1 - e^{-\beta_{cc}(T^c - t_k^c)} \right) - \sum_{t_k < T^c} \frac{\alpha_{cv}}{\beta_{cv}} \left(1 - e^{-\beta_{cv}(T^c - t_k^v)} \right) \\ &\quad + \sum_{i=1}^{N_T^c} \ln(\mu_c + \alpha_{cc} R^{cc}(i) + \alpha_{cv} R^{cv}(i)), \end{aligned} \quad (4.9)$$

e por,

$$\begin{aligned}
\mathcal{L}_T(\Theta_V | \mathbf{N}_T, \mathbf{t}^c, \mathbf{t}^v) &= -\mu_v T^v - \sum_{i=1}^{N_T^V} \left\{ \frac{\alpha_{vv}}{\beta_{vv}} \left[\left(1 - e^{-\beta_{vv}(t_i^v - t_{i-1}^v)}\right) A^{vv}(i-1) \right] + \right. \\
&\quad \left. - \sum_{i=1}^{N_T^C} \frac{\alpha_{vc}}{\beta_{vc}} \left[\left(1 - e^{-\beta_{vc}(t_i^v - t_{i-1}^v)}\right) A^{vc}(i-1) + \sum_{t_{i-1}^v \leq t_k^c < t_i^v} \left(1 - e^{-\beta_{vc}(t_i^v - t_k^c)}\right) \right] + \right. \\
&\quad \left. \sum_{i=1}^{N_T^V} \ln(\mu_v + \alpha_{vv} R^{vv}(i) + \alpha_{vc} R^{vc}(i)) \right\}.
\end{aligned} \tag{4.10}$$

Capítulo 5

Resultados

Neste capítulo, iniciamos a análise dos dados, com base nos conceitos apresentados em capítulos anteriores. A avaliação tem como objetivo diagnosticar e caracterizar o comportamento da série de ordens do *Book* de ofertas do fundo de índice ETF mediante um processo de Hawkes. Inicialmente, descrevemos o processo de coleta dos dados e em seguida, é apresentada uma breve análise descritiva dos dados. Nas seções subsequentes exibimos os resultados do processo de estimação dos parâmetros, primeiramente para o caso univariado e posteriormente o caso multivariado. Em ambos os casos, utilizamos o algoritmo de *Nelder-Mead* para maximizar a função verossimilhança. Consideramos também, tanto a intensidade base constante como a ajustada a sazonalidade intradiária. No final, apresentamos os resultados da implementação da estratégia que leva em conta as características das intensidades de compra e venda de ativos.

5.1 Descrição dos Dados

Para investigar as taxas com que os agentes inserem as ordens no mercado e como ela se comporta ao longo do tempo, foram coletados os tempos de chegadas destas ofertas ao book. Durante o período de negociações foram captados os tempos em que houve

alterações nas melhores ofertas. Para o registro, foram consideradas os casos em que houve a inserção de uma nova melhor oferta ou quando houve uma mudança no número de ofertas sobre o mesmo preço. A contabilização destas chegadas segue-se tanto para as ofertas de compra como para as de venda. Estas taxas representam o fluxo de ordens encaminhadas por fração de tempo.

Sobre a perspectiva das taxas de chegadas de oferta de compra, foram computados tanto os tempos em que houve aumento do valor da melhor oferta de compra quanto no aumento da melhor oferta de venda, pois em ambos os casos as mudanças foram geradas por uma ordem de compra. O mesmo raciocínio é válido sobre a perspectiva das taxas de oferta de venda. Desta maneira, foram extraídas duas séries, uma para as taxas de compra e outra para as taxas de venda. Por meio desta forma de coleta não foi possível segregar as ordens a mercado *market orders* e as ordens do tipo *limits orders*.

Para coletar a amostra foi desenvolvido um programa em VBA (*Visual Basic for Applications*), que registrava a cada mudança no book o tempo em milissegundos. Os dados são fornecidos diretamente através da plataforma Enforque Ltda. O ativo em análise é a cota do fundo de índice ETF *iShare Ibovespa* cotada sobre o código BOVA11. A amostra data de 10 de Outubro de 2013 com um total de 4.048 observações para as compra e 4.945 para as observações de venda.

5.2 Resultados: Preliminares

Nesta etapa apresentamos uma análise descritivas dos dados. A série abaixo mostra a evolução dos preços realizados, com uma frequência de 1 minuto, ao longo do dia. O comportamento dos preços, em última instância, é gerado pelos contrastes dos desbalanços entre as ofertas de compra e de venda.

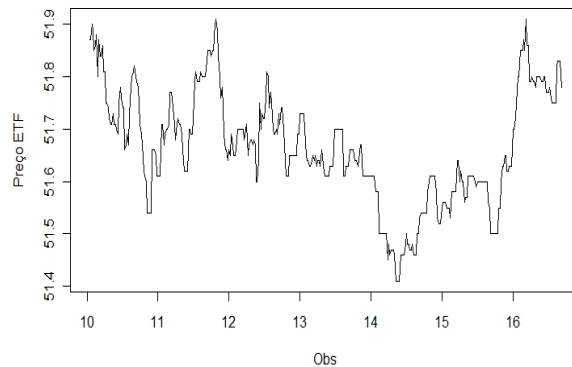


Figura 5.1: Série preços do BOVA11

Quando observamos as séries das durations, elas apresentam comportamentos similares. Isto pode ser visto por meio da tabela 5.1 de estatísticas descritivas de ambas séries.

Tipo	Média	Variância	Skewness	Kurtose
Compra	0,10	0,04	9,83	179,53
Venda	0,082	0,02	5,22	39,85

Tabela 5.1: Estatísticas Descritivas das Durations - ETF BOVA11

Como estamos diante de um processo de contagem, preliminarmente verificamos se os intervalos entre as ordens são exponencialmente distribuídos. Caso os dados se ajustem bem a esta distribuições temos um caso de um processo de Poisson homogêneo. Para investigarmos isso, esboçamos o QQ-Plot das *duration* tanto das ordens de compra como de venda.

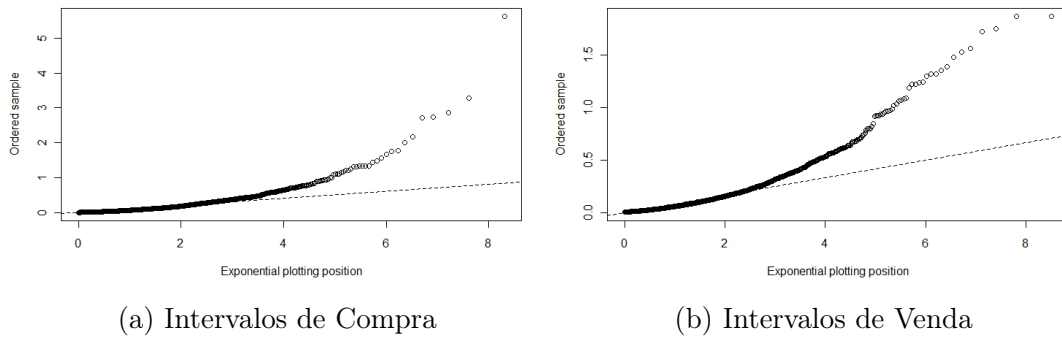


Figura 5.2: QQ-Plot *Durations* das Melhores Ofertas de Compra e Vendas do ETF

Pelo gráfico podemos observar claramente que o processo gerador dos dados não sugere seguir uma distribuição exponencial. Uma das possíveis razões pode ser devido a estrutura de dependência temporal. A seguir é plotado a série das *durations*.

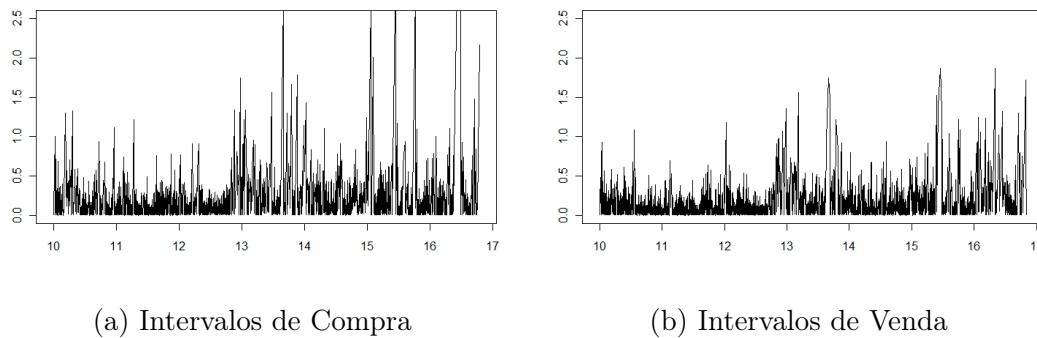


Figura 5.3: Série das *Durations* das Ofertas de Compra e Vendas do ETF

O gráfico mostra que há acúmulos de intervalos, o que sugere uma dependência entre os intervalos. Outro ponto ressaltado é o aumento das *durations* a partir das 13h em ambos os gráficos.

Os gráficos acima corroboram com a ideia de dependência temporal entre os intervalos. No entanto, não é possível verificar a dependência entre compra e venda. A partir das descrição dos dados, podemos perceber que de fato o processo gerador do número

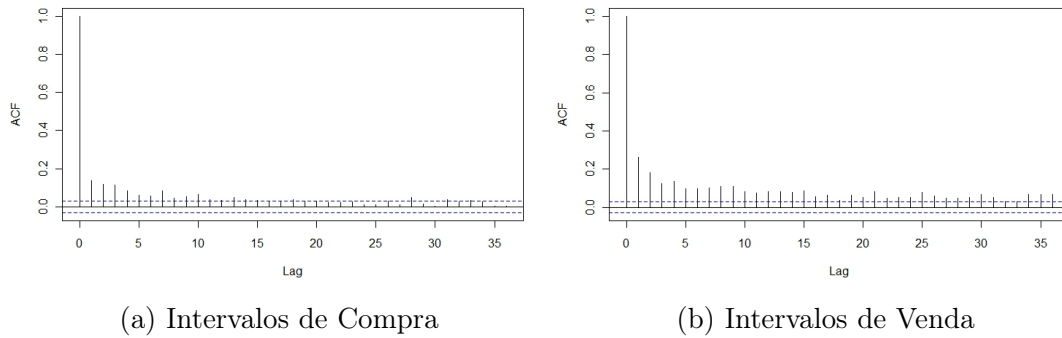


Figura 5.4: Funções de Autocorrelações da Série das *Durations* das Ofertas de Compra e Vendas do ETF

de chegadas de ofertas não segue um processo de Poisson homogêneo. O passo seguinte é verificar como os eventos passados influenciam na mudança da taxa de chegadas de ofertas iniciando com o caso do processo de Hawkes univariado.

5.3 Resultados: Processo Hawkes Univariado

Em uma análise mais simplista, vamos considerar o caso univariado. Neste ponto, apenas as chegadas de ofertas passadas de mesma natureza influencia nas taxas futuras. Como explicitado no Capítulo 4 os parâmetros foram estimados via máxima verossimilhança. A intensidade foi calculada na escala de minutos. A tabela 5.2 mostra as estimativas dos parâmetros com a intensidade base constante.

Os valores com (*) significam que a ordem é menor que 10^{-5} . A estrutura de compra e venda são muitos similares. O parâmetro α chamados de auto excitação instantânea tanto na compra quanto na venda tem o mesmo aumento da taxa de oferta. Instantaneamente podemos verificar que a taxa de oferta dobra em relação a taxa de longo prazo, μ .

O parâmetro β que mede o decaimento exponencial das influências passadas também foram similares para os dois casos. Quanto maior o valor de β menor a influência

	Coeficiente	Erro Padrão	p-Valor
Compra			
$\hat{\mu}_c$	2,36	0,200	0,000*
$\hat{\alpha}_c$	3,18	0,210	0,000*
$\hat{\beta}_c$	4,18	0,318	0,000*
Venda			
$\hat{\mu}_v$	2,99	0,250	0,000*
$\hat{\alpha}_v$	3,20	0,232	0,000*
$\hat{\beta}_v$	4,26	0,353	0,000*

Tabela 5.2: Estimativa Função Intensidade - ETF BOVA11

dos eventos passados. Os resultados dos parâmetros α e β sugerem que os agentes compradores e vendedores agem do mesmo modo.

Embora os parâmetros μ tenham se distanciado um do outro, a discrepância não foi estatisticamente significativa com p-valor de 0,08, o que mostra que as taxas de chegadas de ofertas são similares.

A partir do Teorema 1 podemos verificar o ajuste dos dados a um processo de Hawkes univariado. Caso a taxa de chegadas sejam governadas por este processo os *F-Compensators* devem seguir uma distribuição exponencial. Por meio do gráfico de QQ-Plot podemos ter uma ideia do ajuste.

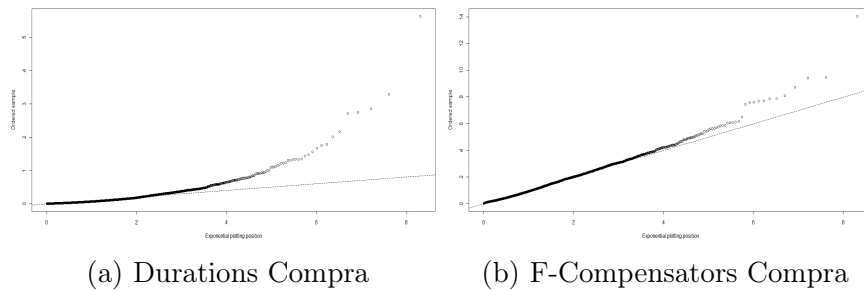


Figura 5.5: Gráfico QQ-Plot da Série das *Durations* das Ofertas de Compra do ETF

Nos gráficos 5.5 e 5.6 são dispostos no lado esquerdo as durations das chegadas de ofertas e no lado direito os *F-Compensators*. Em ambos os gráficos é possível perceber

uma mudança significativa nos ajustes a uma distribuição exponencial. De forma geral,

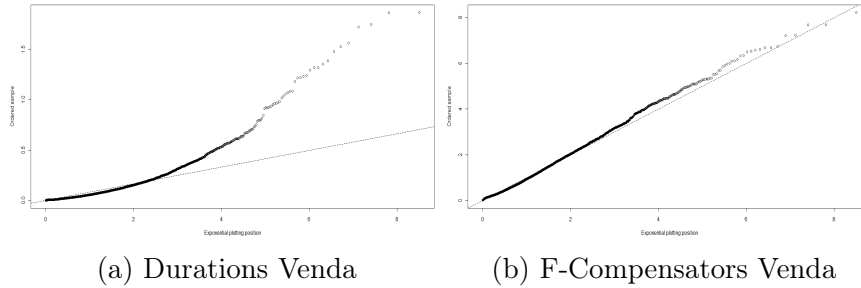


Figura 5.6: Gráfico QQ-Plot da Série das *Durations* das Ofertas de Venda do ETF

podemos observar que o processo pode ser bem descrito por um processo de Hawkes. A intensidade média, ao longo do período de negociação, foram calculadas a partir da Equação 3.18 e estão dispostas na tabela a baixo. Entretanto, é conhecido que ao longo

	Compra	Venda
$E[\lambda(t)]$	9,8648	12,0164

Tabela 5.3: Intensidade Média por Minuto - ETF BOVA11

do período de negociação a intensidade de negociações sofrem uma sazonalidade, em que há, em geral, picos no início e no final do pregão com uma acentuada diminuição de atividade no intervalo do almoço, conhecido com "efeito almoço". Através dos gráficos 5.7 de $\phi(t)$ que aproxima $\lambda_0(t)$ podemos verificar que este comportamento também ocorre, seguindo um formato de "U".

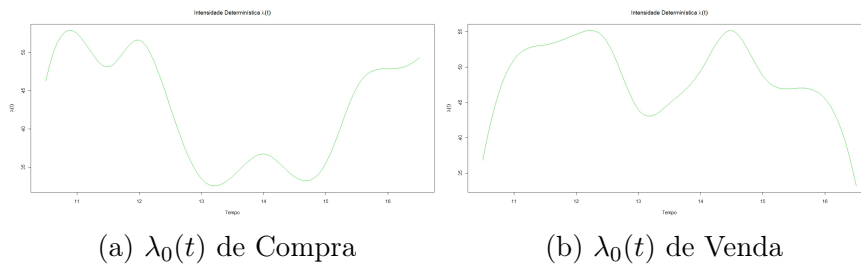


Figura 5.7: Gráfico da interpolação $\phi(t)$ - ETF BOVA11

Os gráficos foram gerados por meio de um *spline* cúbico, em que os pontos foram calculados por 3.24 espaçados a cada meia hora. Mesmo levando em conta este comportamento, os parâmetros não tiveram grandes alterações. A estimação dos parâmetros são apresentados na tabela 5.4.

	Coeficiente	Erro Padrão	p-Valor
Compra			
$\hat{\kappa}_c$	0,58	0,005	0,000*
$\hat{\alpha}_c$	3,08	0,219	0,000*
$\hat{\beta}_c$	4,12	0,33	0,000*
Venda			
$\hat{\kappa}_v$	0,07	0,005	0,000*
$\hat{\alpha}_v$	3,17	0,238	0,000*
$\hat{\beta}_v$	4,36	0,378	0,000*

Tabela 5.4: Estimativa Função Intensidade com Ajuste Sazonal- ETF BOVA11

Na medida em que o tempo passa a função intensidade de chegadas de ofertas vão sendo alteradas e sofrem aumentos instantâneos com a ocorrência de um evento e decaem exponencialmente até a chegada de um novo evento. O gráfico abaixo mostra as taxas $\lambda(t)$ ao longo dos 30 primeiros minutos de negociação.

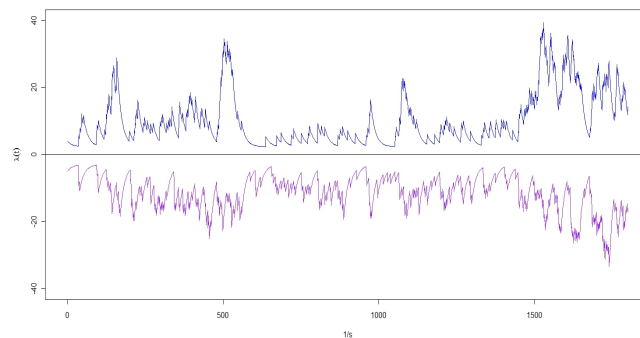


Figura 5.8: Intensidade $\lambda(t)$ de Compra (Azul) e Venda (Lilás) do ETF BOVA11

Fica evidente através do gráfico 5.8 que não somente a intensidade $\lambda(t)$ tem uma dependência temporal bem como deve possuir uma relação entre os tipos de ordens. Isto

pode ser visto na sincronia nos valores de $\lambda_c(t)$ e $\lambda_v(t)$ plotado no gráfico. Para investigarmos esta relação lançaremos mão do processo de Hawkes multivariado apresentado na etapa seguinte.

5.4 Resultados: Processo Hawkes Multivariado

A abordagem multivariada do processo de Hawkes, no contexto da modelagem do book, será a bivariada. A estrutura de relação entre as taxas de chegadas de ofertas de compra e venda que investigamos, segue o sistema de Equações 3.31. Diferentemente do caso univariado, esta configuração permite visualizar a interação entre os tipos de ofertas explicitado pelo conjunto de parâmetros $\alpha_{cv}, \beta_{cv}, \alpha_{vc}, \beta_{vc}$. Os parâmetros da Equação 3.31 foram estimados por máxima verossimilhança por meio da função 4.9. Os resultados da estimação estão apresentados na tabela a seguir.

	Coefficiente	Erro Padrão	p-Valor
Compra			
$\hat{\mu}_c$	1,49	0,28	0,000*
$\hat{\alpha}_{cc}$	3,36	0,25	0,000*
$\hat{\beta}_{cc}$	5,02	0,46	0,000*
$\hat{\alpha}_{cv}$	0,14	0,06	0,0125
$\hat{\beta}_{cv}$	0,92	0,41	0,0121
Venda			
$\hat{\mu}_v$	2,52	0,44	0,000*
$\hat{\alpha}_{vv}$	3,60	0,32	0,000*
$\hat{\beta}_{vv}$	6,20	0,69	0,000*
$\hat{\alpha}_{vc}$	0,18	0,06	0,00137
$\hat{\beta}_{vc}$	0,60	0,20	0,00175

Tabela 5.5: Estimativa Função Intensidade - ETF BOVA11

Os valores com (*) significa que a ordem é menor que 10^{-5} . Todos os parâmetros foram estatisticamente significativos a 95% de confiança. Da mesma forma que no caso univariado, os parâmetros foram similares para ambos casos. Este fato reforça a tese que os agentes atuam do mesmo modo tanto na execução da venda quanto na compra,

exposta no caso univariado.

Podemos perceber que as chegadas de ofertas induzem, de fato, a taxa de ofertas do outro lado do book indicada pelos parâmetros $\alpha_{cv}, \beta_{cv}, \alpha_{vc}, \beta_{vc}$. Outro ponto importante ressaltado pela tabela 5.5, é que embora o aumento instantâneo da intensidade seja menor quando ocorre a chegada de uma oferta de natureza inversa, sua influência perdura durante um maior tempo pois $\beta_{cv} < \beta_{cc}$ e $\beta_{vc} < \beta_{vv}$.

O gráfico abaixo mostra a função intensidade nos primeiros 30 minutos de negociação.

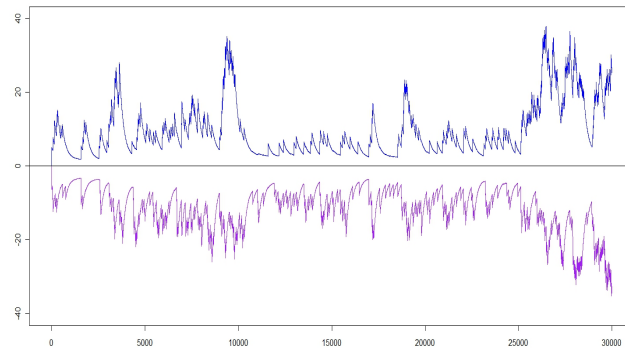


Figura 5.9: Intensidade $\lambda(t)$ de Compra (Azul) e Venda (Lilás) do ETF BOVA11

Novamente utilizando o Teorema 1 podemos verificar a qualidade do ajuste da série. Caso a realização seja governada por um processo de Hawkes bivariado os $\Lambda(t_{i-1}, t_i)$ devem possuir uma distribuição exponencial.

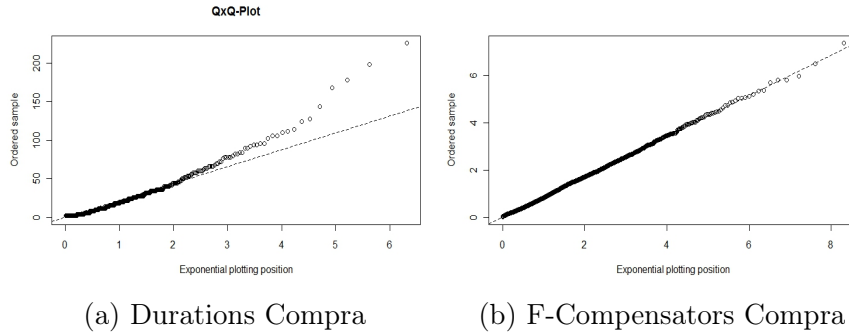


Figura 5.10: Gráfico QQ-Plot da Série das *Durations* das Ofertas de Compra do ETF

A partir dos gráficos de QQ-Plot descritos a baixo podemos averiguar que houve um bom ajuste a uma distribuição exponencial em ambos os casos.

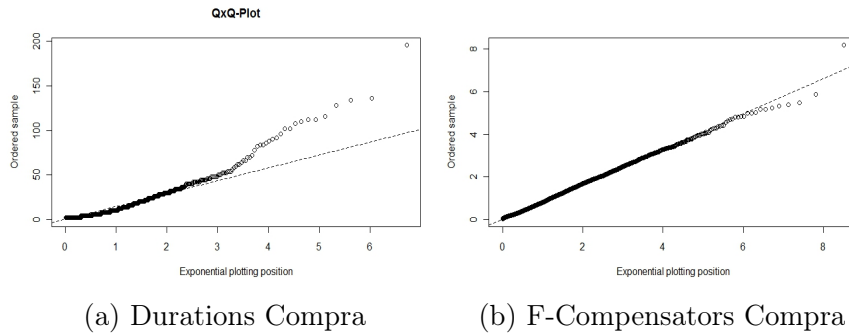


Figura 5.11: Gráfico QQ-Plot da Série das *Durations* das Ofertas de Venda do ETF

Para verificarmos a hipótese nula de $\beta_{cv} = \beta_{vc}$, isto é, se a propagação das influências são iguais tanto na venda quando na compra utilizamos o teste de razão de verossimilhança.

$$\Psi = -2 \log \left(\frac{\sup \mathcal{L}_T(\Theta_0 | \mathbf{N}_T, \mathbf{t}^c, \mathbf{t}^v)}{\sup \mathcal{L}_T(\Theta | \mathbf{N}_T, \mathbf{t}^c, \mathbf{t}^v)} \right) \sim \chi^2_{10-8}$$

Como $\Psi = 476,96 > \chi^2_{0,95;gl=2}$ logo temos evidência pra rejeitar a hipótese nula de que tanto compradores quanto vendedores de ativos possuem a mesma influência de eventos

passados gerados pelas ações do outro lado do book.

Quando consideramos a sazonalidade da função intensidade base, dispostas no gráfico 5.7, os resultados da estimação do modelo se alteram.

	Coeficiente	Erro Padrão	p-Valor
Compra			
$\hat{\kappa}_c$	0,04	0,004	0,000*
$\hat{\alpha}_{cc}$	2,93	0,24	0,000*
$\hat{\beta}_{cc}$	4,14	0,39	0,000*
$\hat{\alpha}_{cv}$	0,60	0,26	0,0112
$\hat{\beta}_{cv}$	7,7	3,52	0,0143
Venda			
$\hat{\kappa}_v$	0,04	0,009	0,000*
$\hat{\alpha}_{vv}$	0,74	0,097	0,000*
$\hat{\beta}_{vv}$	1,26	0,176	0,000*
$\hat{\alpha}_{vc}$	3,71	0,653	0,000*
$\hat{\beta}_{vc}$	11,56	1,95	0,000*

Tabela 5.6: Estimativa Função Intensidade - ETF BOVA11

Todos os parâmetros foram significativo a 95% de confiança. Quando comparamos os resultados da tabela 5.5 e da tabela 5.6 podemos perceber que a estrutura geral de reação dos agentes aos eventos do book muda. Os compradores se mostraram menos reativos a eventos instantâneos do outro lado do book quando comparados aos vendedores, comportamento traduzido pela relação $\alpha_{cv} = 0,6 < \alpha_{vc} = 3,71$. Entretanto, o decaimento exponencial da propagação da influencia dos compradores é menor que os vendedores. Para termos uma ideia desta relação podemos usar a fórmula,

$$\tau_c = \frac{\ln\left(\frac{\alpha_{cv}}{\alpha_{vc}}\right)}{\beta_{vc} - \beta_{cv}} \times 60 \quad \tau_s = \frac{\ln\left(\frac{\alpha_{cc}}{\alpha_{vv}}\right)}{\beta_{cc} - \beta_{vv}} \times 60 \quad (5.1)$$

que indica em quantos segundos a influência dos dois eventos se equivalem, isto é, em qual ponto no tempo as curvas do decaimento exponencial se igualam. No caso em que olhamos apenas para os eventos passados da mesma natureza, o $\tau_s = 28,3$, o que significa que os compradores reagem mais fortemente aos eventos apenas nos primeiros

28,3 segundos após este período as influências são maiores para os vendedores. O mesmo raciocínio é válido quando analisamos os eventos de natureza inversa. Como o $\tau_c = 28,1$, os vendedores reagem mais as ofertas de compra do que os compradores as de venda por aproximadamente 28,1 segundos.

A partir destes resultados podemos entender melhor como os investidores agem diante das chegadas de ofertas no book. Outro ponto é a investigação das similaridades de comportamento entre compradores e vendedores. Com o entendimento da reação dos agentes diante dos eventos de chegadas de ofertas podemos também verificar se a estratégia que leve em conta estas reações pode gerar ganhos superiores a média de mercado. Na próxima seção investigaremos a aplicação deste tipo de estratégia e qual é o seu desempenho diante dos riscos.

5.5 Simulação: Estratégia de Operação

O preço dos ativos são formulados mediante aos desbalanços entre as ofertas de compra e venda registrados no book de ofertas. Como as taxas de chegadas captam as mudanças com que os agentes ofertam e demandam ativos logo podemos utilizar destes movimentos para operar no mercado, como discutido na Seção 3.5. Após estimados os parâmetros que governam a reação dos agentes perante os eventos de chegadas destas ofertas estamos aptos a testar se a estratégia ativa de operação descrita pela função estratégia (3.38) produzem ganhos consistentes. Caso a estratégia ativa gere retornos anormais temos de verificar se superam os riscos das operações.

Todavia, ao realizar a simulação descrita na seção 3.5 não foi possível encontrar os pontos ótimos de π^b e π^s . No caso do ativo BOVA11 todos os pontos gerados levaram a retornos totais menores que um. Em quase toda sua totalidade as operações geradas pela função 3.38 de compra e venda para cada par de pontos π^b e π^s levaram à perdas.

A primeira vista os resultados nos leva a crer que o uso das razões entre as taxas

não possui qualquer poder preditivo do movimento dos preços. Relembrando a seção 3.5, a função estratégia esta sujeita ao risco de liquidez expresso pelo *Bid-Ask Spread*. Desta forma, para gerar retorno sobressalentes não basta ter capacidade preditiva, mas a indicação de movimentação entre compradores e vendedores deve superar a diferença entre a melhor oferta de compra e venda.

Na tentativa de filtrar a capacidade preditiva das mudanças das taxas e o risco de liquidez traçamos um *grid*. No primeiro eixo temos o π^b na escala de desvio padrão de $\lambda_C(t)$. O eixo "tempo" representa o instante δ após observar $\lambda_C(t) > \pi^b$. Caso os melhores preços subam 100% das vezes depois de δ segundos após ter ultrapassado o ponto π^b , temos uma assertividade de 1.

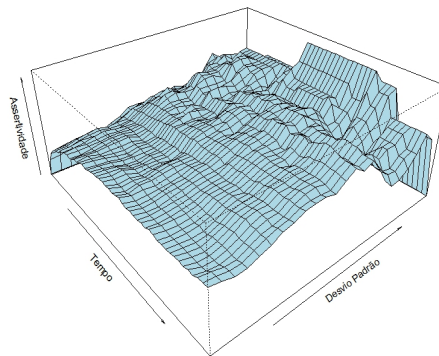


Figura 5.12: Gráfico do Nível de Assertividade

Por meio deste gráfico podemos perceber a partir de qual valor da taxa de compra e quanto tempo demora até o preço subir.

Max	π^b	δ
85%	4,55	156

Tabela 5.7: Resultado da Assertividade Máxima

O melhor ponto ocorre quando a taxa de compra excede 4,55 desvios padrões e

observamos a oferta de melhor compra após 156 segundos. Com estes dois parâmetros acertamos no aumento do valor ofertado 85% das vezes. Logo podemos perceber que as taxas de compra permite algum grau de previsibilidade porém o *bid-ask spread* impede a obtenção de ganhos no caso do ativo BOVA11. Possivelmente, com outros ativos com menores *spreads* seja possível gerar ganhos anormais.

Capítulo 6

Considerações Finais

A presente monografia investigou as taxas de chegadas de ofertas de compra e venda de ativos no mercado acionário brasileiro. Este estudo foi feito por meio da modelagem com a utilização de uma classe de processos pontuais, chamados de *Linear Self-Exciting*, em especial o processo de Hawkes. A partir dos parâmetros do processo que governam os dados, é possível descrever alguns comportamentos dos agentes tanto compradores quanto vendedores de ações ao longo do pregão evidenciados no Book de ofertas.

A apresentação da monografia iniciou-se com a introdução de alguns conceitos básicos que alicerçam este trabalho. No capítulo seguinte, foi introduzido os processos pontuais bem como um caso particular destes processos, o de Hawkes. As características do processo de Hawkes foram apresentadas mediante sua modelagem no Book de ofertas. Abordamos o caso usual de Hawkes, com kernel exponencial. No contexto da dinâmica do Book discutimos também o caso multivariado do processo de Hawkes, pois, a partir dele é possível mensurar o grau de interação entre compradores e vendedores. Outro ponto discutido, foi a descrição de uma estratégia de operação que capte a estrutura de interação no Book, a fim de adquirir ganhos acima da média do mercado. No capítulo 4 foi abordado o método de estimação dos parâmetros do processo de Hawkes.

A etapa de análise de todo resultados da monografia compreendeu apenas um único

ativo, a saber o fundo ETF *iShare Ibovespa*¹. Nesta etapa preliminar objetivamos verificar se as taxas de chegadas de ofertas são independentes e se seguem uma distribuição exponencial. Caso este quadro fosse observado teríamos um processo de Poisson do tipo homogêneo. Nesta circunstância, a descrição de um processo de Hawkes não seria útil. No entanto, os dados não apontam para isto, ou seja, os intervalos entre as chegadas das ofertas tanto de compra como as de venda não se ajustaram a uma distribuição exponencial. Nas estimações dos parâmetros do processo de Hawkes, tanto no caso univariado como no multivariado, os agente se mostraram com comportamentos similares quando atuam como compradores ou como vendedores de ativos. A busca de calibrar uma estratégia por meio das taxas de ofertas de compra e venda que gerasse retornos anormais foi frustrada. Os resultados mostraram não ser possível obter ganhos com esta estratégia. Entretanto, estes resultados foram atribuídos ao tamanho do *bid-ask spread*.

Em estudos futuros é possível estender os resultados para outros ativos do mercado acionário brasileiro. Outro ponto interessante é a investigação da adequação de kernels de outras distribuições a dados financeiros como a Weibull, gama, Burr, Hiperbólica, Pareto Generalizada entre outras. De posse destes resultados, é possível desenvolver alternativas aos modelos de avaliação de probabilidade de negociação com informação privilegiada, como [Easley et al. \(1996\)](#), que considera a taxa de chegadas de ofertas constantes.

¹Ativo registrado com o código BOVA11 na Bovespa

Referências Bibliográficas

Assaf Neto, A. (2009). *Mercado Financeiro*. Atlas.

Clark, P. K. (1973). A subordinated stochastic process model with finite variance for speculative prices. *Econometrica*, 41:135–155.

Cox, D. R. e Ishan, V. (1980). *Point Process*. CHAPMAN HALL/CRC.

Daley, D. J. e Vere-Jones, D. (2002). *An Introduction to the Theory of Point Processes: Volume I*. Springer.

Darolles, S., Gourioux, C., e Jasiak, J. (2001). Compound autoregressive models. *CREST and York University working paper*.

Easley, D., Engle, R. F., O'Hara, M., e Wu, L. (2008). Time-varying arrival rates of informed and uninformed trades. *Journal of Financial Econometrics*, 6:171–207.

Easley, D., Kiefer, N. M., OHara, M., e Paperman, J. B. (1996). Liquidity, information and infrequently traded stocks. *The Journal of Finance*, 51:1405–1436.

Engle, R. e Russel, J. (1998). Autoregressive conditional duration: A new model for irregularly spaced transaction data. *Econometrica*, 66:1127–1162.

Fama, E. F. (1970). Efficient capital markets: a review of theory and empirical work. *The Journal of Finance*, 25:383–417.

- Gourioux, C. e Jasiak, J. (2003). Autoregressive gamma models. *CREST and York University working paper*.
- Gro-Klumann, A. e Hautsch, N. (2011). Predicting bid-ask spreads using long memory autoregressive conditional poisson models. *working Paper*.
- Hawkes, A. (1971). Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, 58:83 – 90.
- Hewlett, P. (2006). Clustering of order arrivals, price impact and trade path optimisation. *working Paper*.
- Hoel, P. G., Port, S. C., e Stone, C. J. (2004). *Introduction to Probability Theory*. Houghton Mifflin Company.
- Jondeau, E., Poon, S., e Rockinger, M. (2007). *Financial Modeling Under Non-Gaussian Distribution*. Springer.
- Jones, C. M., Kaul, G., e Lipson, M. L. (1994). Transaction, volume and volatility. *Review of Financial Studies*, 65:631–651.
- Klebaner, F. C. (2004). *Introduction to Stochastic Calculus With Applications*. Imperial Colege Press.
- Lamoureux, C. G. e Lastrapes, W. D. (1990). Heteroscedasticity in stock return data: Volume versus garch effects. *Journal of Finance*, 45:221–230.
- Lawler, G. F. (2006). *Introduction to Stochastic Process*. CHAPMAN HALL/CRC.
- Ogata, Y. (1978). The asymptotic behaviour of maximum likelihood estimators for stationary point process. *Annals of the Institute of Statistical Mathematics*, 30:243–261.

Schwartz, R. A. (2010). *Micro Markets: A market Structure Approach to Microeconomic Analysis*. Wiley Finance.

Shek, H. (2011). Modeling high frequency market order dynamics using self-excited point process. *Working Paper - Disponl em SSRN 1668160*.

Apêndice A

Primeiro apêndice

Para $m < n$ tem-se que a esperança

$$\mathbb{E} [\mathbb{E} [Y|\mathcal{F}_n] | \mathcal{F}_m] = \mathbb{E} [Y|\mathcal{F}_m]$$

Inicialmente expressando o lado esquerdo da equação a cima temos

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \mathbb{E} [Y|\mathcal{F}_n] \frac{f(x_1, x_2, \dots, x_n, y)}{f(x_1, x_2, \dots, x_m)} dy dx_n dx_{n-1}, \dots, dx_1$$
$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \left[\frac{\int_{-\infty}^{\infty} z f(x_1, x_2, \dots, x_n, z) dz}{\int_{-\infty}^{\infty} f(x_1, x_2, \dots, x_n, z) dz} \right] \frac{f(x_1, x_2, \dots, x_n, y)}{f(x_1, x_2, \dots, x_m)} dy dx_n dx_{n-1}, \dots, dx_1$$

Integrando a função $f(x_1, x_2, \dots, x_n, y)$ com relação a y , a expressão fica

$$\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \left[\frac{\int_{-\infty}^{\infty} z f(x_1, x_2, \dots, x_n, z) dz}{\int_{-\infty}^{\infty} f(x_1, x_2, \dots, x_n, z) dz} \right] \frac{f(x_1, x_2, \dots, x_n)}{f(x_1, x_2, \dots, x_m)} dx_n dx_{n-1}, \dots, dx_1$$

O denominador da esperança $\mathbb{E} [Y|\mathcal{F}_n]$ se cancela com a função $f(x_1, x_2, \dots, x_n)$

$$\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z f(x_1, x_2, \dots, x_n, z) dz \frac{1}{f(x_1, x_2, \dots, x_n)} \frac{f(x_1, x_2, \dots, x_n)}{f(x_1, x_2, \dots, x_m)} dx_n dx_{n-1}, \dots, dx_1$$

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z \frac{f(x_1, x_2, \dots, x_n, z) dz}{f(x_1, x_2, \dots, x_m)} d_{x_n} d_{x_{n-1}} \cdots d_{x_1} = \mathbb{E}[Y | \mathcal{F}_m]$$

Assim, concluímos a igualdade $\mathbb{E}[\mathbb{E}[Y | \mathcal{F}_n] | \mathcal{F}_m] = \mathbb{E}[Y | \mathcal{F}_m]$.

Apêndice B

Segundo apêndice

A série 3.4 representa a probabilidade acumulada de um distribuição de Poisson, logo temos que esta série converge. Desta forma, podemos diferenciar termo a termo para obter a função densidade de S_n ,

$$\begin{aligned}\frac{d}{dt}F_{S_n} &= -\sum_{j=n}^{\infty} \lambda e^{-\lambda t} \frac{(\lambda t)^j}{j!} + \sum_{j=n}^{\infty} \lambda e^{-\lambda t} \frac{(\lambda t)^{j-1}}{(j-1)!}, \\ &= \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!} + \sum_{j=n+1}^{\infty} \lambda e^{-\lambda t} \frac{(\lambda t)^{j-1}}{(j-1)!} - \sum_{j=n}^{\infty} \lambda e^{-\lambda t} \frac{(\lambda t)^j}{j!}, \\ &= \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!} = \frac{\lambda^n}{\Gamma(n)} t^{n-1} e^{-\lambda t}.\end{aligned}\tag{B.1}$$