

2

Análise dos dados intradiários

2.1

Estatísticas descritivas dos dados intradiários

Nesta seção, realizaremos uma análise descritiva dos dados intradiários, os quais são compostos por volume de transação, ágio de compra e venda, preço e o tempo de duração entre os negócios.

Os dados referem-se aos papéis do *Shenzen Development Bank*, no período de 18/07/2001 a 25/01/2002, com 59.878 observações. O horário de funcionamento da bolsa chinesa é no período de 9h30min às 15h sendo, o intervalo de almoço entre 11h30min e 13h. No arquivo temos registrado o ano, dia, hora, minuto, segundo, preço da negociação, volume, preço de compra e de venda e os intervalos de tempo entre negócios. Foram excluídos do estudo os negócios referentes a primeira meia hora de negociações¹, o intervalo de almoço e os negócios registrados depois das 15h, ou seja, os negócios realizados no *after-market*².

Iniciaremos a análise dos dados intradiários plotando uma amostra de 30 minutos dos preços das negociações, Figura 2.1, onde fica evidente a irregularidade no tempo da duração.

Segundo Engel & Russel [10], os modelos econométricos tradicionais são específicos para intervalos fixos de tempo ocasionando dificuldades para a análise. Sendo assim, uma solução para este problema seria fixar os intervalos definindo uma regra de interpolação a ser usada quando não ocorre negociação ao final do intervalo ou podemos optar pelo uso de intervalos estocásticos.

Na Figura 2.2, apresentamos o histograma da variação de preço³ de negociação para o período de 7 meses. Observamos que 96% da variação do preço de negociação está entre -3 e +3 centavos, tal fato é uma consequência

¹Desta forma eliminamos negócios muito discrepantes.

²Desta forma eliminamos negócios que apresentam pouca liquidez.

³A variação de preço é definida como $P_t - P_{t-1}$.

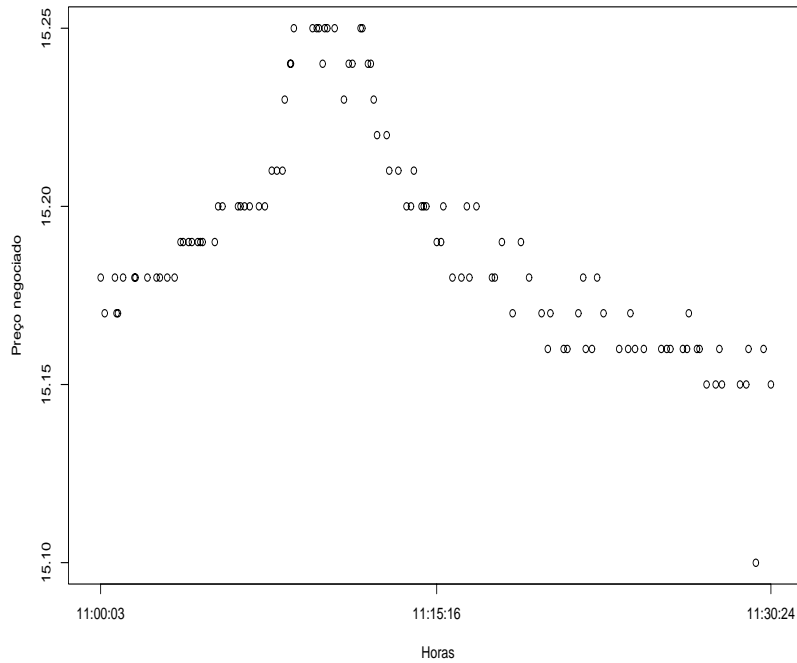


Figura 2.1: Amostra dos preços de negociação

das regras para as negociações impostas pelas Bolsas (tais regras se fazem necessárias para evitar fraudes e garantir a liquidez do mercado), ou seja, grandes variações nos preços, ou volume, não são aceitas pelos parâmetros do sistema de negociação das Bolsas. Isto se aplica também aos preços de compra e venda e ao ‘preço médio de compra e venda’⁴ que apresentarão comportamentos semelhantes conduzindo a uma curtose elevada⁵ como podemos exemplificar na tabela de estatísticas descritivas abaixo da variação do preço de negociação.

⁴O preço médio de compra e venda é definido como $\frac{P_{venda} + P_{compra}}{2}$.

⁵Tal característica é comum em dados de alta frequência como visto em Engle[8].

Descritivas	Valor
Min.	$-7.9000e - 01$
1st Qu.	$-1.0000e - 02$
Median	$0.0000e + 00$
Mean	$-4.4000e - 05$
3rd Qu.	$1.0000e - 02$
Max.	$1.0800e + 00$
Kurtosis	$4.8336e + 02$
Skewness	$1.8455e + 00$
Variance	$2.2499e - 04$

Tabela 2.1: Estatísticas descritivas da variação do preço de negociação

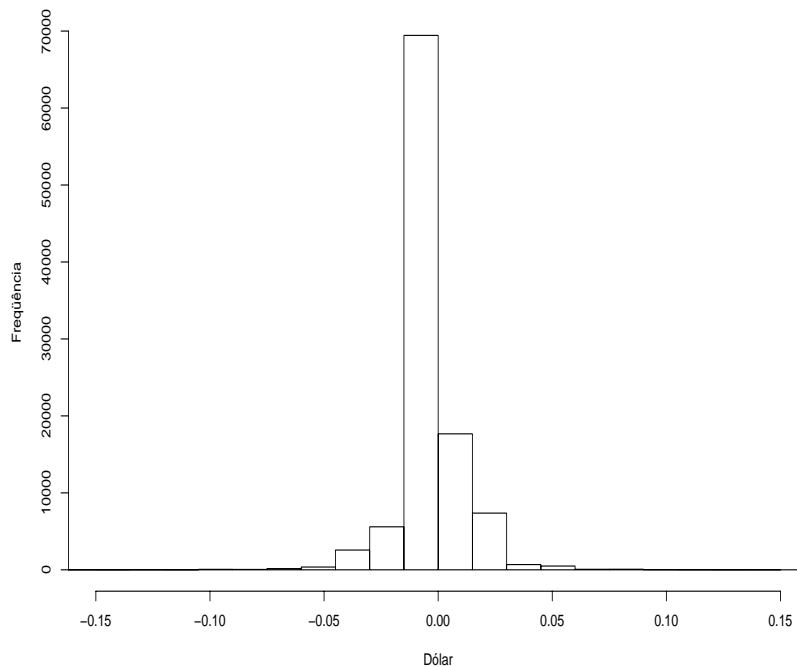


Figura 2.2: Histograma da variação de preço de negociação para as ações do Shenzhen Development Bank

Como exemplo dos parâmetros de negociações usados pelas bolsas, citamos o caso brasileiro onde os negócios diretos são submetidos a leilão quando ultrapassarem os parâmetros de mercado, os quais são estabelecidos por grupos de Ativos, de acordo com sua liquidez. A seguir, apresentamos os parâmetros da BOVESPA.

- Papéis negociados no Pregão a Viva Voz: Oscilação igual ou superior a 0,5% em relação ao último preço registrado, ou com quantidade igual ou superior a 20% da média negociada nos últimos 30 pregões.

- Papéis mais negociados no Sistema Eletrônico de Negociação e demais papéis pertencentes à carteira teórica do Ibovespa e do IBrX-100⁶: Oscilação igual ou superior a 2,0% em relação ao último preço registrado, ou com quantidade igual ou superior a 40% da média nos últimos 30 pregões.
- Demais papéis negociados no Sistema Eletrônico de Negociação: Oscilação igual ou superior a 3,0% em relação ao último preço registrado, ou com quantidade igual ou superior a 80% da média negociada nos últimos 30 pregões.

Engel & Russel [10][3] e Zhang, Russel & Tsay[9] observam um comportamento periódico em relação a frequência dos negócios, volume e a diferença entre os preços de compra e venda exibindo uma forma de U invertido ao longo do dia. A volatilidade do preço é maior na abertura e próximo ao fechamento do mercado, o volume e o *spread* apresentam o mesmo comportamento. Já o tempo entre os negócios, ou durações, tende a ser menor na abertura e próximo ao fechamento do mercado.

Na próxima seção, estimaremos a componente sazonal dos dados intradiários fazendo uso de *splines* e a seguir os dessazonaremos.

2.2

Estimação da componente sazonal dos dados de duração

A fim de estimar a componente sazonal determinística intradiária do mercado para o tempo entre os negócios, adotamos o mesmo procedimento que Zhang, Russel & Tsay[9]. Primeiramente, usando os dados de duração, onde $x_i^* = t_i - t_{i-1}$ para $i = 1, \dots, N$ onde N representa o i -ésimo negócio, obtemos uma estimativa suavizada da componente $\phi(t_{i-1}) \equiv E[X_i^* | t_{i-1}]$ usando *splines cúbico*⁷. Para o cálculo da componente sazonal realizamos o seguinte tratamento nos dados. Agregamos os dados de 30 em 30min⁸ e calculamos a média entre os intervalos de duração. Em seguida, ajustamos

⁶O IBrX - Índice Brasil é um índice de preços que mede o retorno de uma carteira composta por 100 ações selecionadas entre as mais negociadas na BOVESPA, em termos de número de negócios e de volume financeiro.

⁷A técnica de *splines* consiste em dividir valores da variável preditiva em intervalos (“janelas”) predefinidos e ajustar um polinômio (em geral cúbico) para cada intervalo, de forma que, os polinômios se juntem suavemente. A função das janelas é essencialmente considerar uma proporção do conjunto de valores para a construção da função suavizadora, pormenores vide Koopman[4]

⁸Agregando-se temporalmente os dados, a dependência na variação de preços tende a decrescer; porém, para intervalos de meia hora ou autocorrelações negativas de primeira ordem a dependência permanece. Tal como visto em Engle & Russel[10].

o *spline cúbico* e obtemos a componente sazonal de 30 em 30min. Os dados, sazonalmente ajustados, são obtidos fazendo-se $x_i = \frac{x_i^*}{\phi(t_{i-1})}$ e nos referimos a série de duração dessazonada por X_i , onde $E[X_i] = 1$. Assumimos que a componente determinística intradiária $\phi(\cdot)$ pode ser formulada como uma função multiplicativa tal que

$$E[X_i^* | X_{i-1}^*, \dots, X_1^*] = \phi(t_{i-1}) \psi(X_{i-1}, \dots, X_1),$$

onde

$$\psi(X_{i-1}, \dots, X_1) = E[X_i | X_{i-1}, \dots, X_1] \tag{2-1}$$

A Figura 2.3 mostra a componente estimada para um dia de negociação usando-se o mesmo procedimento descrito anteriormente. Observamos que a componente apresenta o mesmo comportamento que descrito por Engel & Russel[10][3] e Zhang, Russel & Tsay[9]. A Figura 2.4 mostra a componente obtida para o período de 18/07/2001 a 25/01/2002 cuja média amostral é $E[X_i] = 0,98$. Ao observar a Figura 2.4 deve-se tomar cuidado com a escala pois, para cada dia de negociação o comportamento seria igual ao Figura 2.3. Apresentamos na Figura 2.5 e Figura 2.6 as durações e durações dessazonadas, respectivamente.

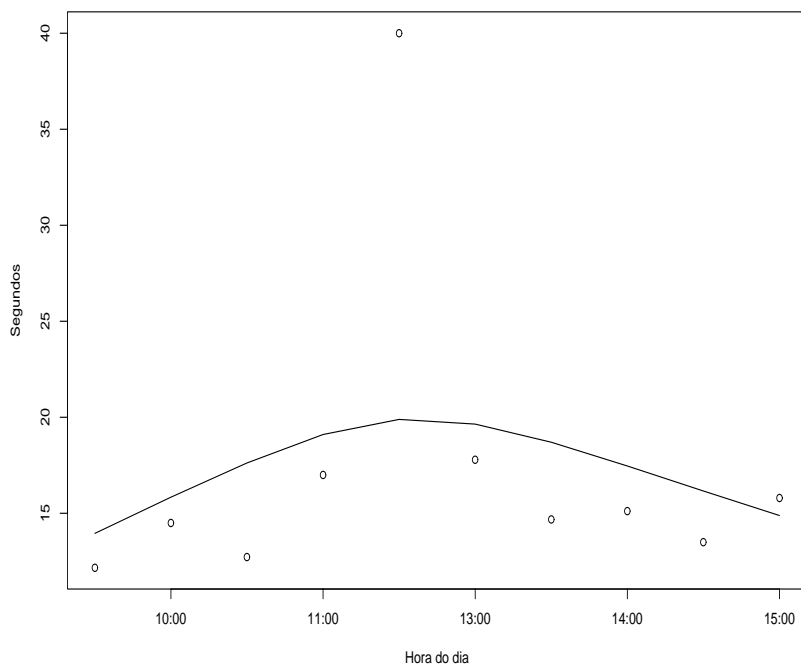


Figura 2.3: Componente sazonal determinística das durações para uma amostra de um dia

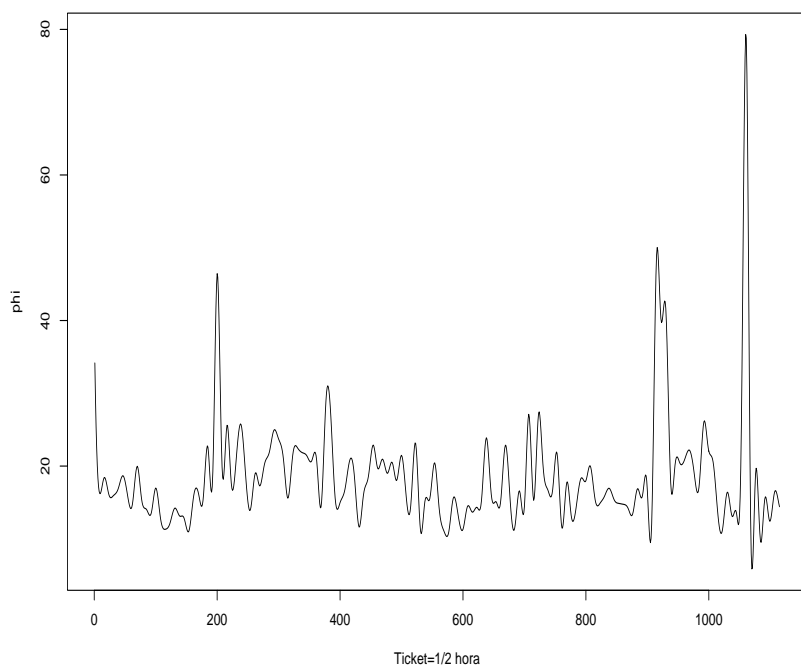


Figura 2.4: Componente sazonal determinística das durações - 18/07/2001 a 25/01/2002

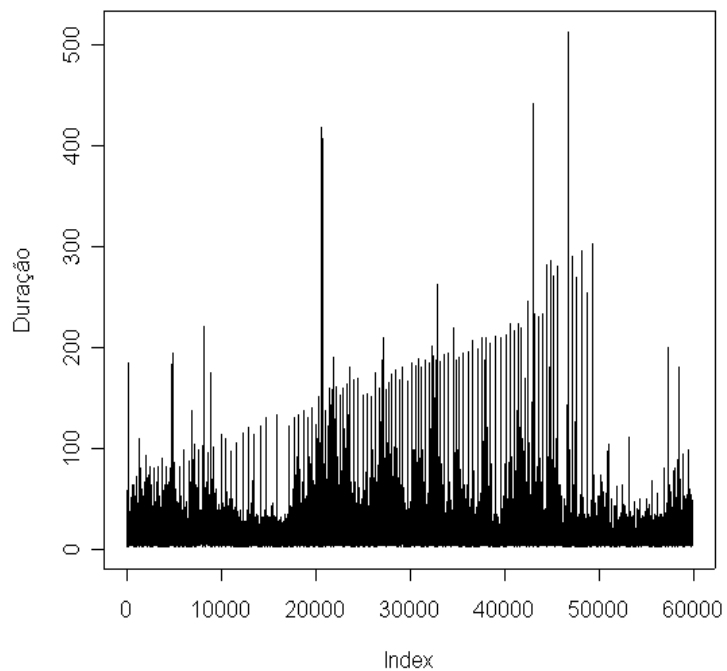


Figura 2.5: Duração

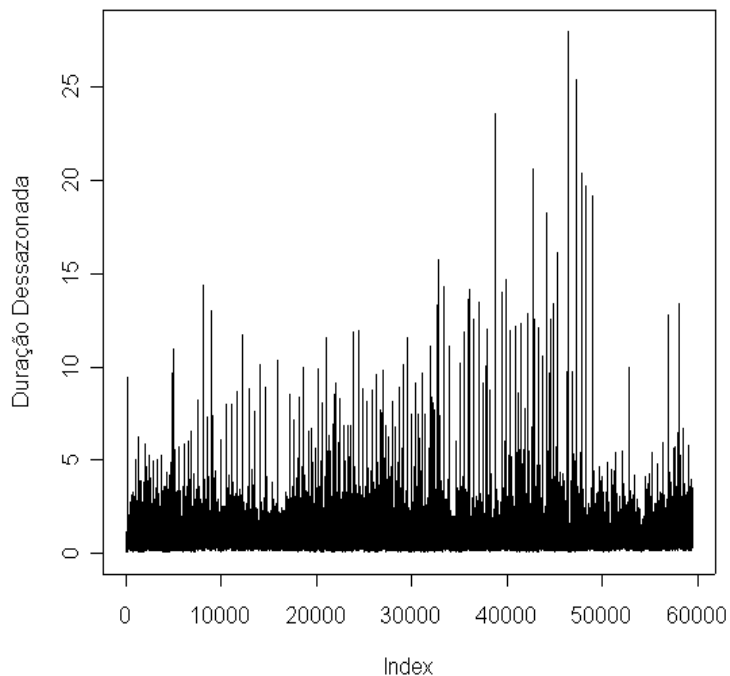


Figura 2.6: Duração dessazonada

Os dados financeiros (volume de transação, ágio de compra e venda, preço e duração) intradiários exibem, tipicamente, forte dependência. Isto é consequência da discretização dos dados e do ágio de compra e venda responsáveis por autocorrelações altas de primeira ordem. Outro fator que induz a dependência na variação de preço são as ordens de compra e venda promovidas pelos investidores para a obtenção do melhor preço. Para grandes horizontes de negociações, por vezes, temos autocorrelações positivas. As Figuras 2.7 e 2.8 apresentam as autocorrelações para as variações do preço de negociação e do ‘preço médio’ por negociação, respectivamente.

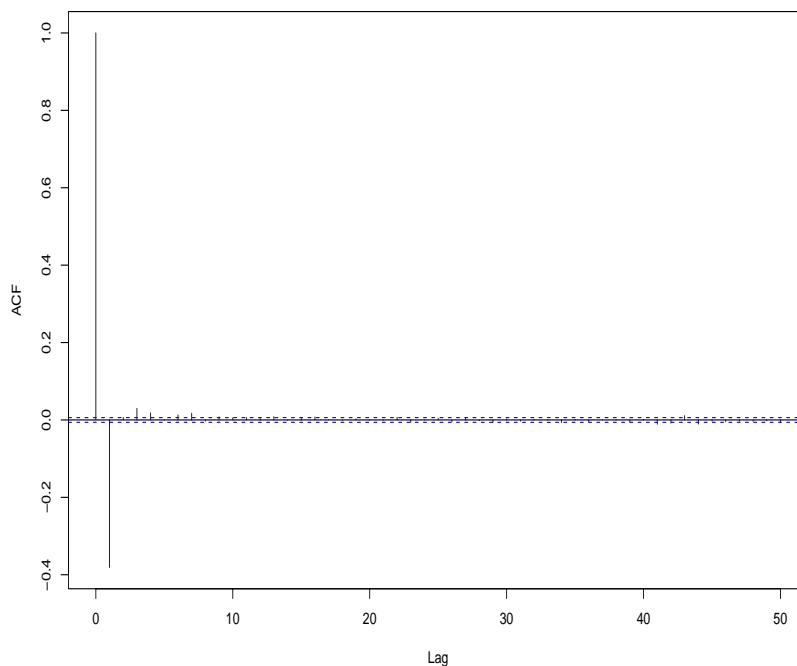


Figura 2.7: Função de autocorrelação da variação do preço de negociação dessazonado

Analogamente aos retornos inter-diários, os dados intradiários tendem a exibir conglomerados de volatilidade. Na Figura 2.9, temos a função de autocorrelação da variação absoluta do preço de negociação⁹, cujo comportamento diurno foi removido, após retirar-mos a componente sazonal, dividindo a variação de preço pelo desvio-padrão de meia em meia hora do dia, onde observamos forte dependência temporal. A variância foi estimada fazendo-se uso de *splines*.

⁹A variação absoluta do preço é definida como $|P_t - P_{t-1}|$.

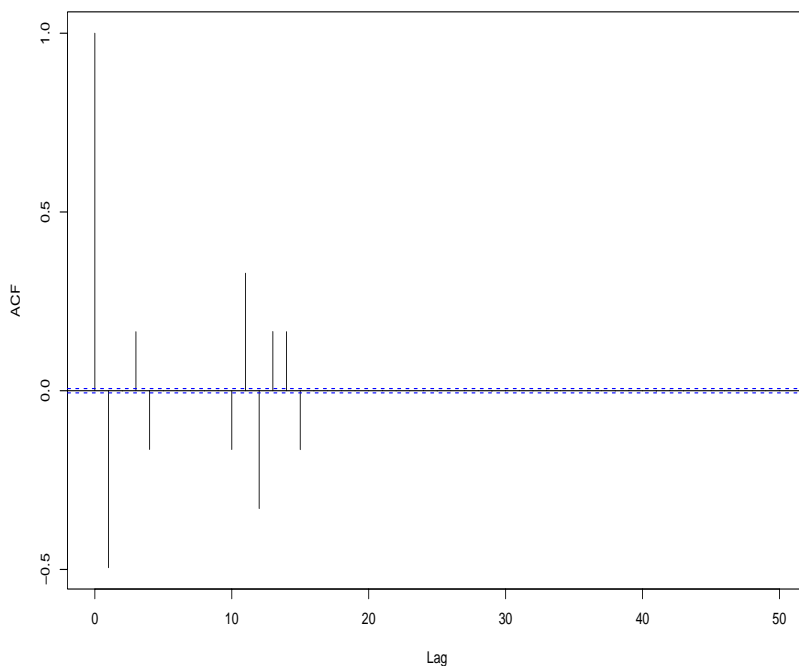


Figura 2.8: Função de autocorrelação da variação do preço médio de negociação dessazonado

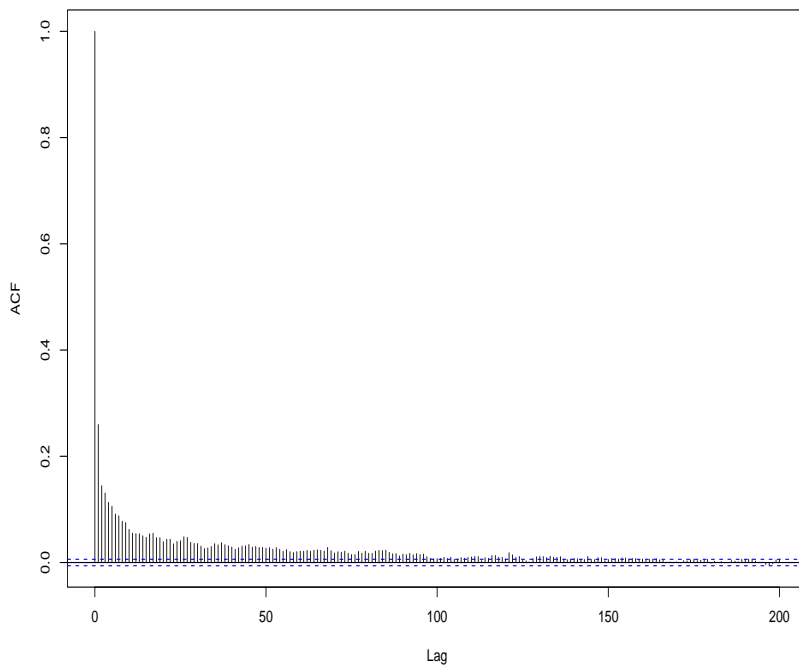


Figura 2.9: Função de autocorrelação da variação do preço de negociação absoluto dessazonado

O número de negócios exibe alta dependência temporal. A Figura 2.10 exibe as autocorrelações para as durações dessazonadas entre os negócios

e a Figura 2.11 apresenta a autocorrelação para o logaritmo do volume. O volume foi dessazonado agrupando em intervalos de 30 minutos, calculando-se a média e ajustando-se um *spline* para obtermos a componente sazonal. Em seguida, dividimos o volume referente ao intervalo de meia hora pela componente correspondente. As duas séries, duração e volume, apresentam autocorrelações elevadas e persistentes¹⁰. Temos também a indicação de conglomerados de duração e volume, respectivamente.

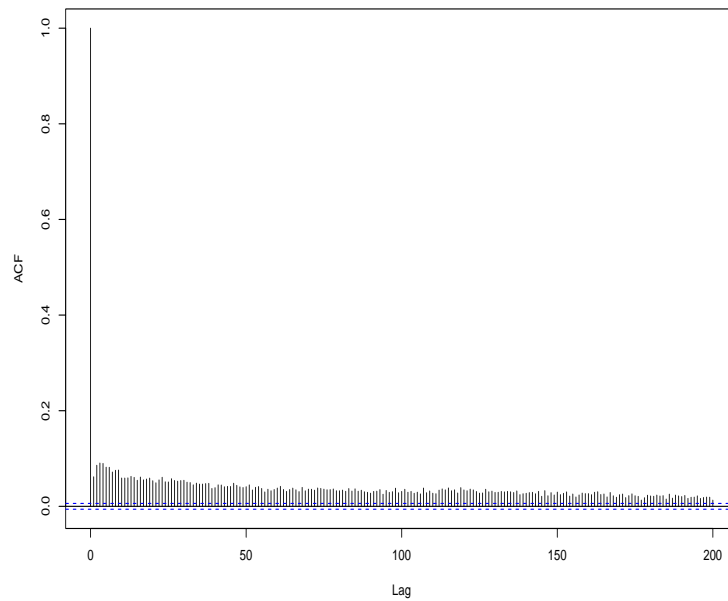


Figura 2.10: Função de autocorrelação da duração dessazonada

¹⁰O leitor pode estranhar o fato de usarmos o termo persistência onde seria mais adequado o termo memória longa. Porém, os artigos de Engle & Russel[10][3] e Engle[8] utilizam o termo persistência. Doravante usaremos o termo persistência para acordarmos com a literatura.

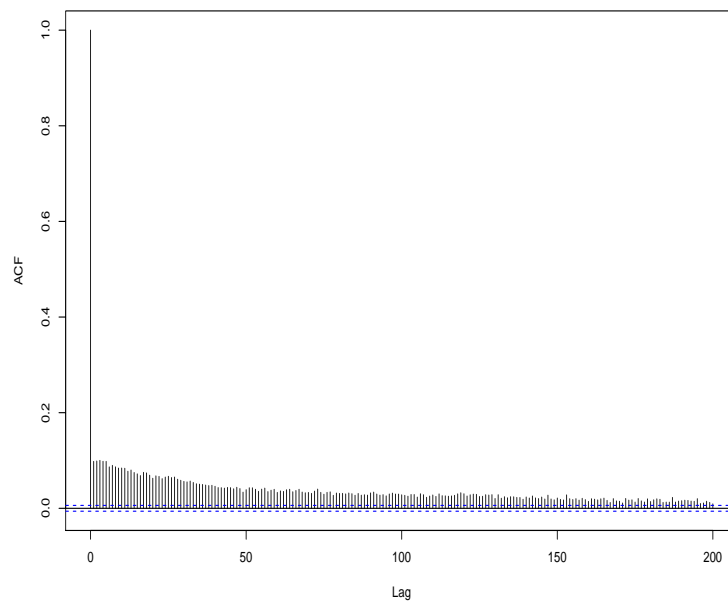


Figura 2.11: Função de autocorrelação do volume dessazonado

Assim, podemos observar que os dados intradiários possuem uma estrutura de correlação temporal e persistência. Sendo assim, podemos propor o modelo que capture esta estrutura. No próximo capítulo, apresentaremos o modelo para a previsão das durações.