

# Análise de Dados de Alta Frequência

Pedro A. Morettin

9 de Novembro de 2007

## 1 Introdução

Dados financeiros de alta frequência são aqueles obtidos em intervalos muito pequenos de tempo, usualmente em escala intra-diária e irregularmente espaçados no tempo. Algumas vezes, dados diários são também considerados de alta frequência. Para fixar nossa nomenclatura, consideraremos *dados de alta frequência* (DAF) como aqueles observados no decorrer de um dia de transações numa bolsa de valores, num mercado de taxas de câmbio etc.

No Capítulo 1 citamos duas séries de retornos intradiários: do IBOVESPA e da TELEMAR PN, ambas observadas a cada quinze minutos (veja o exemplo 1.3). Vimos, também, a forma típica como esses dados são apresentados, na página 5. Os dados brutos dessas duas séries, também chamados dados “tick-by-tick”, ocorreram em instantes aleatórios de tempo e depois foram “limpos” e amostrados a cada 15 minutos. Para detalhes sobre o tratamento prévio de dados intra-diários, veja Dacorogna et al. (2001) e Zivot (2005). Esses dados foram obtidos da Bolsa de Valores de São Paulo (BOVESPA). Outras bases de dados do mercado acionário são a TAQ (“Trades and Quotes”) da NYSE (New York Stock Exchange), AMEX e NASDAQ. Como exemplo de base de dados de opções temos a Berkeley Options Data Base, e com exemplo de base de dados de taxas de câmbio temos aquela organizada por Olsen Associates. Normalmente, essas bases de dados não são acessíveis gratuitamente.

No mercado brasileiro temos ainda relativamente poucos dados de ativos negociados em alta frequência (AF). Na BOVESPA, um número pequeno de ações têm liquidez suficiente para fornecer DAF. Tipicamente, os tempos de negociação são medidos em segundos e podemos ter várias negociações no mesmo instante de

tempo. As variações de preços de um ativo são dadas em múltiplos de um “tick”, que pode ser por exemplo um centavo de real ou um centavo de dólar.

DAF são importantes no estudo da micro-estrutura dos mercados financeiros. Nesses mercados existem grandes movimentos dos preços dos ativos em intervalos de tempo menores do que um dia. No entanto, DAF também apresentam desafios, como a filtragem dos dados, pois esses podem ser registrados com erros e têm que ser corrigidos antes que qualquer análise possa ser feita.

Para as séries do IBOVESPA e TELEMAR o número de observações de cada série era bastante grande: 37.961 para a primeira e 21.429 para a segunda, respectivamente. Para séries de taxas de câmbio esses valores são muito maiores, pois os mercados de taxas de câmbio funcionam praticamente durante 24 horas, dadas as diferenças geográficas entre eles (veja o Apêndice 10). Assim, tamanhos de séries acima de 100.000 observações são comuns.

DAF são em geral não-sincronizados. Ativos diferentes têm freqüências de negociações diferentes. Além disso, para um determinado ativo, a intensidade das negociações varia ao longo do dia, sendo maior na abertura e fechamento dos pregões e menor no horário de almoço. Isso faz com que os tempos entre as negociações (ou durações, veja a seção 10.3) apresentem um ciclo diário, na forma de “U” invertido.

Para retornos diários de ações, negociações não-sincronizadas podem introduzir (Tsay, 2005):

- (a) correlação não-nula no lag 1 para retornos de uma ação;
- (b) correlações e correlações cruzadas não-nulas no lag 1 para uma carteira de ações;
- (c) em algumas situações, correlação negativa para retornos de uma ação.

Com referência ao item (c), considere o log-retorno  $r_t$  de uma ação e suponha que  $r_t \sim \text{i.i.d.}(\mu, \sigma^2)$ . Seja  $\pi$  a probabilidade que a ação não seja negociada em qualquer instante  $t$ . Chamemos  $r_t^*$  o retorno observado no instante  $t$ . Se não houver negociação no instante  $t$ ,  $r_t^* = 0$  e se existir,  $r_t^* = r_t + r_{t-1} + \dots + r_{t-k_t}$ , onde  $k_t$  é o maior inteiro não-negativo tal que não ocorram negociações nos períodos  $t - k_t, t - k_t + 1, \dots, t - 1$ . Então (veja Tsay, 2005, para uma prova) temos que  $\text{Cov}(r_t^*, r_{t-j}^*) = -\pi^j \mu^2$ ,  $j \geq 1$ . Ou seja, quando  $\mu \neq 0$ , negociações não-sincronizadas provocam correlação negativa na série de retornos.

Esta correlação negativa pode ser causada pelo chamado “bid-ask spread”. Chamemos de  $p_{b,t}$  o “bid price” (preço de venda para o público) e  $p_{a,t}$  o “ask price” (preço de compra para o público), ambos no instante  $t$ . A diferença  $s_t = p_{a,t} - p_{b,t}$  é chamada “bid-ask spread” no instante  $t$ , e é em geral um valor pequeno (alguns ticks).

O preço de um ativo num mercado “sem fricção” usualmente é definido por  $p_t^* = (p_{a,t} + p_{b,t})/2$ . Roll (1984) sugere um modelo para o verdadeiro preço de mercado de um ativo, em função de  $p_t^*$ . Em particular, para esse modelo, a correlação de lag 1 para os retornos é negativa e função do spread  $s_t$ . Se não tivermos interesse nos efeitos relacionados a  $s_t$ , podemos trabalhar com preços médios logarítmicos, ou

seja,  $x_t = [\log(p_{a,t} \cdot p_{b,t})]/2$ , de modo que os retornos de período  $k$  serão dados por  $r_t[k] = x_t - x_{t-k}$ .

Outra característica marcante de DAF é a ocorrência de múltiplas transações no mesmo instante, mesmo com preços diferentes. Mas a grande maioria das transações não apresenta mudanças de preços.

**Exemplo 10.1.** Consideremos os dados de negociações das ações da Telemar no período de 8 a 10 de setembro de 2004, perfazendo 6.734 negociações. Foram retiradas as negociações que ocorreram fora do horário de funcionamento da bolsa, entre 10 horas e 17 horas, resultando 6.588 transações (146 negociações ocorreram no chamado período “after market”, horário em que a bolsa está fechada). Vários autores consideram que estas observações têm um padrão diferente das negociações que ocorrem durante o período normal de funcionamento da bolsa. Veja Engle (2000) e Zhang et al. (2001). Neste caso, aproximadamente 2,17% das negociações referem-se às operações feitas por via eletrônica. Veja também Stoll e Whaley (1990).

Há inúmeras negociações que ocorreram ao mesmo tempo, a saber, 2.026, representando aproximadamente 30% do total. Destas, apenas 35% apresentaram alterações de preços. Ou seja, transações simultâneas parecem não trazer grande informação sobre a micro-estrutura do mercado relevante para um tratamento estatístico. Estas negociações simultâneas são, em geral, excluídas na modelagem estatística de tempos entre negociações.

Na Tabela 10.1 selecionamos 4.096 valores selecionados entre os 4.532 que não apresentaram transações simultâneas, classificando os movimentos de preços da ação. As alterações de preços foram classificadas como aumento (+), estabilidade (0) e diminuição (-). A tabela mostra os movimentos de preços entre duas negociações consecutivas, ou seja, da  $(t - 1)$ -ésima para a  $t$ -ésima.

Tabela 10.1: Movimentação de preços em negociações consecutivas para a Telemar.

neg. $t$ neg. $t - 1$	+	0	-	Total
+	168(4,1%)	390(9,5%)	263(6,4%)	821(20,0%)
0	397(9,7%)	1.547(37,8%)	444(10,8%)	2388(58,3%)
-	255(6,3%)	451(11,0%)	181(4,4%)	887(21,7%)
Total	820(20,1%)	2.388(58,3%)	888(21,6%)	4.096(100%)

Notamos que:

- (a) aumentos ou diminuições consecutivas de preços são pequenas, representando apenas 8,5% das negociações;
- (b) em 37,8% das negociações, o preço permanece estável, havendo grande tendência de não haver variações de preço;

- (c) a porcentagem de negócios em que os preços mudaram de sentido, isto é, de queda para subida e vice-versa, é semelhante, da ordem de 6,3%, totalizando aproximadamente 13% das transações e indicando um processo de reversão de preços;
- (d) a distribuição marginal de preços é semelhante em duas negociações consecutivas.

Além dos preços de uma ação, os tempos entre negociações ou *durações* são importantes para o estudo da micro-estrutura dos mercados acionários. Chamemos de  $\tau_0, \tau_1, \dots, \tau_T$  os instantes de tempo de negociações de uma ação, sendo que  $\tau_0$  é o instante inicial e  $T$  é o número de negociações no período avaliado. Então  $x_t = \tau_t - \tau_{t-1}$ ,  $t = 1, \dots, T$  são as durações. Na Seção 10.3 trataremos de um modelo importante para as durações, o chamado modelo ACD (“autoregressive conditional duration”).

Na Figura 10.1 ilustramos a série dos 4.096 preços da Telemar considerados acima, a série de retornos, o histograma dos retornos e a respectiva f.a.c. Notamos uma correlação negativa no lag 1 da f.a.c., como discutido anteriormente. Na Figura 10.2 consideramos as durações, com o respectivo histograma, f.a.c. e box-plot. O histograma das durações mostra um decaimento exponencial a partir da origem e a f.a.c. mostra que a série de durações tem memória longa.

Outra variável de interesse é o número de negociações em dado período, por exemplo por dia, ou em cada intervalo de 5 minutos. A Figura 10.3 contém os gráficos do número de negociações em intervalos de 5 minutos para os dados da Telemar, no período de 4 de agosto a 10 de setembro de 2004 e da f.a.c. estimada para os lags de 0 a 260. A figura também mostra as durações das negociações, em intervalos de 5 minutos e respectiva f.a.c. Ambas as funções mostram uma periodicidade evidente de 84, correspondente ao número de intervalos de 5 minutos em cada dia de negociação. A Figura 10.4 traz os gráficos do número médio de negociações e do tempo médio entre as negociações em intervalos de 5 minutos, no mesmo período. Ao todo são 84 médias. O primeiro gráfico apresenta um maior número de negócios nos períodos das 10:15 às 12:00 e das 14:40 às 16:55. O segundo gráfico apresenta obviamente um padrão inverso.

A Tabela 10.2 apresenta algumas estatísticas das negociações da Telemar neste período, excluídas as simultâneas. O fato de a mediana das durações ser bem menor do que a média e o valor do coeficiente de assimetria ser relativamente alto (4,51 segundos) indicam assimetria positiva dos dados, com concentração nos menores valores. Há, ainda, um grande número de valores extremos na cauda direita da distribuição das durações.

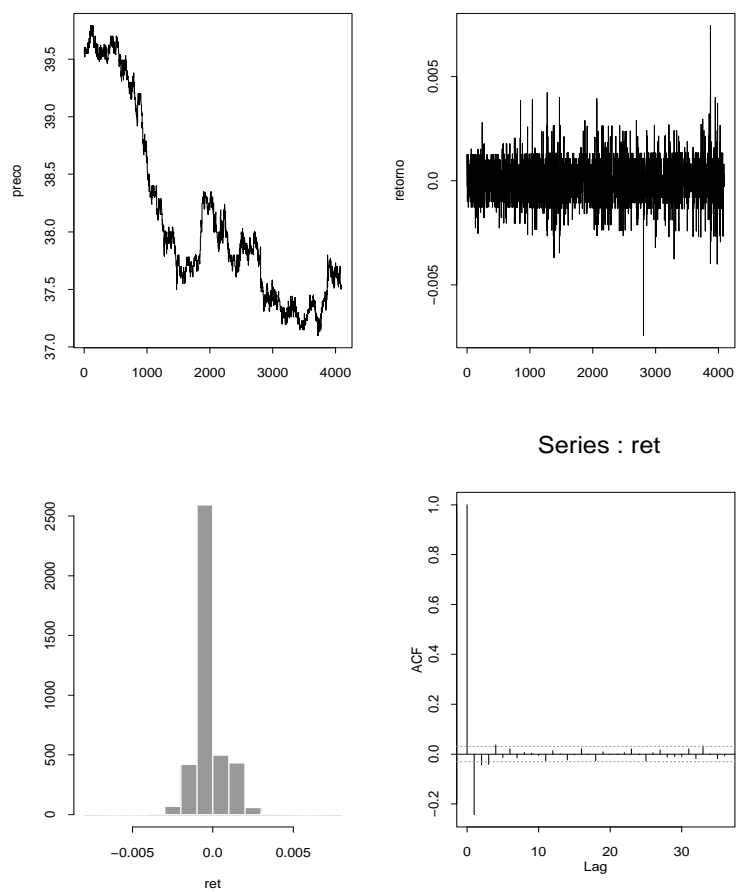


Figura 1: Séries de preços e retornos da Telemar, histograma e f.a.c. dos retornos.

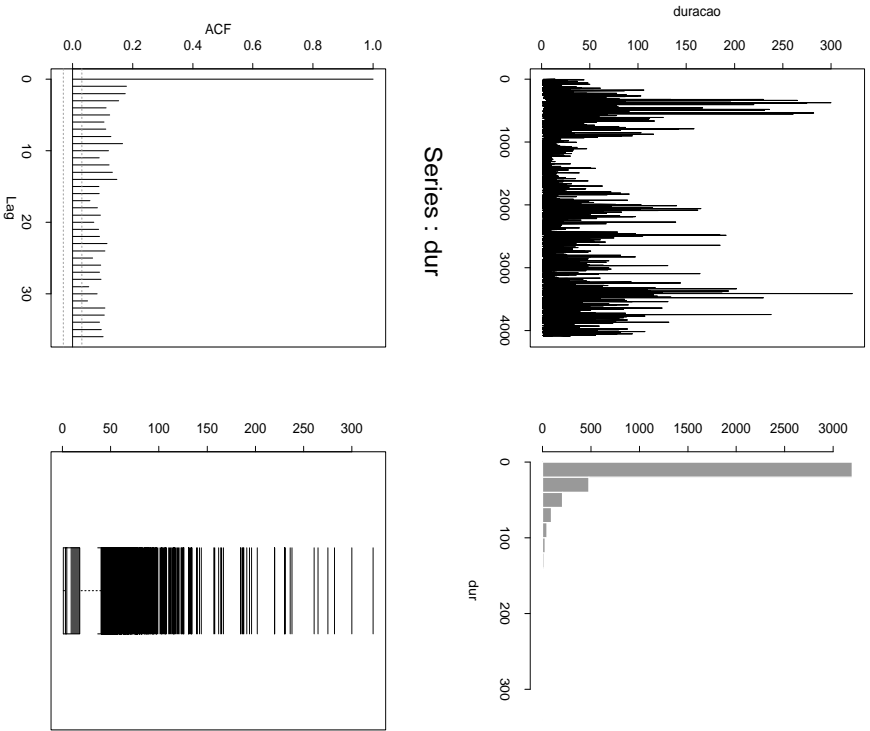


Figura 2: Série de durações da Telemar, histograma, f.a.c. e boxplot.

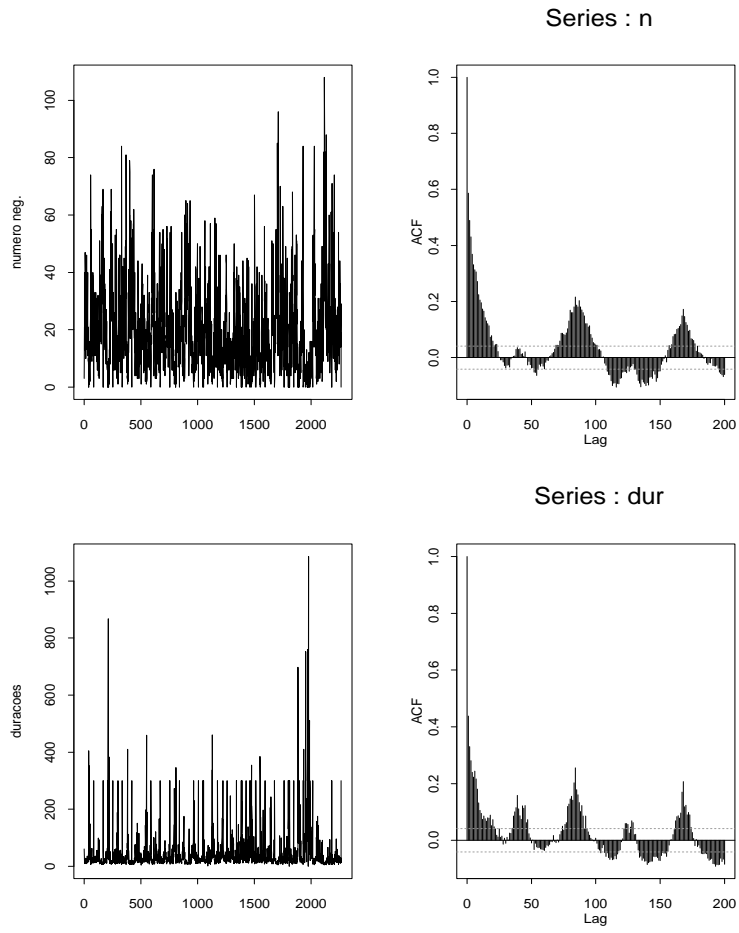


Figura 3: Número de negociações e durações da ação da Telemar em intervalos de 5 minutos e f.a.c.'s.

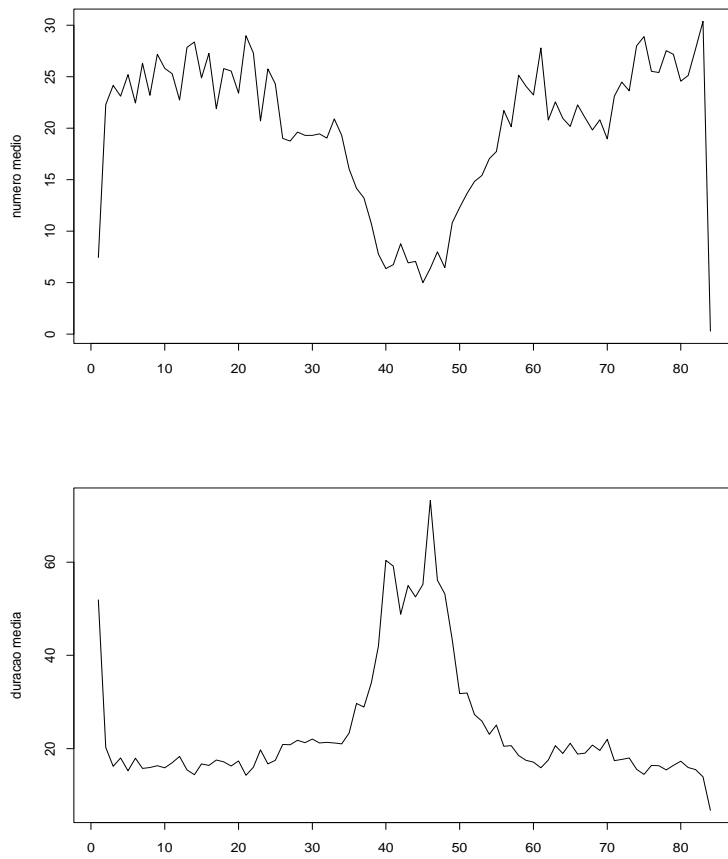


Figura 4: Número médio de negociações e durações médias da ação da Telemar em intervalos de 5 minutos.



Tabela 10.2: Negociações da Telemar no período 8-10 de setembro de 2004.

Estatística	Negociações (por dia)	Durações (em segundos)
Média	2.225	16,25
desvio padrão	575	26,43
mínimo	1.804	1
máximo	2.895	322
mediana		7
assimetria		4,51

## 2 Volatilidade Realizada

No Capítulo 5 tratamos do problema da modelagem da volatilidade de um ativo, considerando-a como uma variável latente (não-observada). Utilizamos modelos da família ARCH e modelos de volatilidade estocástica. Outras possibilidades para estimar volatilidade são usar a volatilidade implícita em modelos de opções ou volatilidade histórica.

Volatilidade e correlação são de importância fundamental em várias áreas de finanças, como apreçamento de ativos e gestão de riscos. Vimos que volatilidade dos ativos de uma carteira e as correlações entre os ativos são necessários para o cálculo do VaR, valor em risco.

Todavia, a utilização de modelos para volatilidade é, em geral, restrita a dados diários ou de menor frequência e podem resultar em previsões não-satisfatórias. Lembremos que os modelos GARCH, por exemplo, tendem a super-estimar volatilidades futuras, incorporando resultados de eventos extremos que tenham ocorrido no passado. A escolha da distribuição dos erros nestes modelos também é um problema importante, a fim de modelar adequadamente os fatos estilizados presentes nos retornos, como agrupamentos de volatilidades e caudas pesadas. Os programas disponíveis aos usuários apresentam poucas possibilidades (normal, t, GED, veja o Apêndice ao Capítulo 5). Considerando-se uma série multivariada de retornos, a situação torna-se ainda mais complicada, pois modelos heteroscedásticos condicionais são complexos e difíceis de estimar, sendo seu uso limitado a uma carteira com poucos ativos.

Recentemente, atenção está voltada para o uso de DAF intradiários para modelar e prever volatilidade e correlação por meio da chamada *volatilidade realizada*. A idéia é considerar a soma dos quadrados dos retornos obtidos em intervalos de alguns minutos durante um dia para estimar a volatilidade desse dia. Obteremos, então, uma série de volatilidades diárias observadas, que pode ser modelada com modelos como ARIMA, ARFIMA etc.

Na impossibilidade de se obter dados intradiários, podemos considerar os preços

de abertura, fechamento, máximo e mínimo de uma ação observados em cada dia, para estimar a volatilidade desse dia, usando-se, por exemplo, a volatilidade de Garman-Klass, como mostrado no Apêndice 9.

Consideremos, primeiramente, o caso de um único ativo. Denote por  $P_t$  seu preço no instante  $t$  (um dia de negócios, por exemplo) e por  $p_t = \log P_t$ , sendo o logaritmo na base  $e$ . Então, com vimos, o log-retorno (que chamaremos simplesmente retorno) no dia  $t$  é dado por  $r_t = p_t - p_{t-1}$ .

Suponha, agora, que tenhamos preços intradiários desse ativo, observados em instantes irregularmente espaçados, mas depois amostrados a cada  $\Delta t$  unidades de tempo, por exemplo a cada 30 segundos ou a cada 15 minutos. Suponha, também, que fixemos a duração do pregão de uma bolsa (a BOVESPA, por exemplo) em 7 horas, das 10 horas às 17 horas (ou das 11 horas às 18 horas, no caso de horário de verão). Se amostramos os preços a cada  $\Delta t = 15$  minutos, por exemplo, teremos  $M = 28$  preços ou  $M$  intervalos de 15 minutos por dia de negociações.

Denotemos por  $p_{d,m}$  o log-preço do ativo no período  $m$  ( $m = 1, 2, \dots, M$ ) do dia  $d$  ( $d = 1, 2, \dots, D$ ), ou seja, há  $D$  dias, de modo que o número total de observações é  $T = M \times D$ . O correspondente retorno intradiário no período  $m$  e dia  $d$  será denotado por

$$r_{d,m} = p_{d,m} - p_{d,m-1}, \quad m = 1, \dots, M, \quad d = 1, \dots, D, \quad (1)$$

com  $t = M(d-1) + m$ ,  $t = 1, \dots, T$ . Usaremos as notações  $r_t$  e  $r_{d,m}$  para denotar os retornos intradiários.

Com esta notação, o retorno diário  $r_d$  é dado por  $p_{d,M} - p_{d-1,M}$ , por exemplo se  $M = 28$ ,  $r_t = p_{d,28} - p_{d-1,28}$ .

Além dos retornos definidos acima, há o chamado retorno *overnight*, que incorpora as informações relativas ao intervalo entre o fechamento do pregão do dia anterior e a abertura do pregão do dia atual, sendo definido por

$$r_{d,0} = p_{d,0} - p_{d-1,M}, \quad (2)$$

sendo  $p_{d,0}$  o log-preço do ativo na primeira cotação do dia  $d$  e  $p_{d-1,M}$  o log-preço do ativo na última cotação do dia  $d-1$ . Incluindo-se este retorno no caso de  $\Delta t = 15$  minutos, teremos  $M = 29$  observações em cada dia.

**Exemplo 10.2.** Vamos considerar, agora, uma outra base de dados já mencionada no exemplo 1.3, ou seja, dados intradiários do IBOVESPA, obtidos de 1.381 dias úteis, em duas frequências de amostragem:

- (a)  $\Delta t = 15$  minutos, de 02/01/98 a 31/10/02, com 1.189 dias úteis, sendo que em cada dia há de 29 a 33 observações, totalizando 35.227 dados;
- (b)  $\Delta t = 30$  segundos, de 01/11/02 a 13/08/03, com 192 dias úteis, sendo que em cada dia há 865 a 872 observações, totalizando 166.236 dados.

Para a utilização desses dados originais, procedeu-se a uma limpeza dos mesmos. Os detalhes estão em Berti (2005). Como salientamos antes, a duração de cada pregão foi fixada em 7 horas, que corresponde à duração normal do pregão de viva-vós, período em que é maior o volume de negociações. Um resumo dos ajustes é o seguinte:

- (a) foram excluídos os dados anteriores a 06/04/98, por apresentarem períodos irregulares e dados discrepantes;
- (b) foram excluídos os dados de negociação eletrônica;
- (c) foi excluída a primeira hora de negociação nos casos em que o pregão viva-vós possuía duração de 8 horas;
- (d) foram excluídos dias com expediente reduzido (feriados, jogos da copa mundial de futebol etc);
- (e) foram feitas correções em dados com sinais claros de erros de digitação, por meio de interpolações lineares.

Sobre o intervalo de amostragem,  $\Delta t$ , há vários estudos, dentre os quais destacamos Andersen e Bollerslev (1998), Giot e Laurent (2004) e Oomen (2001), que analisaram intervalos variando de 5 a 25 minutos. Sob o ponto de vista teórico, para se obter consistência das medidas de volatilidade realizada é necessário que  $\Delta t \rightarrow 0$ , o que significaria amostrar continuamente, que por sua vez é uma impossibilidade. Segue-se que medidas de volatilidade realizada conterão erros de mensuração. Bandi e Russel (2003) propõem um método para escolher  $\Delta t$  que minimiza o erro quadrático médio do erro de mensuração. Bandi e Russell (2006) estudam a influência de efeitos da microestrutura do mercado (spread bid-ask, efeitos de calendário etc) no viés de medidas de volatilidade.

Consideramos, então, que um intervalo de amostragem de 15 minutos é razoável e como a maior parte dos dados está amostrada nesta frequência, no período de 01/11/02 a 13/08/03 (em que a frequência é a cada 30 segundos) também foram consideradas amostras a cada 15 minutos, resultando em cotações do IBOVESPA de 06/04/98 a 13/08/03, ou seja, 1.309 dias, com 29 observações intradiárias cada um, totalizando 37.961 pontos.

Na Figura 10.5 temos o gráfico da série de índices do IBOVESPA a cada 15 minutos, dos log-retornos, histograma dos retornos e f.a.c. Na Figura 10.6 temos os gráficos dos retornos diários e horários, com seus histogramas e fa.c.'s, respectivamente.

Pelas figuras vemos que há presença de correlação serial nas séries, mais evidentes para frequência mais altas. Já os quadrados dos retornos exibem autocorrelações maiores. Os valores da estatística de Ljung-Box mostram estes dois fatos, o que nos leva à rejeição da hipótese de ruído branco para retornos. Note a sazonalidade presente nas autocorrelações dos quadrados dos retornos intradiários a cada 15 minutos e horários, na Figura 10.7.

Na Tabela 10.3 apresentamos algumas estatísticas das séries de retornos, para um número menor de dias, 838, correspondendo a 24.302 observações da série do IBOVESPA. Verifica-se que série de retornos diários possui um alto valor da curtose, que aumenta com a frequência de amostragem. A assimetria é positiva para retornos de um dia e negativa para frequências maiores. Os valores da estatística de Jarque-Bera mostra a não-normalidade dos retornos em todas as frequências.

Tabela 10.3: Estatísticas para as séries de retornos do IBOVESPA

Estatística	15 minutos	1 hora	1 dia	LVOLR	log(GK)
Mínimo	-10,15	-10,69	-17,27	-1,50	-1,89
Primeiro quartil	-0,16	-0,40	-1,47	0,49	0,33
Média	0,00	0,0	-0,01	1,11	0,99
Mediana	0,0	0,01	-0,04	0,99	0,93
Terceiro quartil	0,16	0,41	1,41	1,58	1,56
Máximo	7,93	13,30	28,82	4,93	5,67
Desvio padrão	0,43	0,97	2,84	0,91	1,03
Assimetria	-0,73	-0,24	1,16	0,80	0,56
Curtose	42,90	17,64	16,33	1,21	1,30
Ljung-Box*	666,6	48,55	30,90	2626	1569
Ljung-Box(quadrado*)	1389,0	972,1	106,0		
Jarque-Bera*	1864750	75964	9385		

\* todas as estatísticas têm p-valor < 0,001

Veamos como construir uma medida da volatilidade realizada. Primeiramente, definimos a *variância realizada* ( $VR_d$ ) para o ativo no dia  $d$  por meio de

$$VR_d = r_{d,0}^2 + \sum_{m=1}^M r_{d,m}^2, \quad d = 1, \dots, D, \quad (3)$$

com  $r_{d,m}$  e  $r_{d,0}$  definidos em (10.1) e (10.2), respectivamente.

A *volatilidade realizada* ( $VOLR_d$ ) para o ativo no dia  $d$  é, então, definida por

$$VOLR_d = \sqrt{VR_d}, \quad (4)$$

enquanto a *log-volatilidade realizada* no dia  $d$  é definida por

$$LVOLR_d = \log(VOLR_d) \quad (5)$$

onde, novamente,  $\log x = \ln x$ .

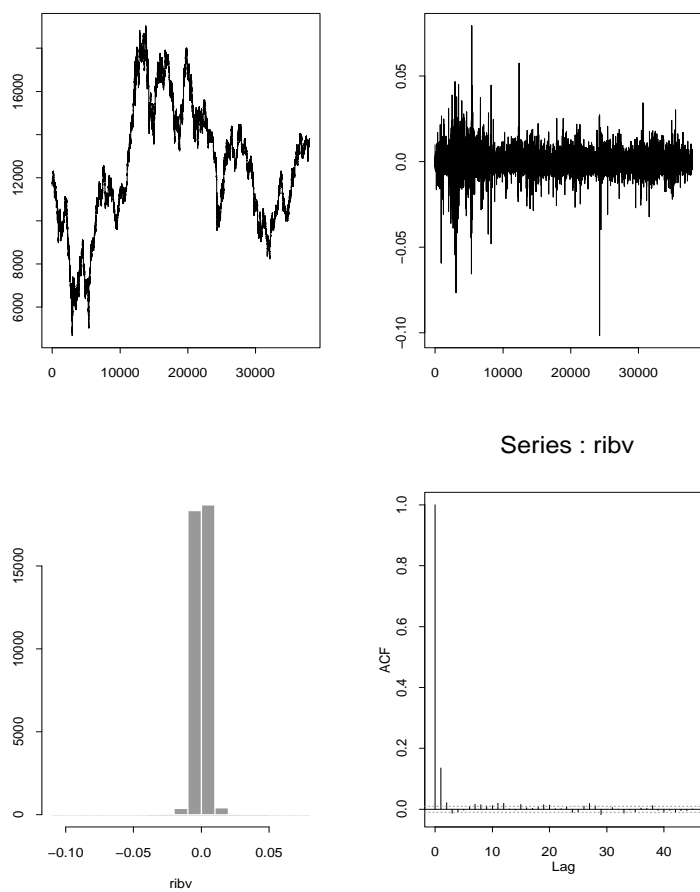


Figura 5: Série do IBOVESPA a cada 15 min, retornos, histograma e f.a.c.

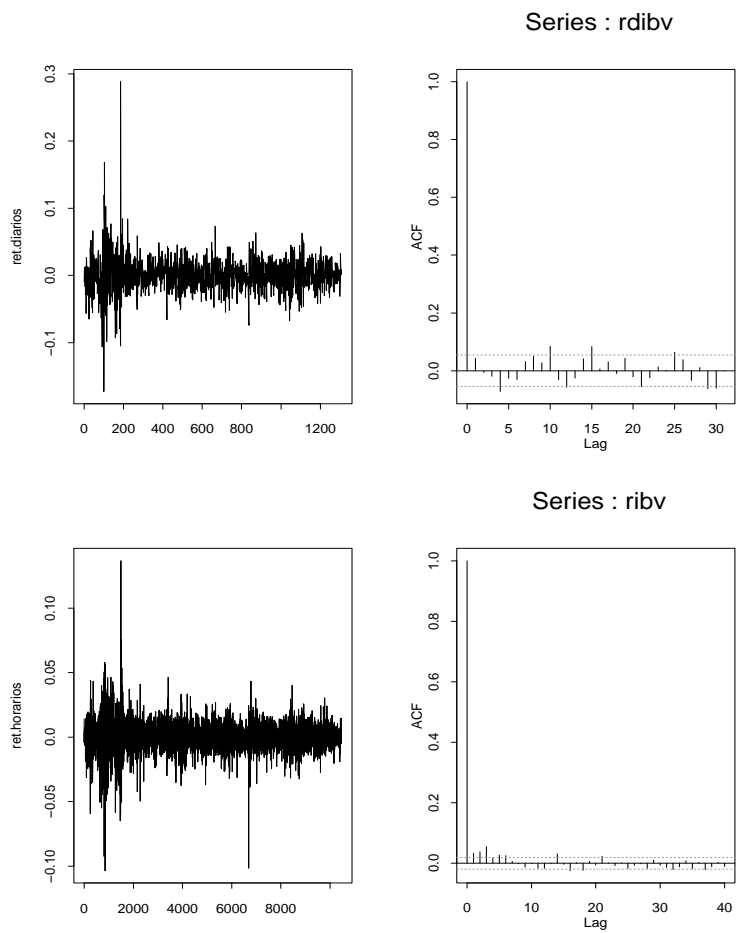


Figura 6: Retornos diários e horários do IBOVESPA, e f.a.c.'s.

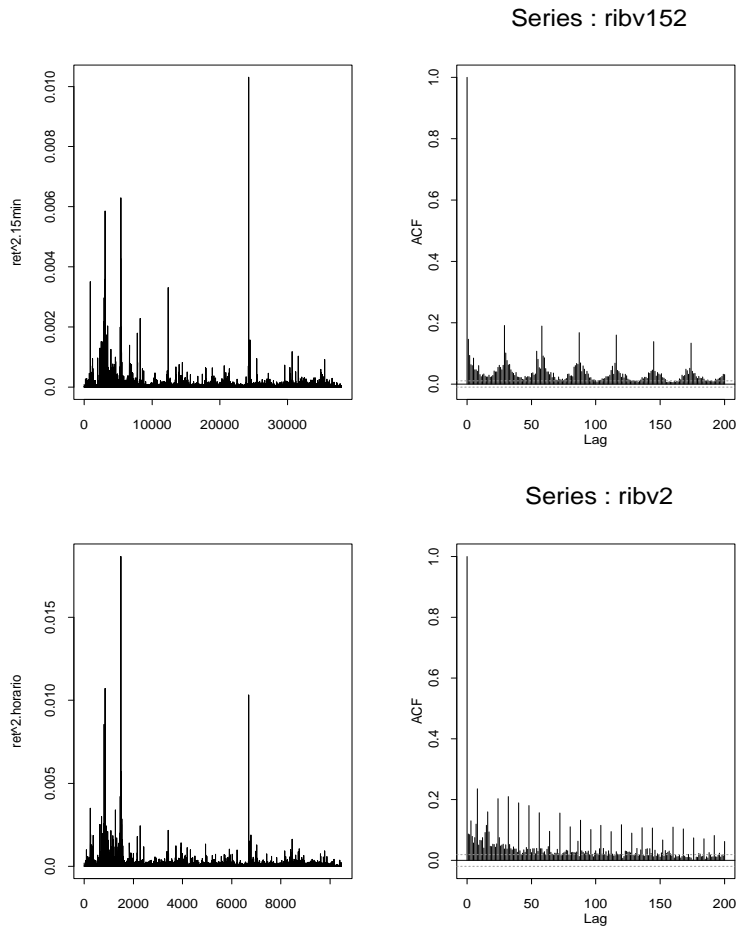


Figura 7: Quadrados dos retornos intradiários de 15 minutos e horários do IBO-VESPA e f.a.c.'s.

No caso específico do IBOVESPA intradiário, alguns dias foram excluídos no processo de filtragem dos dados, e neste caso temos que adaptar a equação (10.3), que toma a forma

$$VR_d = \begin{cases} \sum_{m=1}^{28} r_{d,m}^2, & \text{se o dia } d - 1 \text{ foi excluído,} \\ r_{d,0}^2 + \sum_{m=1}^{28} r_{d,m}^2, & \text{caso contrário,} \end{cases} \quad (6)$$

para  $d = 1, \dots, 1.308$ .

Suponha, agora, que temos  $n$  ativos, com retornos  $r_{1,d;m}, \dots, r_{n,d;m}$  e considere o vetor  $\mathbf{r}_{d,m} = (r_{1,d;m}, \dots, r_{n,d;m})'$ ,  $m = 1, \dots, M$ ,  $d = 1, \dots, D$ . Então, para cada ativo  $i$  teremos a VR dada por (10.3), que chamaremos  $VR_{i,d}$  e as respectivas volatilidade realizada ( $VOLR_{i,d}$ ) e log-volatilidade realizada ( $LVOLR_{i,d}$ ).

Para podermos calcular, por exemplo, o valor em risco de uma carteira com esses  $n$  ativos, será necessário, eventualmente, calcular as correlações entre os retornos dos ativos. Definamos a matriz de covariâncias entre os retornos dos ativos para o dia  $d$  como sendo

$$COVR_d = \sum_{m=1}^M \mathbf{r}_{d,m} \mathbf{r}_{d,m}', \quad d = 1, \dots, D. \quad (7)$$

Se  $COVR_d = (c_{i,j,d}) = ([COVR_d]_{i,j})$ , então  $c_{i,j,d} = \sum_{m=1}^M r_{i,d;m} r_{j,d;m}$ ,  $i, j = 1, \dots, n$ ,  $d = 1, \dots, D$ . Esta matriz será positiva definida se  $n < M$ . A correlação entre os retornos dos ativos  $i$  e  $j$  é dada por

$$\rho_{i,j,d} = \frac{c_{i,j,d}}{\sqrt{VR_{i,d} \times VR_{j,d}}}. \quad (8)$$

**Exemplo 10.2** (continuação) Na Figura 10.8 temos o gráfico da volatilidade realizada,  $VOLR_t$ , que estima a volatilidade diária, obtida a partir dos retornos de 15 minutos do IBOVESPA, bem como a sua f.a.c. Na figura temos, também, o gráfico da  $LVOLR_d$  e sua f.a.c. Na Figura 10.9 temos a volatilidade de Garman-Klass, obtida a partir de preços de abertura, fechamento, mínimo e máximo de cada dia (conforma Apêndice 9), sua f.a.c., o logaritmo da volatilidade de GK e sua f.a.c., para efeito de comparação. Na Tabela 10.3 encontram-se as estatísticas do  $LVOLR_d$  e  $\log(\text{GK})$ . Nas duas figuras, verificamos que as f.a.c.'s das volatilidades estimadas ( $VOLR, LVOLR, \text{GK}, \log(\text{GK})$ ) decaem lentamente para zero, indicando que todas essas séries de volatilidades apresentam memória longa.

Analisando a f.a.c. dos quadrados dos retornos da Figura 10.7 notamos uma sazonalidade de 29 minutos para os dados de 15 minutos e de 7 horas, para os dados horários. Estes padrões sazonais são importantes no procedimento de modelagem da



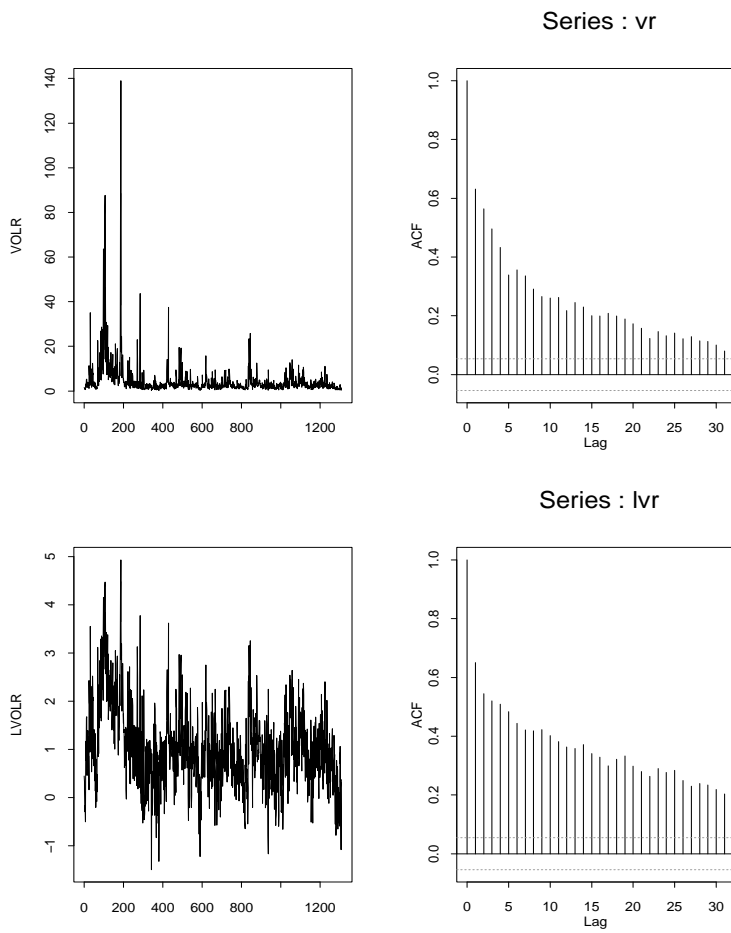


Figura 8: Volatilidade realizada e f.a.c., log(volatilidade realizada) e f.a.c.

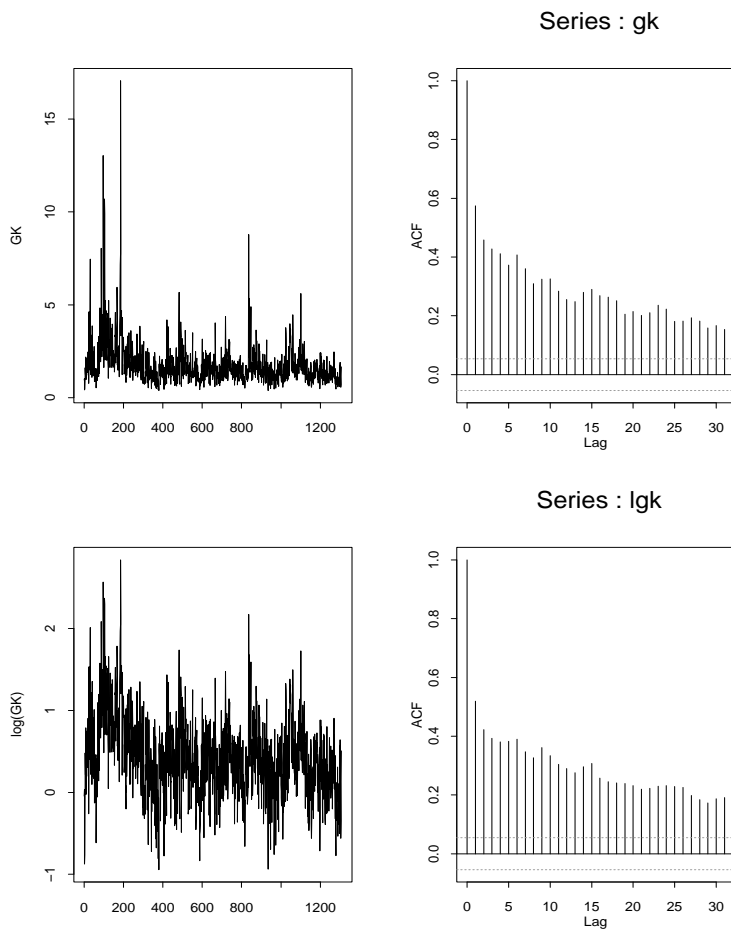


Figura 9: Volatilidade de GK e f.a.c, log(volatilidade de GK) e f.a.c.

volatilidade dos retornos intradiários. Aqui, é comum supor que estes sejam dados por  $r_t = \sigma_t s_t \varepsilon_t$ , ou seja, além da modelagem usual da volatilidade, introduzimos uma componente sazonal  $s_t$ .

Questões importantes são: (i) a retirada do padrão sazonal melhora a previsão dos modelos de volatilidade? (ii) qual é o melhor método para remover a componente sazonal? Para uma discussão dessas questões e dos diversos procedimentos de ajustamento sazonal veja Martens et al. (2002). Vamos nos limitar aqui em utilizar o seguinte procedimento, sugerido por Taylor e Xu (1997). Retomemos os retornos intradiários  $r_{d,m}$  e como estimativa da componente sazonal no período  $m$  considere

$$s_m = \left( \frac{1}{D} \sum_{d=1}^D r_{d,m}^2 \right)^{1/2}, \quad m = 1, \dots, M. \quad (9)$$

Segue-se que os retornos livres de componente sazonal são dados por  $r_t^* = r_t/s_t$ . Na Figura 10.10 temos a f.a.c dos quadrados dos retornos de 15 minutos e 1 hora dos dados do IBOVESPA, após a retirada da componente sazonal, onde não se notam mais os padrões sazonais, os quais são mostrados na Figura 10.11.

### 3 Modelo de Duração Condicional

Usando conceitos similares aos dos modelos ARCH e GARCH desenvolvidos para a volatilidade, Engle e Russell (1998) e Bollerslev (1986) introduziram o modelo auto-regressivo de duração condicional (ACD-autoregressive conditional duration) para descrever os tempos entre negociações de um particular ativo financeiro. Designemos por  $x_i$  a  $i$ -ésima duração e chamamos  $\eta_i = E(x_i|\mathcal{F}_{i-1})$ , onde  $\mathcal{F}_{i-1}$  denota a informação disponível até a  $(i-1)$ -ésima negociação.

O modelo auto-regressivo de duração condicional de ordens  $r$  e  $s$ , denotado por ACD( $r, s$ ), pode ser escrito na forma

$$x_i = \eta_i \varepsilon_i, \quad (10)$$

$$\eta_i = \omega + \sum_{j=1}^r \delta_j x_{i-j} + \sum_{j=1}^s \gamma_j \eta_{i-j}, \quad (11)$$

onde  $\varepsilon_i$  é uma seqüência de v.a. i.i.d. não negativas, com média um. Podemos considerar para  $\varepsilon_i$  distribuições como a exponencial, Weibull e Gama generalizada. Como  $x_i$  é positivo, devemos ter  $\omega \geq 0$  e  $\delta_j > 0$ ,  $\gamma_j > 0$ .

Como no caso de modelos GARCH, o processo  $\nu_i = x_i - \eta_i$  é uma seqüência de diferenças martingales, com  $E(\nu_i|\mathcal{F}_{i-1}) = 0$ , não-correlacionadas e com variância não-constante. O modelo pode, então, ser escrito na forma

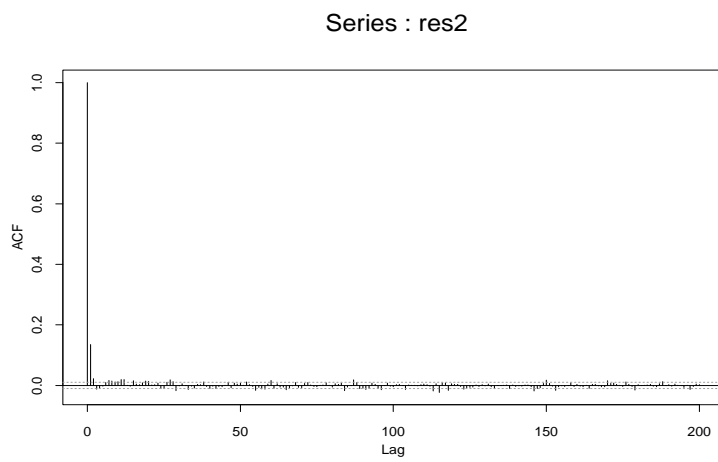
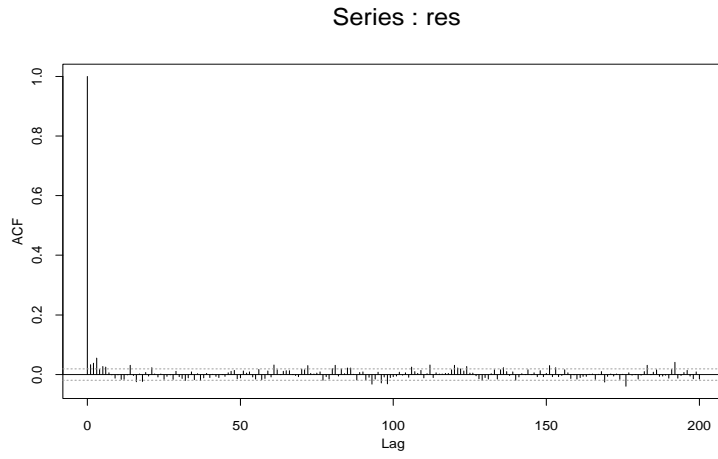


Figura 10: F.a.c.'s dos quadrados dos retornos horários e a cada 15 minutos, sem a componente sazonal.

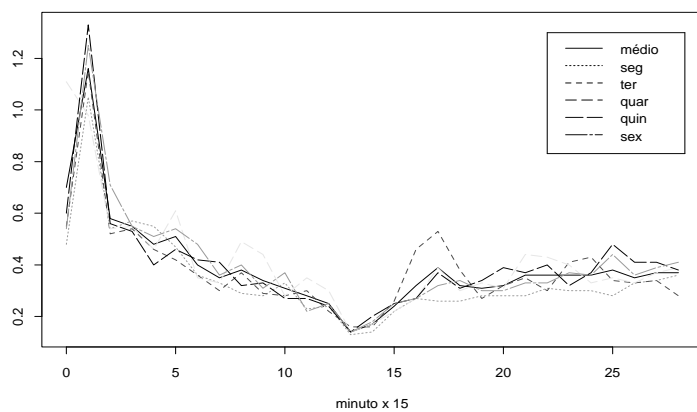
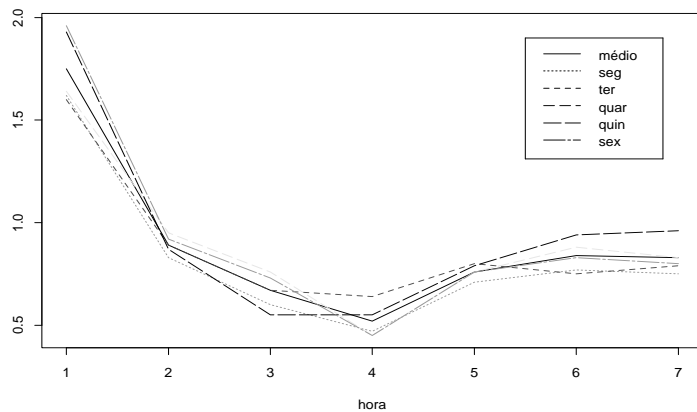


Figura 11: Padrões sazonais para os dados intradiários.

$$x_i = \omega + \sum_{j=1}^q (\delta_j + \gamma_j)x_{i-j} + \sum_{j=1}^s \gamma_j \eta_{i-j} + \nu_i, \quad (12)$$

onde  $q = \max(r, s)$ , ou seja, na forma de um modelo ARMA( $q, s$ ) com inovações não-gaussianas e heteroscedásticas. Segue-se que

$$E(x_i) = \frac{\omega}{1 - \sum_{j=1}^q (\delta_j + \gamma_j)}, \quad (13)$$

com  $\sum_{j=1}^q (\delta_j + \gamma_j) < 1$ .

O modelo mais simples é o ACD(1, 1),

$$x_i = \eta_i \varepsilon_i, \quad (14)$$

$$\eta_i = \omega + \delta x_{i-1} + \gamma \eta_{i-1}, \quad (15)$$

supondo por exemplo  $\varepsilon_i \sim$  i.i.d.  $E(1)$  (exponencial de média um). Usando os mesmos argumentos usados para o modelo GARCH(1, 1) e supondo  $x_i$  estacionária, podemos obter (veja o Problema 2):

$$E(x_i) = \mu_x = \frac{\omega}{1 - \delta - \gamma}, \quad (16)$$

$$\text{Var}(x_i) = \mu_x \left( \frac{1 - \gamma^2 - 2\delta\gamma}{1 - 2\delta^2 - 2\delta\gamma - \gamma^2} \right). \quad (17)$$

De (10.17) devemos ter  $\gamma^2 + 2\delta\gamma + 2\delta^2 < 1$ , para que a variância incondicional seja constante no tempo.

Usando a forma ARMA(1, 1) teremos

$$x_i = \omega + (\delta + \gamma)x_{i-1} + \nu_t - \gamma\nu_{i-1},$$

da qual obtemos que a f.a.c de um modelo ACD(1, 1) é dada por

$$\rho_1 = \frac{\delta(1 - \gamma^2 - \delta\gamma)}{1 - \gamma^2 - 2\delta\gamma}, \quad (18)$$

$$\rho_j = (\delta + \gamma)\rho_{j-1}, \quad j \geq 2. \quad (19)$$

Veja o Problema 3.

Há várias alternativas sugeridas na literatura para o modelo ACD, dentre as quais destacamos: o modelo de duração estocástico (Bauwens e Vereda, 2000), a versão logaritmica (Bauwens e Giot, 2000), a versão não-linear (Zhang et al., 2001),

o modelo ACD assimétrico (Bauwens e Giot, 2003) e o modelo ACD aumentado (Fernandes e Gramming, 2005).

### Identificação

Assim como para um modelo GARCH, a especificação da ordem de um modelo ACD é uma tarefa complicada, e portanto sugere-se que se ajuste modelos de ordens não muito altas e se utilize algum critério (como AIC, BIC ou log-verossimilhança) para a escolha do modelo.

### Estimação

Supondo que os erros  $\varepsilon_t$  sigam alguma distribuição, como uma das mencionadas acima, podemos usar o método de máxima verossimilhança para obter os estimadores dos parâmetros do modelo ACD. Chamando de  $\boldsymbol{\theta}$  o vetor de parâmetros e  $\mathbf{x} = (x_1, \dots, x_T)'$  o vetor de durações observadas, obtemos a função de verossimilhança

$$L(\boldsymbol{\theta}; \mathbf{x}) = \left[ \prod_{i=2}^T f(x_i | \mathcal{F}_{i-1}, \boldsymbol{\theta}) \right] f(x_1 | \boldsymbol{\theta}). \quad (20)$$

A densidade marginal  $f(x_1 | \boldsymbol{\theta})$  tem importância decrescente, quando  $T$  cresce e pode ser omitida (Tsay, 2005), quando da maximização de (10.20) para obter os EMV.

**Exemplo 10.3.** Suponha o modelo ACD(1,1) com erros seguindo uma distribuição exponencial com média unitária. Então,

$$f(x_i | \mathcal{F}_{i-1}, \boldsymbol{\theta}) = \exp \left\{ -\frac{x_i}{\eta_i} \right\} \frac{1}{\eta_i},$$

e a log-verossimilhança fica

$$\begin{aligned} \ell(\boldsymbol{\theta}; \mathbf{x}) &= \sum_{i=2}^T [-\ln(\eta_i) - x_i/\eta_i] \\ &= \sum_{i=2}^T [-\ln(\omega + \delta x_{i-1} + \gamma \eta_{i-1}) - x_i/(\omega + \delta x_{i-1} + \gamma \eta_{i-1})], \end{aligned} \quad (21)$$

com  $\boldsymbol{\theta} = (\omega, \delta, \gamma)'$ .

### Diagnóstico

A verificação da adequação do modelo é feita nos moldes usuais, analisando a f.a.c. dos resíduos e dos quadrados dos resíduos do modelo ajustado e calculando as respectivas estatísticas de Box-Pierce-Ljung.

Como a série de durações apresenta uma periodicidade, como vimos na seção 10.1, antes de ajustar o modelo esta sazonalidade tem que ser removida, usando a mesma metodologia que foi vista para a volatilidade realizada.

**Exemplo 10.4.** Retomemos as durações em intervalos de 5 minutos da Telemar, no período de 4 de agosto a 10 de setembro de 2004. Removemos a sazonalidade usando um procedimento similar ao da Seção 10.2 e consideramos as 4096 durações no período de 8 a 10 de setembro de 2004. A esta série livre de componente sazonal ajustamos um modelo ACD(1,1), com erros exponenciais de média 1. O programa RATS foi utilizado para o procedimento de estimação. Veja Tsay (2005, pag. 246) para um exemplo de tal programa. O modelo obtido é

$$x_i = \eta_i \varepsilon_i, \quad (22)$$

$$\eta_i = 0,01 + 0,07x_{i-1} + 0,918\eta_{i-1}, \quad (23)$$

onde os respectivos desvios padrões das estimativas são 0,002, 0,004 e 0,005. Na Figura 10.12 temos a f.a.c. dos resíduos e o plot  $Q \times Q$  (quantis empíricos dos resíduos versus quantis da  $Exp(1)$ ). Vemos que o modelo ajustado não consegue captar algumas durações extremas, que estão destacadas no plot  $Q \times Q$ . Uma possibilidade é considerar modelos ACD com coeficientes variando no tempo a fim de levar em conta a não-estacionariedade das durações e a sazonalidade presente. Veja Bruscato (2006). Para fazer previsões com o modelo ajustado é necessário recompor a série, por meio do padrão sazonal previamente estimado, ou seja,  $\hat{y}_i = \hat{x}_i s_5$ .

## 4 Modelagem da Volatilidade

No Capítulo 5 vimos como modelar a volatilidade, supondo-a latente (não-observada), usando modelos da família ARCH. Como a volatilidade realizada é efetivamente observada, podemos tentar modelá-la por meio de modelos da família ARIMA, por exemplo. Contudo, vimos que a volatilidade realizada apresenta características de processos de memória longa, logo modelos da classe ARFIMA podem ser úteis. Outra possibilidade é considerar modelos ARCH, GARCH e extensões para dados intra-diários e usá-los para fornecer informação adicional para os dados diários. Veja Hol e Koopman (2002), por exemplo. A idéia é ajustar modelos GARCH, digamos, para frequências de uma hora ou 15 minutos e realizar previsões da volatilidade nessas frequências  $k$  passos à frente até cobrir o período de um dia ( $k = 7$  para dados horários e  $k = 29$  para dados a cada 15 minutos), agregando-se as previsões para obter a volatilidade diária. Observemos que antes de ajustar



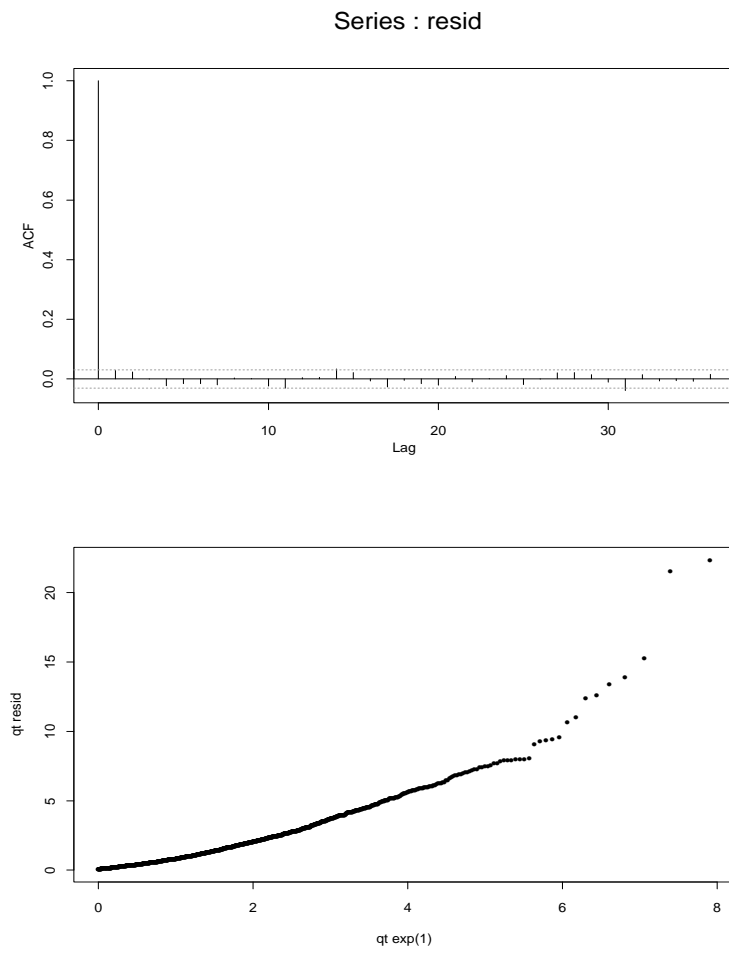


Figura 12: F.a.c dos resíduos do modelo (10.22)-(10.23) e plot  $Q \times Q$ .

um modelo ARMA-GARCH aos dados intra-diários originais é necessário retirar a sazonalidade, conforme visto na seção 10.2.

**Exemplo 10.5.** Consideremos os dados intra-diários do IBOVESPA, com 1309 dias de observações. Vamos considerar 838 dias para estimar modelos e 471 dias para previsões. No caso de dados horários teremos 24302 dados e no caso de dados a cada 15 minutos, teremos 5886 dados para estimação. Para obter previsões da volatilidade diária, faremos previsões  $k$  passos à frente,  $k = 7$  ou  $k = 29$ , respectivamente.

A avaliação da acurácia preditora dos modelos de volatilidade é geralmente baseada em funções de perda, comparando os valores previstos do quadrado da volatilidade pelo modelo usado,  $\hat{h}_t^2$ , com os quadrados dos retornos,  $r_t^2$ . No entanto, Andersen e Bollerslev (1998) sugerem usar a volatilidade realizada, no lugar dos quadrados dos retornos. Se  $T$  indica o número total de dados e  $T_0$  o número de observações usadas para estimar o modelo, algumas medidas usadas são:

(a) Erro médio:

$$EM = \frac{1}{T - T_0} \sum_{t=T_0+1}^T (\hat{h}_t^2 - VOLR_t^2).$$

(b) Erro quadrático médio:

$$EQM = \frac{1}{T - T_0} \sum_{t=T_0+1}^T (\hat{h}_t^2 - VOLR_t^2)^2.$$

(c) Erro absoluto médio:

$$EAM = \frac{1}{T - T_0} \sum_{t=T_0+1}^T |\hat{h}_t^2 - VOLR_t^2|.$$

Berti (2005) ajustou várias combinações de modelos ARMA-GARCH (por GARCH aqui entendemos GARCH, EGARCH, PGARCH etc), sendo que os melhores modelos segundo medidas de perda, incluindo aquela acima, foram:

(i) Para log-retornos a cada 15 minutos: modelo AR(2)-GARCH(1,2), com distribuição  $t$  para os erros.

(ii) Para log-retornos horários: modelo AR(5)-GARCH(1,2), com distribuição  $t$  para os erros.

(iii) Para log-retornos diários: modelo AR(4)-PGARCH(1,1), com distribuição  $t$  para os erros e com a variância realizada  $VR_{t-1}$  como co-variável. Veja o Capítulo 5, Problema 13, para a definição do modelo “Power GARCH”.

(iv) Para a volatilidade realizada: modelo ARFIMA(0,d,0), com  $\hat{d} = 0,438$ .

Veja Berti (2005) para detalhes e os Problema 8, 9,10 e 11. Duas observações sobre o ajuste desses modelos: (a) Os modelos GARCH intradiários em geral superestimam as previsões da volatilidade; (b) os modelos ARFIMA em geral subestimam a volatilidade em períodos de alta volatilidade. Seria interessante também avaliar a eficácia dos modelos de volatilidade estocástica aplicados a dados intradiários.

## 5 Comentários Adicionais

[1] Andersen et al. (2000, 2001, 2003) (designados brevemente por ABDL) e Barndorff-Nielsen e Shephard (2002, 2004) (designados brevemente por BS) desenvolveram os aspectos teóricos relativos à variância e volatilidade realizada. Seja  $p(t)$  o log-preço de um ativo,  $r(d, d-1)$  o respectivo retorno diário e  $r(d) = p(d) - p(0)$  o retorno acumulado até  $d$ . A *variação quadrática* (VQ) é definida por

$$[r](d) = p \lim \sum_{j=0}^{n-1} [p(s_{j+1}) - p(s_j)]^2,$$

onde  $0 = s_0 < s_1 < \dots < s_n = d$  e o limite é para  $\max_{1 \leq j \leq n} |s_j - s_{j-1}| \rightarrow 0$ ,  $n \rightarrow \infty$ . Então, BS provam que

$$VR_d \xrightarrow{P} [r](d) - [r](d-1) = VQ_d, \quad n \rightarrow \infty.$$

Ou seja, a variância realizada diária converge, em probabilidade, para incrementos diários da VQ. ABDL provam que

$$\text{Var}(r(d, d-1) | \mathcal{F}_{d-1}) = E[VQ_d | \mathcal{F}_{d-1}],$$

ou seja, a variância condicional dos retornos diários é igual à esperança condicional do processo de variação quadrática diária.

Para a classe de processos de Ito contínuos, caracterizados pela equação diferencial estocástica

$$dp(t) = \mu(t)dt + \sigma(t)dW(t),$$

onde  $\sigma(t)$  é a volatilidade e  $W(t)$  é o MBP, resultados mais fortes podem ser obtidos, em particular,  $r(d, d-1)/VOLR_d$  tem uma distribuição aproximadamente normal. Também é possível provar que a VOLR e LVOLR têm distribuições assintoticamente normais. Para detalhes, veja os autores acima mencionados.

[2] Podemos considerar outras classes de modelos para DAF. Müller et al. (1997) introduziram os modelos HARCH (heterogeneous ARCH). Um modelo HARCH(n) para retornos é dado por

$$r_t = \sigma_t \varepsilon_t,$$

$$\sigma_t^2 = c_0 + \sum_{j=1}^n c_j \left( \sum_{i=1}^j r_{t-i} \right)^2,$$

com  $c_0 > 0, c_n > 0, c_j \geq 0$ , para  $j = 1, \dots, n-1$ , e os  $\varepsilon_t$  são v.a. i.i.d. com média zero e variância um. Este modelo leva em conta a heterogeneidade de informação do mercado financeiro, permitindo agregar retornos. Por exemplo, no modelo HARCH(2) a volatilidade toma a forma

$$\sigma_t^2 = c_0 + c_1 r_{t-1}^2 + c_2 (r_{t-1} + r_{t-2})^2.$$

Ruilova (2007) generaliza esse modelo para a classe GHARCH(m,p), que permite agregações de diferentes tamanhos, sendo  $m$  o número de componentes do mercado.

## 6 Problemas

1. Considere os dados intradiários da Telemar do exemplo 10.1 (arquivo id-tlm09.04.dat). Obtenha a volatilidade realizada diária, faça seu gráfico e obtenha sua f.a.c. Obtenha a f.a.c. dos quadrados dos retornos e verifique se existe sazonalidade. Caso haja, remova o padrão sazonal.
2. Mesmo problema para os dados intradiários da IBM (arquivo id-ibm.dat).
3. Prove as relações (10.16) e (10.17).
4. Prove as relações (10.18) e (10.19).
5. Obtenha (10.21) usando a distribuição de Weibull padrão para os erros  $\varepsilon_t$ :

$$f(\varepsilon, \alpha) = \alpha \cdot \varepsilon^{\alpha-1} \cdot \exp\{-\varepsilon^\alpha\}, \quad \alpha > 0.$$

6. Ajuste um modelo ACD para as durações da IBM (arquivo id-ibm.dat), com erros exponenciais de média 1.
7. Mesmo problema, com distribuição Weibull padrão.
8. Ajuste os modelos ARMA-GARCH sugeridos na seção 10.4 aos dados de retornos diários e intradiários do IBOVESPA.
9. Ajuste o modelo ARFIMA sugerido na seção 10.4 aos dados de volatilidade realizada do IBOVESPA.

10. Ajuste modelos ARMA-GARCH para os dados intra-diários da Telemar. Escolha o melhor modelo segundo as medidas dadas na seção 10.4.
11. Ajuste modelos ARFIMA( $p, d, q$ ) para a volatilidade realizada da Telemar obtida no Problema 1. Escolha o modelo que fornece o melhor EQM de previsão.
12. Considere os dados de taxa de câmbio Euro/Dólar de 1 de janeiro de 1999 a 31 de dezembro de 2002, a cada 5 minutos, num total de  $T = 288860$  observações (arquivo id-eudo99.02.dat). Obtenha o gráfico, histograma e f.a.c. dos retornos e dos quadrados dos retornos da série.
13. Para os dados do problema anterior, obtenha a volatilidade realizada, o logaritmo da volatilidade realizada, seus histogramas e f.a.c.'s.
14. Ajuste um modelo ARFIMA para a variância realizada dos dados do Euro/Dólar.

## Apêndice 10: Notas Complementares

### A.10.1. Tipos de Mercados

Podemos dividir os mercados acionários em duas categorias:

(a) *price driven markets*, onde há a figura dos especialistas ou *market-makers* (que trabalham para as próprias bolsas) e que são os responsáveis por comprar (ao bid price) e vender (ao ask price) lotes de ações, evitando movimentos abusivos nos preços das ações, e dando liquidez ao mercado. Brevemente, liquidez é definida como a habilidade de comprar e vender rapidamente um grande volume de ações com um mínimo de impacto nos preços. A bolsa NASDAQ adota este mecanismo.

(b) *order driven markets*, onde não há market-makers no processo de negociações, mas as ordens são feitas num *order book*, ou seja, cada ordem é colocada automaticamente no livro de ofertas de cada ação, sendo possível visualizar as melhores compras e vendas por lote, preço e contraparte. A BOVESPA adota este mecanismo, sendo que as corretoras contratadas por empresas para dar liquidez a seus papéis é que colocam as ordens. A NYSE adota um sistema híbrido. A BOVESPA também poderia ser pensada desta forma, supondo que as corretoras fazem o papel dos market-makers.

Para detalhes sobre este assunto e sobre liquidez de mercados veja Bauwens e Giot (2001).

### A.10.2. Mercado de Taxa de Câmbio(FX)

Este é um tipo especial de mercado, funcionando continuamente, 24 horas por dia, do tipo “price driven market”. Os participantes são “market-makers” (“dealers”),

representando grandes bancos.

Podemos considerar três zonas, correspondendo a três mercados geográficos: Ásia, Europa e EUA. Estes mercados operam da seguinte forma: após meia-noite GMT os mercados de Tokyo, Hong Kong e Singapura estão ativos, com um decréscimo de operações ao redor de 4 horas GMT, que corresponde ao horário de almoço. Ao redor das 8 horas GMT as negociações começam na Europa (Londres, Frankfurt e Paris) e param na Ásia. A última zona começa ao redor das 15 horas GMT com a abertura da bolsa de Nova Iorque. Este padrão intradiário é bem conhecido.

Outra característica dos mercados FX é que as negociações ocorrem somente entre os “dealers” e a informação sobre preços e volumes permanece confidencial. Veja Bauwens e Giot (2001) para detalhes.

### A.10.3. Algumas Distribuições Especiais

#### Distribuição exponencial

A v.a.  $X$  tem distribuição exponencial de parâmetro  $\beta > 0$  se sua densidade é dada por

$$f(x, \beta) = \frac{1}{\beta} \exp\{-x/\beta\} . I\{x \geq 0\}.$$

Se  $\beta = 1$  obtemos a distribuição exponencial padrão,  $E(X) = 1$  e escrevemos  $X \sim E(1)$ .

#### Distribuição Gama

Dizemos que a v.a.  $X$  tem distribuição gama com parâmetros  $\alpha > 0$  e  $\beta > 0$ , e escrevemos  $X \sim \text{Gama}(\alpha, \beta)$ , se a densidade é dada por

$$f(x, \alpha, \beta) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} \exp\{-x/\beta\} . I\{x \geq 0\}.$$

Se  $\beta = 1$  obtemos a distribuição gama padrão com parâmetro  $\alpha$ .

#### Distribuição de Weibull

A v.a.  $X$  tem distribuição de Weibull com parâmetros  $\alpha > 0, \beta > 0$  se a densidade é dada por

$$f(x, \alpha, \beta) = \frac{\alpha}{\beta^\alpha} x^{\alpha-1} \exp\{-(x/\beta)^\alpha\} . I\{x \geq 0\}.$$

Se  $\alpha = 1$  obtemos a distribuição exponencial e se  $Y = \frac{X}{\beta \Gamma(1+1/\alpha)}$  obtemos a distribuição de Weibull padrão, com  $E(Y) = 1$  e densidade

$$f(y, \alpha) = 1 - \exp\{-[\Gamma(1 + 1/\alpha)y]^\alpha\}.I\{y \geq 0\}.$$

### Distribuição Gama Generalizada

Dizemos que  $X$  tem distribuição gama generalizada com parâmetro de escala  $\beta$  e parâmetros de forma  $\alpha$  e  $\gamma$ , todos positivos, se a densidade é da forma

$$f(x, \gamma, \beta, \alpha) = \frac{\gamma}{\beta^\gamma \Gamma(\alpha)} x^{\gamma\alpha-1} \exp\left[-\left(\frac{x}{\beta}\right)^\gamma\right].I\{x \geq 0\}.$$

Se  $Y = \lambda X/\beta$ , com  $\lambda = \Gamma(\alpha)/\Gamma(\alpha + 1/\gamma)$ , então  $E(Y) = 1$  e  $Y$  tem distribuição gama generalizada padrão. Se  $\alpha = 1$  obtemos a distribuição de Weibull, logo exponencial e Weibull são casos particulares da gama generalizada.

### Referências

- Akaike, H. (1973). Maximum likelihood identification of Gaussian autoregressive moving average models. *Biometrika*, **60**, 255–265.
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, AC-19, 716–723.
- Akaike, H. (1977). On entropy maximization principle. In *Applications of Statistics* (P.R. Krishnaiah, ed.), 27–41. Amsterdam: North-Holland.
- Alexander, C. (2001). *Market Models*. Wiley.
- Andersen, T., Bollerslev, T., Diebold, F.X. and Labys, P. (2000a). Exchange rates returns standardized by realized volatility are (nearly) Gaussian. *Multinational Finance Journal*, **4**, 159–179.
- Andersen, T., Bollerslev, T., Diebold, F.X. and Labys, P. (2000b). Great realizations. *Risk*, **13**, 105–108.
- Andersen, T., Bollerslev, T., Diebold, F.X. and Labys, P. (2001). The distribution of realized exchange rate volatility. *Journal of the American Statistical Association*, **96**, 42–55.
- Andersen, T., Bollerslev, T., Diebold, F.X. and Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, **71**, 579–626.
- Andersen, T. and Bollerslev, T. (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review*, **39**, 885–905.

- Andersson, F., Mausser, H., Rosen, D. and Uryasev, S. (2001). Credit risk optimization with conditional value-at-risk criterion. *Mathematical Programming, Series B*, **89**, 273–291.
- Aparicio, F.M. and Estrada, J. (2001). Empirical distributions of stock returns: European securities markets, 1990-1995. *European Journal of Finance*, **7**, 1–21.
- Artzner, P., Delbaen, F., Eber, J.M. and Heath, D. (1997). Thinking coherently. *Risk*, **10**, 68–71.
- Artzner, P., Delbaen, F., Eber, J.M. and Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, **9**, 203–228.
- Baillie, R. T. (1996). Long memory processes and fractional integration in econometrics. *Journal of Econometrics*, **73**, 5–59.
- Bandi, F.M. and Russell, J.R. (2003). Microstructure noise, realized volatility, and optimal sampling. Working paper.
- Bandi, F.M. and Russell, J.R. (2006). Volatility. Working paper, GSB, University of Chicago.
- Banerjee, A., Dolado, J.J., Galbraith, J.W. and Hendry, D.F. (1993). *Co-integration, Error Correction and the Econometric Analysis of Non-Stationary Data*. Oxford: Oxford University Press.
- Barndorff-Nielsen, O.E. and Shephard, N. (2002a). Estimating quadratic variation using realized variance. *Journal of Applied Econometrics*, **17**, 457–477.
- Barndorff-Nielsen, O.E. and Shephard, N. (2002b). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society, Series B*, **64**, 253–280.
- Barndorff-Nielsen, O.E. and Shephard, N. (2004a). Econometric analysis of realized covariation: High frequency based covariance, regression, and correlation in financial economics. *Econometrica*, **73**, 885–926.
- Barndorff-Nielsen, O.E. and Shephard, N. (2004b). How accurate is the asymptotic approximation to the distribution of realized volatility? In *Identification and Inference for Econometric Models. A Festschrift in Honour of T.J. Rothenberg*. Ed. by D.W.K. Andrews, J. Powell, P.A. Ruud, and J.H. Stock. Cambridge: Cambridge University Press.
- Bauwens, L. and Giot, P. (2000). The logarithmic ACD model: An application to the bid-ask quote process of three NYSE stocks. *Annales d'Economie et de Statistique*, **60**, 117–150.



- Bauwens, L. and Vereda, D. (2000). The stochastic conditional duration model: A latent factor model for the analysis of financial durations. *Journal of Econometrics*, **119**, 381–412.
- Bauwens, L. and Giot, P. (2001). *Econometric Modelling of Stock Market Intraday Activity*. Boston: Kluwer Academic Publishers.
- Bauwens, L. and Giot, P. (2003). Asymmetric ACD models: Introducing price information in ACD models. *Empirical Economics*, **28**, 709–731.
- Bera, A.K. and Jarque, C.M. (1981). An efficient large sample test for normality of observations and regression residuals. Working paper in Econometrics N0. 40, Australian National University, Canberra.
- Beran, J. (1995) Maximum likelihood estimation of the differencing parameter for invertible short and long memory ARIMA models. *Journal of the Royal Statistical Society, Series B*, **57**, 659–672.
- Berkowitz, J. and O'Brien, J. (2002). How accurate are the Value-at-Risk models at commercial banks? *The Journal of Finance*, Vol. LVII, 1093–11.
- Berti, A.F. (2005). Estimação da Volatilidade Diária com Dados de Alta Frequência: Aplicações ao Cálculo do Valor em Risco do IBOVESPA. Dissertação de Mestrado, Instituto de Matemática e Estatística da Universidade de São Paulo.
- Beveridge, W.H. (1921). Weather and harvest cycles. *Economics Journal*, **31**, 429–452.
- Black, F. (1976). Studies of stock market volatility changes. Proc. 1976 Meeting of the American Statistical Association, Business and Economic Statistics Section, 177–181.
- Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, **81**, 635–654.
- Blattberg, R.C. and Gonedes, N.J. (1974). A comparison of stable and Student distributions as statistical models for stock prices. *The Journal of Business*, **47**, 244–280.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, **31**, 307–327.
- Bollerslev, T. (1987). A conditionally heteroskedastic time series model for speculative process and rates of return. *Review of Economics and Statistics*, **69**, 542–547.
- Bollerslev, T. (1988). On the correlation structure for the generalized autoregressive

- conditional heteroskedasticity. *Journal of Time Series Analysis*, **9**, 121–132.
- Bollerslev, T., Chou, R.Y. and Kroner, K.F. (1992). ARCH modeling in finance: A review of the theory and empirical evidence. *Journal of Econometrics*, **52**, 5–59.
- Bollerslev, T., Engle, R.F. and Nelson, D.B. (1994). Arch Models. In *Handbook of Econometrics*, Vol. IV (eds. R.F. Engle and D.L. McFadden), 2959–3038. New York: North Holland.
- Bollerslev, T. and Mikkelsen, H.O. (1996). Modeling and pricing long memory in stock market volatility. *Journal of Econometrics*, **73**, 151–184.
- Box, G.E.P. and Pierce, D.A. (1970). Distribution of autocorrelations in autoregressive moving average models. *Journal of The American Statistical Association*, **65**, 1509–1526.
- Box, G.E.P., Jenkins, G.M. and Reinsel, G. (1994). *Time Series Analysis: Forecasting and Control*. Third Edition. Englewood Cliffs: Prentice Hall.
- Boudjellaba, H., Dufour, J.-M. and Roy, R. (1992). Testing causality between two vectors in multivariate autoregressive moving average models. *Journal of the American Statistical Association*, **87**, 1082–1090.
- Breidt, F.J. and Carriquiry, A.L. (1996). Improved quasi-maximum likelihood estimation for stochastic volatility models. In *Modelling and Prediction: Honoring Seymour Geisser* ( J.C. Lee and A. Zellner, eds.), 228–247. New York: Springer.
- Breidt, F.J., Crato, N. and de Lima, P.J.F. (1993). Modeling long-memory stochastic volatility. Working paper (Johns Hopkins University, Baltimore, MD).
- Brockwell, P.J. and Davis, R.A. (1991). *Time Series: Theory and Methods*. Second Edition. New York: Springer.
- Bruscatto, A. (2006). Modelo Auto-Regressivo de Duração Condicional com Coeficientes Variando no Tempo. Tese de Doutorado, IME-USP.
- Campbell, J.Y., Lo, A.W. and MacKinlay, A.C. (1997). *The Econometrics of Financial Markets*. Princeton: Princeton University Press.
- Carmona, R.A. (2004). *Statistical Analysis of Financial Data in S-Plus*. New York: Springer.
- Coles, S. (2001). *An Introduction to Statistical Modeling of Extremes*. London: Springer.
- Cramér, H. and Leadbetter, M.R. (1967). *Stationary and Related Stochastic Processes*. New York: Wiley.

- da Cunha, D. M. S. (1997). *Análise de Causalidade em Séries Temporais*. Dissertação de mestrado, IME-USP, São Paulo.
- Dacorogna, M.M., Muller, U.A., Nagler, R.J., Olsen, R.B. and Pictet, O.V. (1993). A geographical model for the daily and weekly seasonal volatility in the foreign exchange market. *Journal of International Money and Finance*, **12**, 413–438.
- Dacorogna, M.M., Gençay, R., Muller, U.A., Olsen, R.B. and Pictet, O.V. (2001). *An Introduction to High-Frequency Finance*. San Diego: Academic Press.
- Dickey, D.A. and Fuller, W.A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, **74**, 427–431.
- Dickey, D.A. and Fuller, W.A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica*, **49**, 1052–1072.
- Dickey, D.A. and Pantula, S. (1987). Determining the order of differencing in autoregressive processes. *Journal of Business and Economic Statistics*, **15**, 455–461.
- Ding, Z., Granger, C.W.J. and Engle, R.F. (1993). A long memory property of stock returns and a new model. *Journal of Empirical Finance*, **1**, 83–106.
- Dolado, J.J., Jenkinson, T. and Sosvilla-Rivero, S. (1990). Cointegration and unit roots. *Journal of Economic Surveys*, **4**, 249–273.
- Dunsmuir, W. and Hannan, E.J. (1976). Vector linear time series models. *Advances in Applied Probability*, **8**, 339–364.
- Durbin, J. (1970). An alternative to the bounds test for testing serial correlation in least-squares regression. *Econometrica*, **38**, 422–429.
- Durbin, J. and Koopman, S.J. (1997a). Monte Carlo maximum likelihood estimation for non-Gaussian state space models. *Biometrika*, **84**, 669–684.
- Durbin, J. and Koopman, S.J. (1997b). Time series analysis of Non-Gaussian observations based on state space models. Preprint, London School of Economics.
- Durbin, J. and Koopman, S.J. (2000). Time series analysis of Non-Gaussian observations based on state space models from both classical and Bayesian perspectives. *Journal of The Royal Statistical Society, Series B*, **62**, 3–56.
- Einstein, A. (1905). On the movement of small particles suspended in a stationary liquid demanded by the molecular-kinetic theory of heat. *Ann. Physics*, **17**.
- Elliot, G., Rothenberg, T.J. and Stock, J.H. (1996). Efficient tests for an autoregressive unit root. *Econometrica*, **64**, 813–836.

- Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance*. New York: Springer.
- Enders, W. (1995). *Applied Econometric Analysis*. New York: Wiley.
- Engle, R.F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. *Econometrica*, **50**, 987–1008.
- Engle, R.F. (2000). The econometrics of ultra-high-frequency data. *Econometrica*, **68**, 1–22.
- Engle, R.F. and Granger, C.W.J. (1987). Cointegration and error correction: Representation, estimation and testing. *Econometrica*, **55**, 251–276.
- Engle, R.F. and Yoo, B.S. (1987). Forecasting and testing in cointegrated systems. *Journal of Econometrics*, **35**, 143–159.
- Engle, R.F. and Russell, J.R. (1998). Autoregressive conditional duration: A new model for irregularly spaced transaction data. *Econometrica*, **66**, 1127–1162.
- Engle, R.F. and Manganelli, S. (1999). CAViaR: Conditional autoregressive value at risk by regression quantiles. Preprint, University of California, San Diego.
- Engle, R.F. and Patton, A.J. (2001). What good is a volatility model. *Quantitative Finance*, **1**, 237–245.
- Fernandes, M. and Gramming, J. (2005). A family of autoregressive conditional duration models. Preprint.
- Fernandez, C. and Stell, M. (1998). On Bayesian modelling of fat tails and skewness. *Journal of the American Statistical Association*, **93**, 359–371.
- Fox, R. and Taqqu, M. S. (1986). Large sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *The Annals of Statistics*, **14**, 517–532.
- Franses, P.H. (1998). *Time Series Models for Business and Economic Forecasting*. Cambridge: Cambridge University Press.
- Garman, M.B. and Klass, M.J. (1980). On the estimation of security price volatility from historical data. *Journal of Business*, **53**, 67–78.
- Geweke, J. and Porter-Hudak, S. (1983). The estimation and application of long memory time series models. *Journal of Time Series Analysis*, **4**, 221–238.
- Giot, P. and Laurent, S. (2004). Modelling daily value-at-risk using realized volatility

- and ARCH type models. *Journal of Empirical Finance*, **11**, 379–398.
- Glosten, L.R., Jagannathan, R. and Runkle, D. (1993). Relationship between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance*, **48**, 1779–1801.
- Gordy, M.B. (2000). A comparative anatomy of credit risk models. *Journal of Banking and Finance*, **26**, 1335–1353.
- Granger, C.W.J. (1969). Investigating causal relationships by econometric models and cross-spectral methods. *Econometrica*, **37**, 424–438.
- Granger, C. M. G. and Joyeux, R. (1980). An introduction to long memory time series models and fractional differencing. *Journal of Time Series Analysis*, **1**, 15–29.
- Granger, C.W.J. and Newbold, P.E. (1974). Spurious regression in econometrics. *Journal of Econometrics*, **2**, 111–120.
- Granger, C.W.J. and Andersen, A.P. (1978). *An Introduction to Bilinear Time Series Models*. Gottingen: Vandenhoeck and Ruprecht.
- Hamilton, J.D. (1990). Analysis of time series subject to change in regime. *Journal of Econometrics*, **45**, 39–70.
- Hamilton, J.D. (1994). *Time Series Analysis*. Princeton: Princeton University Press.
- Hannan, E. J. (1973). The asymptotic theory of linear time series models. *Journal of Applied Probability*, **10**, 130–145.
- Hannan, E.J. (1980). The estimation of the order of an ARMA process. *The Annals of Statistics*, **8**, 1071–1081.
- Hannan, E.J. (1982). Testing for autocorrelation and Akaike's criterion. In *Essays in Statistical Science*, special volume 19A of *Journal of Applied Probability*. The Applied Probability Trust, Sheffield, 403–412.
- Hansen, B.E. (1992). Efficient estimation and testing of cointegrating vectors in the presence of deterministic trends. *Journal of Econometrics*, **53**, 87–121.
- Hendry, D.F. and Juselius, K. (2000). Explaining cointegration analysis: Part I. *The Energy Journal*, **21**, 1–42.
- Hendry, D.F. and Juselius, K. (2001). Explaining cointegration analysis: Part II. *The Energy Journal*, **22**, 75–120.
- Hol, E. and Kopman, S.J. (2002). Stock index volatility forecasting with high fre-

- quency data. Discussion Paper N. 2002-068/4, Tinbergen Institute.
- Hosking, J. R. M. (1981). Fractional differencing. *Biometrika*, **68**, 165–176.
- Hsiao, C. (1979). Autoregressive modelling of Canadian money and income data. *Journal of the American Statistical Association*, **74**, 553–560.
- Hurst, H. E. (1951). Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, **16**, 770–799.
- Hurst, H. E. (1957). A suggested statistical model of time series that occur in nature. *Nature*, **180**, 494.
- Hurvish, C.M. and Tsai, C.L. (1989). Regression and time series model selection in small samples. *Biometrika*, **76**, 297–307.
- Jaquier, E., Polson, N.G. and Rossi, P.E. (1994). Bayesian analysis of stochastic volatility models (With Discussion). *Journal of Business and Economic Statistics*, **12**, 371–417.
- Jenkins, G.M. and Watts, D.G. (1968). *Spectral Analysis and its Applications*. San Francisco: Holden-Day.
- Joanes, D.N. and Gill, C.A. (1998). Comparing measures of sample skewness and kurtosis. *The Statistician*, **47**, 183–189.
- Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, **12**, 231–254.
- Johansen, S. (1994). The role of the constant and linear terms in cointegration analysis of nonstationary variables. *Econometric Reviews*, **13**, 205–229.
- Johansen, S. (1995). *Likelihood Based Inference in Cointegrated Vector Error Correction Models*. Oxford: Oxford University Press.
- Kim, S., Shephard, N. and Chib, S. (1998). Stochastic volatility: Likelihood inference and comparison with ARCH models. *Review of Economic Studies*, **85**, 361–393.
- Kim, C.-J. and Nelson, C.R. (1999). *State Space Models With Regime Switching*. Cambridge: The MIT Press.
- Koenker, R.W. and Basset, G.W. (1978). Regression quantiles. *Econometrica*, **46**, 33–50.
- Koop, G. (1992). Aggregate shocks and macroeconomic fluctuations: A Bayesian approach. *Journal of Applied Econometrics*, **7**, 395–411.

- Koopman, S.J., Harvey, A.C., Doornick, J.A. and Shephard, N. (1995). STAMP 5.0: Strucutural Time Series Analyser, Modeller and Predictor. London: Timberlake.
- Kwitkowski, D., Phillips, P.C.B., Schmidt, P. and Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, **54**, 159–178.
- Laurent, S. and Peters, J.P. (2002). GARCH 2.3: An Ox package for estimating and forecasting various ARCH models. *Journal of Economic Surveys*, **16**, 447–485.
- Layton, A. P. (1984). A further note on the detection of Granger instantaneous causality. *Journal of Time Series Analysis*, **5**, 15–18.
- Ljung, G. and Box, G.E.P. (1978). On a measure of lack of fit in time series models. *Biometrika*, **65**, 297–303.
- Lo, A.W. (1991). Long term memory in stochastic market prices. *Econometrica*, **59**, 1279–1313.
- Longerstae, J. and More, L. (1995). Introduction to RiskMetrics<sup>TM</sup>, 4th Edition. New York: Morgan Guaranty Trust Company.
- Lopes, H.F. and Lima, E.C.R. (1995). *Co-integração: Enfoques Clássico e Bayesiano*. Vitória: Sexta Escola de Séries Temporais e Econometria.
- Lütkepohl, H. (1991). *Introduction to Multiple Time Series Analysis*. Heidelberg: Springer Verlag.
- Mandelbrot, B.B. and Van Ness, J.W. (1968). Fractional Brownian motions, fractional Brownian noises and applications. *SIAM Review*, **10**, 422–437.
- Mandelbrot, B.B. and Wallis, J. (1968). Noah, Joseph and operational hydrology. *Water Resources Research*, **4**, 909–918.
- Manteiga, S.M. (2002). Comparação de Metodologias para Estimção de Volatilidades para Cálculo do VaR- Valor-no-Risco e Modelagem de Perdas Não Previstas pelo VaR em Momentos de Crise. Dissertação de Mestrado Profissionalizante, FEA-IME, Universidade de São Paulo.
- Martens, M., Chang, Y.-C. and Taylor, S.J. (2002). A comparison of seasonal adjustment methods when forecasting intraday volatility. *The Journal of Financial research*, **25**, 283–299.
- McLeod, A. I. and Hipel, K. W. (1978). Preservation of the rescaled adjusted range, 1: A reassessment of the Hurst phenomenon. *Water Resources Research*, **14**, 491–508.

- Mendes, B.V.M. (2004). *Introdução à Análise de Eventos Extremos*. Rio de Janeiro: E-Papers Serviços Editoriais.
- Merton, R.C. (1973). Theory of option pricing. *Bell. J. Econ. Manag. Sci.*, **4**, 141–183.
- Mills, T.C. (1999). *The Econometric Modelling of Financial Time Series*. Second Edition. Cambridge University Press.
- Mittnik, S., Rachev, S.T. and Paoletta, M.S. (1998). Stable Paretian modeling in finance: Some empirical and theoretical aspects. In *A Practical Guide to Heavy Tails* (R.J. Adler, R.E Feldman and M.S. Taqqu, editors), p. 79–110. Boston: Birkhäuser.
- Morettin, P.A. (1984). The Levinson algorithm and its applications in time series analysis. *International Statistical Review*, **52**, 83–92.
- Morettin, P.A. and Toloi, C.M.C. (2006). *Análise de Séries Temporais*. Segunda Edição. São Paulo: Editora E. Blücher-Associação Brasileira de Estatística.
- Motta, A.C.O. (2001). Modelos de Espaço de Estados Não-Gaussianos e o Modelo de Volatilidade Estocástica. Dissertação de mestrado, IMECC-UNICAMP.
- Müller, U.A., Dacorogna, M.M., Davé, R.D., Olsen, R.B., Pictet, O.V. and von Weizsäcker, J.E. (1997). Volatilities of different time resolutions - Analyzing the dynamics of market components. *Journal of Empirical Finance*, **4**, 213–289.
- Nelson, D.B. (1991). Conditional heteroskedasticity in asset returns. *Econometrica*, **59**, 347–370.
- Nelson, D.B. and Cao, C.Q. (1992). Inequality constraints in the univariate GARCH model. *Journal of Business and Economic Statistics*, **10**, 229–235.
- Newey, W.K. and West, K.D. (1987). A simple positive semidefinite heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, **55**, 703–708.
- Ng, S. and Perron, P. (1995). Unit roots tests in ARMA models with data-dependent methods for the selection of the truncation lag. *Journal of the American Statistical Association*, **90**, 268–281.
- Oomen, R. (2001). Using high frequency stock market index data to calculate, model and forecast realized return volatility. Working Paper 2001/6, European University Institute.
- Osterwald-Lenum, M. (1992). A note with quantiles of the asymptotic distribution of maximum likelihood cointegration rank statistics. *Oxford Bulletin of Economics and Statistics*, **54**, 461–472.



- Pagan, A.R. and Schwert, G.W. (1990). Alternative models for conditional stochastic volatility. *Journal of Econometrics*, **45**, 267–290.
- Peña, D., Tsay, R.S. and Tiao, G.C. (2001). *A Course in Time Series Analysis*. New York: Wiley.
- Perron, P. (1988). Trends and random walks in macroeconomic time series: Further evidence from a new approach. *Journal of Economic Dynamics and Control*, **12**, 297–332.
- Perron, P. and Ng, S. (1996). Useful modifications to some unit root tests with dependent errors and their local asymptotic properties. *Review of Economic Studies*, **63**, 435–463.
- Phillips, P.C.B. (1986). Understanding spurious regression in econometrics. *Journal of Econometrics*, **33**, 311–340.
- Phillips, P.C.B. (1987). Time series regression with a unit root. *Econometrica*, **55**, 277–301.
- Phillips, P.C.B. (1991). Optimal inference in cointegrated systems. *Econometrica*, **59**, 283–306.
- Phillips, P.C.B. (1994). Model determination and macroeconomic activity. Cowles Foundation Discussion Paper # 1083, Yale University.
- Phillips, P.C.B. and Perron, P. (1988). Testing for unit roots in time series regression. *Biometrika*, **75**, 335–346.
- Phillips, P.C.B. and Ouliaris, S. (1990). Asymptotic properties of residual based tests for cointegration. *Econometrica*, **58**, 73–93.
- Phillips, P.C.B. and Shimotsu, K. (2004). Local Whittle estimation in nonstationary and unit root cases. *The Annals of Statistics*, **32**, 656–692.
- Pierce, D.A. and Haugh, L.D. (1977). Causality in temporal systems: Characterizations and a survey. *Journal of Econometrics*, **5**, 265–293.
- Quenouille, M.H. (1949). Approximate tests of correlation in time series. *Journal of the Royal Statistical Society, Series B*, **11**, 68–84.
- Rissanen, J. (1978). Modelling by shortest data description. *Automatica*, **14**, 465–471.
- Reiss, R.-D. and Thomas, M. (2001). *Statistical Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and Other Fields*. Second Edition.

Birkhäuser.

Rockafellar, R.T. and Uryasev, S. (2000). Optimization of conditional value-at-risk. *Journal of Risk*, **2**, 21–41.

Rockafellar, R.T. and Uryasev, S. (2002). Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance*, **26** 1443–1471.

Roll, R. (1984). A simple implicit measure of the effective bid-ask spread in an efficient market. *Journal of Finance*, **39**, 1127–1140.

Ruilova, J.C. (2007). Modelos ARCH Heterogêneos e Aplicações à Análise de Dados de Alta Frequência. Tese de Doutorado, IME-USP.

Sáfadi, T. and Morettin, P.A. (2000). A Bayesian analysis of threshold autoregressive moving average models. *Sankhya*, Series B, **62**, 353–371.

Said, S.E. and Dickey, D.A. (1985). Hypothesis testing in ARIMA (p,1,q) models. *Journal of the American Statistical Association*, **80**, 369–374.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, **6**, 461–464.

Schwert, W. (1989). Test for unit roots: A Monte Carlo investigation. *Journal of Business and Economic Statistics*, **7**, 147–159.

Seater, J. J. (1993). World temperature – Trend uncertainties and their implications for economic policy. *Journal of Business and Economic Statistics*, **11**, 265–277.

Shephard, N. and Pitt, M.K. (1997). Likelihood analysis of Non-Gaussian measurement time series. *Biometrika*, **84**, 653–667.

Shibata, R. (1976). Selection of the order of an autoregressive model by Akaike's information criterion. *Biometrika*, **63**, 117–126.

Smith, R.L. (1985). Maximum likelihood estimation in a class of non-regular cases. *Biometrika*, **72**, 67–90.

Stock, J.H. and Watson, M.W. (1988). Testing for common trends. *Journal of The American Statistical Association*, **83**, 1097–1107.

Stoll, H. and Whaley, R. (1990). Stock market structure and volatility. *Review of Financial Studies*, **3**, 37–71.

Subba Rao, T. (1981). On the theory of bilinear time series models. *Journal of the Royal Statistical Society*, Series B, **43**, 244–255.

- Subba Rao, T. and Gabr, M.M. (1984). *An Introduction to Bispectral Analysis and Bilinear Time Series Models*. Berlin: Springer.
- Taylor, S.J. (1980). Conjectured models for trend in financial prices tests as forecasts. *Journal of The Royal Statistical Society, Series B*, **42**, 338–362.
- Taylor, S.J. (1986). *Modeling Financial Time Series*. New York: Wiley.
- Taylor, S.J. and Xu, X. (1997). The incremental volatility information in one million foreign exchange quotations. *Journal of Empirical Finance*, **4**, 317–340.
- Tong, H. (1983). *Threshold Models in Non-Linear Time Series Analysis*. Lecture Notes in Statistics, **21**. Heidelberg: Springer
- Tong, H. (1990). *Non-Linear Time Series Models*. Oxford: Oxford University Press.
- Tsay, R.S. (2005). *Analysis of Financial Time Series*. Second Edition. New York: Wiley.
- Tsurumi, H. and Wago, H. (1994). A Bayesian analysis of unit root and cointegration with an application to a Yen–Dollar exchange rate model. Pre-print.
- Venables, W.N. and Ripley, B.D. (2001). *Modern Applied Statistics with S-Plus*. Second Edition. New York: Springer.
- Wiener, N. (1923). Differential space. *Journal of Mathematical Physics*, **2**, 131–174.
- Zakoian, J.M. (1994). Threshold heteroskedasticity models. *Journal of Economic Dynamics and Control*, **18**, 931–955.
- Zhang, M.Y., Russell, J.R. and Tsay, R.S. (2001). A nonlinear autoregressive conditional duration model with applications to financial transaction data. *Journal of Econometrics*, **104**, 179–207.
- Zivot, E. (2005). *Analysis of High Frequency Financial Data: Methods, Models and Software*. Minicurso para a 11<sup>a</sup> Escola de Séries Temporais e Econometria. Vila Velha, ES, Agosto de 2005.
- Zivot, E. and Wang, J. (2006). *Modelling Financial Time Series With SPLUS*. Second Edition. New York: Springer.