Econometria para Séries Financeiras

Fernando Antonio Lucena Aiube
State University of Rio de Janeiro

Econometria para Séries Financeiras

Fernando Antonio Lucena Aiube

Pontifícia Universidade Católica do Rio de Janeiro
Departamento de Engenharia Industrial

aiube@ind.puc-rio.br

Maio 2007

Introdução

A teoria de Finanças tem se desenvolvido muito na últimas décadas. O desenvolvimento de metodologias de apreçamento de derivativos fêz com que muita atenção fosse dedicada à estimação da volatilidade dos ativos subjacentes. Este parâmetro, que não diretamente observável, é fundamental para as metodologias de apreçamento. Por conseguinte muita pesquisa foi empreendida sobre os dados empíricos dos ativos financeiros. Isto fêz surgir a econometria para séries financeiras, uma nova disciplina que agrega as análises empíricas das séries financeiras e os testes econométricos de validação dos modelos desenvolvidos em Finanças.

Este trabalho apresenta alguns conceitos fundamentais para análise das séries de dados financeiros. Inicialmente são apresentados conceitos fundamentais das séries de tempo sob a ótima da metodologia Box e Jenkins. Posteriormente, são analisadas as séries de retornos e finalmente são tratados alguns dos modelos de volatilidade condicional. Ao final reunimos uma extensa lista de trabalhos que discorrem sobre os temas aqui abordados e certamente são úteis para o suporte ao estudo e pesquisa nesta área.

O curso é acompanhado de vários exercícios para a fixação de conceitos e simultâneamente são realizados exercícios práticos usando softwares onde os testes estatísticos estão implementados. Como exercício final o aluno deverá apresentar um modelo que descreva o comportamento de uma série financeira, baseado em dados históricos da mesma, devendo ser acompanhado dos testes empíricos de acordo com a metodologia apresentada durante o curso. Comentários e sugestões serão bem recebidos.

Sumário

Capítulo 1

Conceitos básicos de séries temporais

Neste primeiro capítulo apresentamos os conceitos fundamentais de séries temporais segundo a modelagem Box-Jenkins. O capítulo seguinte inicia a abordagem de séries temporais aplicadas a Finanças.

1.1 Conceitos fundamentais

Definição: Série temporal é qualquer conjunto de observações ordenado no tempo. A abordagem da análise pode ser no domínio do tempo com modelos paramétricos ou no domínio da freqüência com modelos não paramétricos. As séries temporais podem ser classificadas em:

(i) Discretas - quando o conjunto de observações for finito ou infinito enumerável;

(ii) Contínuas - quando o conjunto for infinito não enumerável;

(iii) Estocásticas - quando houver um componente aleatório;

(iv) Determinística - quando não houver componente aleatório e o modelo puder ser definido por funções determinísticas;

(v) Multivariadas - quando a série temporal é representada por um vetor;

(vi) Multidimensional - quanto t assume dimensão superior a 1.

Definição: Um processo estocástico X é uma coleção de variáveis aleatórias

$$(X_t,\ t \in T) = (X_t(\omega),\ t \in T,\ \omega \in \Omega)$$

definidas em algum espaço Ω.

Por exemplo, se a variável X_t representa a taxa de câmbio que está sendo negociada determinada moeda em um mercado, o valor de X está associado tanto ao instante de tempo t quanto a possíveis realizações ω. Assim, para um instante de tempo t fixo a variável aleatória é:

$$X_t = X_t(\omega), \quad \omega \in \Omega$$

Para um determinado estado da natureza, $\omega \in \Omega$, a variável aleatória é uma função do tempo:

$$X_t = X_t(\omega), \; t \in \; T$$

A função é denominada realização, trajetória ou caminho do processo de X.

Definição: Autocovariância γ_k - É a covariância entre duas variáveis da série defasadas por k intervalos de tempo, isto é:

$$\gamma_k = Cov\left(y_t, y_{t+k}\right) = E\left[\left(y_t - \mu\right)\left(y_{t+k} - \mu\right)\right] \tag{1.1}$$

onde μ é a média $\mu = E\left(y_t\right)$. Para uma amostra $y_1, y_2, \ldots, y_N$, temos o estimador de γ_k:

$$\hat{\gamma_k} = \frac{1}{N} \sum_{t=1}^{N-k} \left(y_t - \bar{y}\right)\left(y_{t+k} - \bar{y}\right) \tag{1.2}$$

onde $\bar{y} = \frac{1}{N}\sum_{t=1}^{N} y_t$ e $\hat{\gamma_k}$ é um estimador não tendencioso de (1.1).

Definição: Função de autocorrelação (FAC) é definida por

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{Cov\left(y_t, y_{t+k}\right)}{Var\left(y_t\right)}. \tag{1.3}$$

onde γ_0 é a variância da série. O estimador de ρ_k é

$$\hat{\rho_k} = \frac{\hat{\gamma_k}}{\hat{\gamma_0}} \tag{1.4}$$

Para testar a hipótese conjunta de que todos os ρ_k são simultâneamente iguais a zero pode-se usar a estatística Q desenvolvida por Box e Pierce, definida por:

$$Q(m) = N \sum_{k=1}^{m} \hat{\rho}_k^2\left(\hat{\epsilon}\right) \tag{1.5}$$

onde N é o tamanho da amostra e m a defasagem (ou *lag*) considerado. A estatística Q em grandes amostras é distribuído como uma qui-quadrado com

m graus de liberdade.

Outro teste usual para verificar a significância de ρ_k é o teste de Ljung-Box (LB). A estatística LB é dada por:

$$LB(m) = N\,(N+2)\sum_{k=1}^{m}\frac{\hat{\rho}_k^2\,(\hat{\epsilon})}{N-k} \tag{1.6}$$

que se distribui como uma qui-quadrado com m graus de liberdade em grandes amostras. A estatística LB possui mais poder estatístico para amostras pequenas que a estatística Q.

Definição: A função de auto-correlação parcial (FACP) é a correlação entre as variáveis y_t e y_{t+k} dado que são conhecidos y_{t+1},y_{t+2},$\ldots$, y_{t+k-1}.

A FACP para um processo estacionário com média zero pode ser obtido a partir da regressão

$$y_{t+k} = \phi_{k1}y_{t+k-1} + \phi_{k2}y_{t+k-2} + \ldots + \phi_{kk}y_t + \epsilon_{t+k}$$

Multiplicando ambos os lados por y_{t+k-j} e calculando o valor esperado e dividindo pela variância, tem-se

$$\rho_j = \phi_{k1}\rho_{j-1} + \phi_{k2}\rho_{j-2} + \ldots + \phi_{kk}\rho_{k-j}$$

Então para $j = 1, 2, \ldots, k$, temos:

$$\begin{aligned}
\rho_1 &= \phi_{k1}\rho_0 + \phi_{k2}\rho_1 + \ldots + \phi_{kk}\rho_{k-1} \\
\rho_2 &= \phi_{k1}\rho_1 + \phi_{k2}\rho_0 + \ldots + \phi_{kk}\rho_{k-2} \\
&\vdots \\
\rho_k &= \phi_{k1}\rho_{k-1} + \phi_{k2}\rho_{k-2} + \ldots + \phi_{kk}\rho_0
\end{aligned} \tag{1.7}$$

Para $k = 1 \Rightarrow \hat{\phi}_{11} = \rho_1$

Para $k = 2 \Rightarrow \rho_1 = \phi_{21} + \phi_{22}\rho_1$ e $\rho_2 = \phi_{21}\rho_1 + \phi_{22}$

Ou podemos escrever a última equação em notação matricial:

$$\begin{bmatrix}\rho_1\\ \rho_2\end{bmatrix} = \begin{bmatrix}1 & \rho_1\\ \rho_1 & 1\end{bmatrix}\begin{bmatrix}\phi_{21}\\ \phi_{22}\end{bmatrix}$$

cuja solução para o estimador $\hat{\rho}_{22}$ é dada pela regra de Cramer:

$$\hat{\phi}_{22} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}}$$

Para $k = 3$ temos as equações:

$$\rho_1 = \phi_{31} + \phi_{32}\rho_1 + \phi_{33}\rho_2$$

$$\rho_2 = \phi_{31}\rho_1 + \phi_{32} + \phi_{33}\rho_1$$

$$\rho_3 = \phi_{31} + \phi_{32}\rho_1 + \phi_{33}$$

Em notação matricial temos:

$$\begin{pmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \end{pmatrix} = \begin{pmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho 1 \\ \rho_2 & \rho_1 & 1 \end{pmatrix} \begin{pmatrix} \phi_{31} \\ \phi_{32} \\ \phi_{33} \end{pmatrix}$$

cuja solução para o estimador $\hat{\phi}_{33}$ é dada por:

$$\hat{\phi}_{33} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}}$$

E assim sucessivamente.

Definição: Seja $\{\epsilon_t\}$ uma seqüência de variáveis aleatórias independentes e identicamente distribuídas (iid) com variância σ_ϵ^2. Esta seqüência é denominada ruído branco (RB). Assim temos $\epsilon_t \sim$ iid tal que $E(\epsilon_t) = 0$; $Var(\epsilon_t) = \sigma_\epsilon^2$; $Cov(\epsilon_t, \epsilon_{t+k}) = \gamma_k = 0 \quad \forall k \neq 0$; $\rho_k = 1$ se $k = 0$, $\rho_k = 0$ se $k \neq 0$.

Definição: Considere ϵ_t um ruído branco, seja

$$y_t = y_{t-1} + \epsilon_t \tag{1.8}$$

O processo descrito pro y_t é definido como sendo um passeio aleatório (*random walk*).

Exercício 1: Seja y_t um passeio aleatório como definido na equação (1.8). Calcule $E(y_t)$, $Var(y_t)$,$Cov(y_t, y_s)$.

Em séries temporais é usual trabalhar com operadores que defasam a variável. Definimos então o operador *lag* L como um operador linear tal que:

$$L^i y_t = y_{t-i} \tag{1.9}$$

São válidas as seguintes propriedades do operador L:

(i) O *lag* de uma constante é a própria constante $Lc = c$

(ii) O operador *lag* segue a propriedade distributiva em relação à soma $(L^i + L^j) y_t = L^i y_t + L^j y_t = y_{t-i} + y_{t-j}$

(iii) É válida a propriedade associativa da multiplicação $L^i L^j y_t = L^i (L^j y_t) = L^i (y_{t-j}) = y_{t-i-j}$. Ou ainda $L^i L^j y_t = L^{i+j} y_t = y_{t-i-j}$

(iv) Potências negativas de L significam um operador de avanço, $L^{-i} y_t = L^j y_t$ fazendo $j = -i$. Então $L^{-i} y_t = L^j y_t = y_{t-j} = y_{t+i}$

(v) Se $|a| < 1$ a soma infinita $(1 + aL + a^2L^2 + \ldots) y_t = \frac{y_t}{1-aL}$

(vi) Se $|a| > 1$ a soma infinita $\left(1 + (aL)^{-1} + (aL)^{-2} + \ldots\right) y_t = -\frac{aL}{1-aL} y_t$

Exercício 2: Mostre a validade das propriedades (v) e (vi) acima do operador L.

Definição: Quando o processo estocástico que gerou a série de observações é invariante no tempo diz-se que é estacionário. Um processo é estritamente estacionário se as distribuições de y_t e y_{t+k} são idênticas (ou seja, a função densidade é a mesma). Um processo é estacionário de segunda ordem se a média e a variância de y_t são idênticas para qualquer t e covariância é função apenas da defasagem.

A motivação para o estudo de séries temporais é definir o processo gerador de dados, fazer previsões futuras da série, identificar ciclos, tendências ou sazonalidades de forma que a decisão que envolve a variável em questão seja a mais acurada possível.

1.2 Formulação dos modelos Box-Jenkins

A metodologia Box e Jenkins é a interpretação e análise de uma série temporal como sendo oriunda de uma realização de um processo estocástico. O objetivo é inferir sobre o processo gerador de dados. Busca-se identificá-lo baseado nas informações contidas na série levando-se em consideração a parcimônia do modelo, ou seja, tratando o modelo com o menor número de parâmetros possível. A estratégia envolve a repetição do processo de indenficação até encontrar o modelo que seja mais satisfatório.

Os modelos Box-Jenkins são tais que a série y_t é escrito como

$$\Phi_p(L)\, y_t = \Theta_q(L)\, \epsilon_t \tag{1.10}$$

onde L é o operador *lag*, Φ e Θ são polinôminos de graus p e q, respectivamente e ϵ_t é $RB\left(0, \sigma_\epsilon^2\right)$. Mais apropriadamente

$$\Phi_p(L) = 1 - \phi_1 L - \phi_2 L^2 - \ldots - \phi_p L^p \tag{1.11}$$

$$\Theta_q(L) = 1 - \theta_1 L - \theta_2 L^2 - \ldots - \theta_p L^p \tag{1.12}$$

O polinômio $\Phi_p(L)$ define a parte auto-regressiva (AR) do modelo enquanto o polinômio $\Theta_q(L)$ define a parte denominada média móvel (MA). Assim, o modelo em (1.10) é denominado $ARMA(p, q)$. Por exemplo, o modelo $ARMA(2, 3)$ é escrito como;

$$\begin{aligned}
\Phi_2(L)\, y_t &= \Theta_3(L)\, \epsilon_t \\
\left(1 - \phi_1 L - \phi_2 L^2\right) y_t &= \left(1 - \theta_1 L - \theta_2 L^2 - \theta_3 L^3\right) \epsilon_t \\
y_t - \phi_1 y_{t-1} - \phi_2 y_{t-2} &= \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \theta_3 \epsilon_{t-3} \\
y_t &= \phi_1 y_{t-1} + \phi_2 y_{t-2} - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \theta_3 \epsilon_{t-3} + \epsilon_t
\end{aligned}$$

No caso em que $\Theta_q(L) = 1$ temos o modelo $ARMA(p, 0)$ ou melhor $AR(p)$. Da mesma forma, para o caso em que $\Phi_p(L) = 1$ temos o modelo $ARMA(0, q)$ ou simplesmente $MA(q)$.

A condição de estacionariedade de um modelo $AR(p)$ deve ser tal que as raízes do polinômio $\Phi_p(L) = 0$ devem estar fora do círculo unitário. Para os modelos $MA(q)$ a estacionariedade é trivial. Para um modelo $ARMA(p, q)$ as condições de estacionariedade são aquelas de um modelo $AR(p)$.

A condição de inversibilidade de um modelo $AR(p)$ é trivial. Para um modelo $MA(q)$ a inversibilidade ocorre sempre que as raízes do polinômio

$\Theta_q(B) = 0$ estiverem fora do círculo unitário. Já um modelo $ARMA(p,q)$ tem a inversibilidade sob as mesmas condições de um $MA(q)$.

Pode-se resumir no quadro abaixo o comportamento dos modelos com relação a estacionariedade e inversiblidade.

Modelo	Condições
$\Phi_p(L)\, y_t = \epsilon_t$	$\Phi_p(L) = 0 \Rightarrow$ raízes fora do círculo unitário $\Rightarrow$ estacionário e trivialmente inversível.
$y_t = \Theta_q(L)\, \epsilon_t$	$\Theta_q(L) = 0 \Rightarrow$ raízes fora do círculo unitário $\Rightarrow$ inversível e trivialmente estacionário.
$\Phi_p(L)\, y_t = \Theta_q(L)\, \epsilon_t$	$\Phi_p(L) \ldots$ raízes fora do cíbulo unitário $\Rightarrow$ estacionário. $\Theta_q(L) \ldots$ raízes fora do círculo unitário $\Rightarrow$inversível.

Exercício 3: Considere o modelo $AR(1)$: $y_t = \phi y_{t-1} + \epsilon_t$, onde $|\phi| < 1$ e $\epsilon_t \sim RB(0, \sigma_\epsilon^2)$.

(i) o modelo é estacionário?

(ii) o modelo é inversível?

(iii) calcule a média $\mu = E(y_t)$

(iv) calcule $\gamma_0, \gamma_1, \ldots, \gamma_k$ e $\rho_0, \rho_1, \ldots, \rho_k$

(v) escreva o modelo sob a forma inversa.

Exercício 4: Considere o modelo $AR(1)$: $y_t = \beta + \phi y_{t-1} + \epsilon_t$ onde $|\phi| < 1$ e $\epsilon_t \sim RB(0, \sigma_\epsilon^2)$.

(i) calule a média $\mu = E(y_t)$

(ii) calcule a variância $\gamma_0 = Var(y_t)$

(iii) calcule as covariâncias $\gamma_1, \ldots, \gamma_k$

Exercício 5: Considere o modelo $MA(1)$: $y_t = \epsilon_t + \epsilon_{t-1}$, onde $\epsilon_t \sim RB(0, \sigma_\epsilon^2)$.

(i) calcule a média $\mu = E(y_t)$

(ii) calcule a variância $\gamma_0 = Var(y_t)$

(iii) calcule a FAC

(iv) o modelo é inversível?

Exercício 6: Suponha que a receita das vendas de petróleo R_t seja modelada pelo seguinte processo estocástico $R_t = \beta + R_{t-1} + \epsilon_t$ onde $\epsilon_t \sim N(0, \sigma^2)$. O que dizer da tendência da receita? Se o processo fosse modelado por $R_t = \beta + \varphi R_{t-1} + \epsilon_t$ onde $|\varphi| < 1$, você mudaria a sua resposta?

Exercício 7: Considere o modelo $MA(1)$: $y_t = \epsilon_t - \theta\epsilon_{t-1}$.

(i) calcule a média $\mu = E\left(y_t\right)$

(ii) calcule a variância $\gamma_0 = Var\left(y_t\right)$

(iii) calcule a FAC

Exercício 8: Considere o modelo

$$y_t = 0,8y_{t-1} - 0,3\epsilon_{t-1} + \epsilon_t$$

(i) verifique se é estacionário e inversível

(ii) calcule a média e a variância

(iii) calcule a FAC

(iv) escreva o modelo como um MA(∞)

Na prática os processos sobre os quais fazemos inferência através de uma série temporal, são geralmente não estacionários. Trataremos dos processos não estacionários homogêneos, ou seja, processos cuja a diferenciação produz processos estacionários. A diferenciação é definida por $\Delta y_t = y_t - y_{t-1} = \left(1 - L\right) y_t$.

Seja então Z_t um processo não estacionário e y_t um processo estacionário obtido de Z_t por diferenciação sucessivas. Inversamente, pode-se dizer que Z_t é obtido a partir de y_t por integração. Tem-se que

$$\Delta^d Z_t = y_t \tag{1.13}$$

onde d representa o número de diferenciações. O processo estacionário y_t pode ser representado por um modelo $ARMA(p, q)$, logo $\Phi_p(L)\, y_t = \Theta_q(L)\, \epsilon_t$, ou então

$$\Phi_p(L)\, \Delta^d Z_t = \Theta_q(L)\, \epsilon_t \tag{1.14}$$

Dizemos que o modelo Z_t é auto-regressivo-integrado-médias móveis, ou $ARIMA(p, d, q)$.

Exercício 9: Considere o modelo $y_t = 1,5y_{t-1} - 0.5y_{t-2} + \epsilon_t + 0,6\epsilon_{t-1}$.

(i) identifique o modelo

(ii) escreva o modelo como um $AR(\infty)$.

1.3 Identificação dos modelos

Para os modelos $ARMA(p, q)$ a estratégia de idenficação da ordem p e q mais apropriada é através da função de auto-correlação parcial FACP. Desta forma busca-se a idenficação do modelo comparando-se a FAC e FACP teóricas com aquelas oriundas do modelo.

As séries temporais que apresentam comportamento não estacionários são diferenciadas até que seja identificada a estacionariedade. Os testes de estacionariedade, comumente referidos como testes da raíz unitária, serão apresentados em detalhes no próximo capítulo. Os diversos padrões teóricos para as FAC e FACP são apresentados pelo software EVIEWS, junto com o correlograma das séries calculadas a partir dos dados. No caso de um ou mais modelos serem selecionados, baseados nos critérios da FAC e FACP, pode-se investigar qual o modelo apresenta melhor ajuste dentro da amostra. Infelizmente a medida R^2 não é útil para os modelos de séries temporais lineares por estar relacionado somente aos valores dos parâmetros. Os critérios de seleção mais apropriados são os critérios de informação de Akaike (1974) e Schwarz(1978). O critério de Akaike é referido como AIC (Akaike Information Criteria). O critério de Schwarz é referido como BIC (Bayesian Information Criteria). Estes critérios comparam o ajuste dentro da amostra, que é medido pela variância dos resíduos, contra o número de parâmetros estimados. O critério de Akaike é

$$AIC(k) = N \ln \hat{\sigma}^2 + 2k \tag{1.15}$$

ond e $k = p + q + 1$, $\hat{\sigma}_2 = \frac{1}{N}\sum_{t=1}^{N} \hat{\epsilon}_t^2$, sendo $\hat{\epsilon}_t$ os resíduos do modelo ARMA. Os valores de p e q que minimizam $AIC(k)$ são as ordens apropriadas do modelo $ARMA$.

O cirtério de Schwarz, ou ainda o critério BIC é calculado por

$$BIC\,(k) = N \ln \hat{\sigma}^2 + k \ln N \tag{1.16}$$

Como $\ln N > 2$ para $N > 8$, o critério BIC penaliza mais fortemente que o critério AIC a introdução adicional de parâmetros. Portanto, usando o critério BIC o modelo selecionado tende a ser mais parciomonioso que aquele orindo do critério AIC.

1.4 Função impulso resposta

A análise da estacionariedade de uma série pode ser feita pelos teste Dickey-Fuller (DF) e Augmented Dickey-Fuller (ADF) que serão detalhados no próximo capítulo. Vale ressaltar entretanto que através de técnicas de simulação pode-se verificar os efeitos que os choques causam em um modelo. Tomemos o caso do modelo $AR\,(p)$:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \epsilon_t \quad t = 1, 2, \ldots, N \tag{1.17}$$

Fazendo-se substituições recursivas otem-se para y_t

$$y_t = \sum_{i=1}^{p} y_{1-i} + \sum_{i=0}^{t-1} c_i \epsilon_{t-i} \tag{1.18}$$

onde a seqüência de $c_1, \ldots, c_{t-1}$ é obtida por

$$c_i = \sum_{j=1}^{min(i,p)} \phi_j c_{i-j} \quad i = 1, \ldots, t-1 \tag{1.19}$$

A seqüência c_k é denominada de função impulso reposta. Outra maneira de definir a função impulso resposta é por

$$\begin{aligned} IRF\,(k, \delta) = E\left[y_{t+k} | \epsilon_t = \delta, \epsilon_{t+1} = \ldots \epsilon_{t+k} = 0\right] - \\ E\left[y_{t+k} | \epsilon_t = \epsilon_{t+1} = \ldots \epsilon_{t+k} = 0\right] \end{aligned} \tag{1.20}$$

Isto é a função $IRF\,(k, \delta)$ mede o efeito de um choque δ em t em relação à situação de não existência do choque. Quando os choques são transientes o efeito do choque em t nos futuros valores de y_t desfaz-se rapidamente. Em um processo $AR\,(p)$ a presença dos choques transientes ocorre se raízes características estão fora do círculo unitário, ou seja, se o processo for estacionário. Para as situações onde observa-se que o efeito do choque não se desfaz, dizemos que o choque apresenta presistência. O software EVIEWS apresenta graficamente o efeito dos choques nas futuras observaões da série. Desta forma, a identificação visual torna claro o entendimento do conceito de estacionariedade.

Exercício 10: Mostre que a função impulso resposta para o modelo $AR(p)$ é dada por $IRF(k,\delta) = c_k\delta$.

1.5 Estratégia para especificação dos modelos

A modelagem de séries temporais por modelos lineares $ARMA$ deve seguir as seguintes etapas:

(i) cálculo de algumas estatísticas básicas para série temporal;

(ii) comparar o valor de tais estatísticas com valores teóricos caso estes sejam adequados;

(iii) estimar os parâmetros para o modelo sugerido no passo anterior, observando caso necessário, os critérios AIC e BIC;

(iv) avaliar o modelo usando as medidas diagnósticas;

(v) caso não esteja adequado reespecificar o modelo;

(vi) usar o modelo para descrever a variável e fazer previsões.

Como dito na seção anterior os modelos $ARIMA$ requerem previamente a identificação do parâmetro d (ordem de diferenciação) através de testes de estacionariedade.

Exercício 11: Considere os dados da série temporal do arquivo *teste.xls*. Analise o modelo segundo os critérios definidos na estratégia de especificação definida anteriormente.

1.6 Previsão

Seja agora a situação em que foi realizada a estimação do modelo e feitos os testes de adequação do mesmo. Desejamos realizar previsões para os valores das séries y_t. Considere que estamos no instante t (na origem) e desejamos fazer previsões ℓ períodos à frente. Esta previsão é denominada $\hat{y}_t(\ell) = E_t(y_t(\ell))$. Vale ressaltar que $E_t(\epsilon_{t+\ell}) = 0$ para $\ell \geq 1$. Aqui trataremos da previsão com erro quadrático mínimo. Para maiores detalhes veja Morettin e Toloi (2004).

Seja o modelo MA-1 dado por

$$y_t = \mu + (1 - \theta L)\,\epsilon_t \tag{1.21}$$

onde $\epsilon_t \sim NID\left(0, \sigma_\epsilon^2\right)$. O valor de y_t ℓ passos à frente é

$$y_t(\ell) = \mu + \epsilon_{t+\ell} - \theta\epsilon_{t+\ell-1} \tag{1.22}$$

Tomando o valor esperado dadas as informações até o instante t, temos

$$\hat{y}_t(\ell) = \mu + E_t(\epsilon_{t+\ell}) - \theta E_t(\epsilon_{t+\ell-1}) \tag{1.23}$$

Considere $\ell = 1$ para prever 1 passo à frente, então ficamos com

$$\hat{y}_t(1) = \mu + E_t(\epsilon_{t+1}) - \theta E_t(\epsilon_t) = \mu + 0 - \theta\epsilon_t$$

onde $\epsilon_t = y_t - \hat{y}_{t-1}(1)$

Para $\ell = 2$, faremos a previsão 2 passos à frente:

$$\hat{y}_t(2) = \mu + 0 - \theta \times 0 = \mu$$

E para $\ell > 2$ a situação acima repete-se. Deste modo podemos escrever:

$$\hat{y}_t(\ell) = \begin{cases} \mu - \theta_1\left(y_t - \hat{y}_{t-1}(1)\right) & \text{para } \ell < 2 \\ \mu & \text{para } \ell \geq 2 \end{cases}$$

A variância do erro de previsão é dada por

$$Var\left(e_t(\ell)\right) = E_t\left(e_t^2(\ell)\right)$$

onde $e_t(\ell) = y_t(\ell) - \hat{y}_t(\ell)$.

A variância do erro para $\ell = 1$ é dada por

$$Var\left(e_t(1)\right) = E_t\left(e_t^2(1)\right)$$

onde $e_t(1) = y_t(1) - \hat{y}_t(1)$. Ainda temos que

$$y_t(1) = \mu + \epsilon_{t+1} - \theta\epsilon_t \;\text{ e }\; \hat{y}_t(1) = \mu - \theta\epsilon_t$$

Logo temos $e_t(1) = \epsilon_{t+1}$, portanto $Var(e_t(1)) = \sigma_\epsilon^2$

A variância do erro para $\ell \geq 2$ é

$$Var(e_t(\ell)) = E_t\left[(y_t(\ell) - \hat{y}_t(\ell))^2\right]$$

Mas temos que $y_t(\ell) = \mu + \epsilon_{t+\ell} - \theta\epsilon_{t+\ell-1}$ e $\hat{y}_t(2) = \mu$. Logo teremos

$$y_t(2) - \hat{y}_t(2) = \epsilon_{t+\ell} - \theta\epsilon_{t+\ell-1}$$

Consequentemente a variância será

$$Var(e_t(\ell)) = E_t\left[(\epsilon_{t+\ell} - \theta\epsilon_{t+\ell-1})^2\right]$$

$$Var(e_t(\ell)) = E_t\left[\epsilon_{t+2} + \theta^2\epsilon_{t+1} - 2\theta\epsilon_{t+2}\epsilon_{t+1}\right]$$

$$\sigma_\epsilon^2 + \theta^2\sigma_\epsilon^2 = \left(1 + \theta^2\right)\sigma_\epsilon^2$$

Exercício 12: Seja o modelo AR-1

$$(1 - \phi L)\, y_t = \delta + \epsilon_t$$

onde $\epsilon_t \sim NID(0, \sigma_\epsilon^2)$. Calcule as previsões para $\ell = 1$, $\ell = 2$ e $\ell = \lambda$ passos à frente. Calcule as respectivas variâncias dos erros de previsão.

Capítulo 2

Séries financeiras - retornos

Este capítulo analisa o comportamento das séries financeiras. A questão da previsibilidade é tratada sob os diversos modelos de passeio aleatório (*random walk*) conforme a classificação de Campbell, Lo e MacKinlay (1997). Posteriormente analisamos os diversos testes nas séries financeiras identificando a dependência linear. A dependência não linear é tratada no capítulo seguinte.

2.1 Série de retornos

A maior parte dos estudos financeiros concentra-se na análise da série de retornos ao invés do uso da série de preços. A razão desta preferência, conforme Campbell et al (1997), está relacionada a dois fatos. Em primeiro lugar o retorno de um ativo financeiro contém as informações que atendem aos interesses dos investidores. Em segundo lugar a série de retornos possui propriedades estatisticamente mais atrativas que a série de preços.

O retorno de um ativo entre os instantes de tempo entre $t-1$ e t é dado por

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} \tag{2.1}$$

Ou ainda podemos escrever

$$R_t = \frac{P_t}{P_{t-1}} - 1 \Rightarrow 1 + R_t = \frac{P_t}{P_{t-1}}$$

O retorno em k períodos entre os intervalos $t-k$ e t é dado por

$$R_t\left(k\right) = \frac{P_t - P_{t-k}}{P_{t-k}} \Rightarrow 1 + R_t\left(k\right) = \frac{P_t}{P_{t-k}} = \frac{P_t}{P_{t-1}} \times \frac{P_{t-1}}{P_{t-2}} \times \ldots \times \frac{P_{t-k+1}}{P_{t-k}}$$

Ou ainda

$$1 + R_t(k) = \frac{P_t}{P_{t-k}} \tag{2.2}$$

O retorno capitalizado continamente significa que os instantes t e $t - \Delta t$ tornam-se muito próximos com Δt sendo infinitesimal. Neste caso $R_t << 1$. Definimos então o log-retorno como:

$$r_t = \ln(1 + R_t) = \ln\left(\frac{P_t}{P_{t-1}}\right) = \ln P_t - \ln P_{t-1} \cong R_t \tag{2.3}$$

O retorno multiperíodo capitalizado continuamente é dado por:

$$r_t(k) = \ln(1 + R_t(k)) = \ln\left[(1 + R_t)(1 + R_{t-1})\ldots(1 + R_{t-k+1})\right]$$

$$r_t(k) = \ln(1 + R_t) + \ln(1 + R_{t-1}) + \ldots + \ln(1 + R_{t-k+1})$$

$$r_t(k) = r_t + r_{t-1} + \ldots + r_{t-k+1}$$

2.2 Fatos estilizados

Definição: Os fatos estilizados são regularidades estatísticas observadas em um grande número de séries financeiras de retornos, a partir de estudos empíricos em diversos mercados. Pode-se resumir os principais fatos estilizados em:

(i) estacionariedade;

(ii) fraca dependência linear e dependência não linear;

(iii) caudas pesadas da distribuição ou excesso de curtose;

(iv) comportamento heterocedástico condicional.

O comportamento heterocedástico condicional reúne características como aglomerados de volatilidade e efeito alavanca. O efeito alavanca aponta para o efeito do comportamento dos choques. Choques negativos afetam a volatilidade condicional em maior magnitude que os choques positivos.

Os fatos estilizados serão tratados ao longo deste capítulo através de testes estatísticos, o capítulo seguinte dedica-se aos modelos heterocedásticos condicionais.

2.3 Distribuição dos retornos

A distribuição dos retornos escrita de forma mais geral envolve a análise das séries de retornos $\{r_{it}\}$ onde $i = 1, 2, \ldots, N$ e $t = 1, \ldots, T$. Em vários estudos tal como o $CAPM$ o foco é a análise seccional (*cross-section*) onde observam-se os retornos em um instante de tempo $\{r_{1t}, \ldots, r_{Nt}\}$. Para a análise do comportamento de um ativo específico teremos $\{r_{it}\}_{t=1}^{T}$. Iremos nos deter neste caso. É usual o tratamento do retorno como variáveis aleatórias contínuas e neste a função de densidade conjunta é dada por;

$$f(r_{i1}, r_{i2}, \ldots, r_{iT}; \theta) = f(r_{i1}) f(r_{i2}|r_{1t}) \ldots f(r_{iT}|r_{iT-1} \ldots r_{1i})$$

$$f(r_{i1}, r_{i2}, \ldots, r_{iT}; \theta) = f(r_{i1}|\theta) \prod_{t=2}^{T} f(r_{it}|r_{it-1} \ldots r_{i1}; \theta) \tag{2.4}$$

O aspecto relevante é a observação de como as distribuições do ativo evoluem no tempo, ou seja, a especificação da distribuição condicional. Por exemplo, uma das versões do passeio aleatório RW, que será vista adiante, pressupõe que a distribuição condicional é igual à distribuição incondicional $f(r_{it}|\cdot) = f(r_{it})$. Assim, os retornos são independentes e conseqüentemente não previsíveis.

Exercício 13: Uma consideração usual para a distribuição de retornos r_t é

$$r_t \sim NID\left(\mu, \sigma^2\right)$$

Portanto, R_t será uma distribuição log-normal iid. Calcule a média e a variância de R_t.

2.4 Previsibilidade dos retornos

A previsibilidade dos retornos é um fato intrigante em Finanças e ao qual muitos pesquisadores têm devotado atenção. Não há uma conclusão definitiva sobre o problema e a questão continua aberta ao debate acadêmico. A análise da previsibilidade é considerada diante das informações passadas dos retornos e consequentemente da distribuição dos retornos. Seguiremos a classificação de Campbell et al (1997). Esta classificação baseia-se nos vários tipos de passeio aleatório e na propriedade martingal. Assim temos os seguintes modelos para retorno $RW1$, $RW2$, martingal e $RW3$.

2.4.1 Modelo $RW1$

O modelo $RW1$ é a versão mais simples dentre as apresentadas acima e pressume que os retornos são normais e iid. Em outras palavras:

$$\ln P_t = \beta + \ln P_{t-1} + \epsilon_t \tag{2.5}$$

onde $\epsilon_t \sim NID\,(0, \sigma^2)$. Os logarítmos dos preços constituem um passeio aleatório com *drift*. Escrito de outra forma, temos:

$$r_t = \beta + \epsilon_t \tag{2.6}$$

Sob esta condição não há possibilidade nenhuma de previsão quer seja na média ou nos momentos superiores. Este modelo considera que a distribuição dos retornos é a mesma em qualquer instante t. O modelo $RW1$ contraria dois fatos estilizados quais sejam: a distribuição dos retornos não é normal e os retornos apresentam variância condicional variando com o tempo.

2.4.2 Modelo $RW2$

O pressuposto de que os retornos são iid no modelo $RW1$ é questionável como visto anteriormente. A hipótese de que os retornos são identicamente distribuídos é relaxada no modelo $RW2$. Porém a condição de independência é mantida. Então podemos escrever:

$$\ln P_t = \beta + \ln P_{t-1} + \epsilon_t \tag{2.7}$$

onde $\epsilon_t \sim INID\,(0, \sigma_t^2)$ onde $INID$ significa independente não identicamente distribuído. Ou ainda:

$$r_t = \beta + \epsilon_t \tag{2.8}$$

O modelo $RW2$ acomoda a possibilidade da variância ser diferente ao longo do tempo que é uma característica empírica das séries de retorno.

2.4.3 Modelo martingal

O modelo martingal está relacionado ao jogo justo onde considera-se que é impossível lucrar em um jogo dadas as informações passadas. Ou melhor, o processo martingal considera que a melhor previsão para o valor da variável aleatória amanhã é o seu valor hoje. Formalmente definimos o modelo martingal abaixo.

Definição: Seja $\{r_t\}_{t=1}^{T}$ um processo descrito pela variável aleatória r_t, dizemos que r_t é martingal com relação às informações por ele geradas se:

(i) $E\left(|r_t|\right) < \infty$

(ii) r_t contém todas as informações geradas pelo seu processo

(iii) $E\left(r_t | r_{t-1}, r_{t-2} \ldots\right) = r_{t-1}$

A condição (iii) estabelece que a previsão do valor de r_t dada as informações em $t-1$ é o seu valor em $t-1$, ou seja, r_{t-1}. Ainda podemos dizer que $E\left(r_t - r_{t-1} | r_{t-1}\right) = 0$. Esta é a interpretação do jogo justo onde os ganhos incrementais em qualquer instante de tempo, dada as informações passadas do jogo, é zero.

Exercício 14: Considere o modelo $r_t = \beta + \sigma_t \epsilon_t$ onde $\epsilon_t \sim NID\left(0,1\right)$ e $\sigma_t^2 = \omega_0 + \omega_1 r_{t-1}^2$. Mostre que o processo $r_t - \beta$ é uma diferença martingal.

2.4.4 Modelo $RW3$

O modelo $RW3$ relaxa a hipótese de independência do modelo $RW2$ considerando a dependência e a descorrelação dos incrementos. Assim, pode-se dizer que $Cov\left(\epsilon_t, \epsilon_{t-k}\right) = 0$, porém $Cov\left(\epsilon_t^2, \epsilon_{t-k}^2\right) \neq 0$. Este processo é descorrelatado mas não independente pois os quadrados dos resíduos são correlacionados.

Exercício 15: Seja o modelo $\ln P_t = \beta + \ln P_{t-1} + \nu_t$ onde $\nu_t = c\epsilon_{t-1}^2 + \epsilon_t$, onde $\epsilon_t \sim NID\left(0, \sigma^2\right)$, e c é constante. Mostre que:

(i) $Cov\left(\nu_t, \nu_{t-k}\right) = 0$

(ii) $Cov\left(r_t, r_{t-k}\right) = 0$

(iii) $E\left(\nu_t, \nu_{t-1}^2\right) \neq 0$

No exercício anterior o item (iii) implica que $E\left(R_t, R_{t-1}^2\right) \neq 0$, ou seja, o quadrado dos retornos defasados de uma unidade podem ser utilizados para construir a média condicional do processo.

2.5 Teste de normalidade

Como foi visto anteriormente, a suposição de que uma série de retornos possui distribuição normal é uma premissa para o Modelo $RW1$. Sabe-se entretanto que as séries de retornos não são normais. Em geral possuem caudas mais pesadas que a distribuição normal, isto é, excesso de curtose. Existem vários testes para verificação de normalidade. Iremos detalhar o teste de Jarque-Bera.

Seja X uma variável aleatória qualquer com média μ e variância σ^2. A assimetria de X é definida por:

$$A(X) = E\left[\frac{(X-\mu)^3}{\sigma^3}\right] \tag{2.9}$$

A curtose desta mesma distribuição é

$$K(X) = E\left[\frac{(X-\mu)^4}{\sigma^4}\right] \tag{2.10}$$

Para a distribuição normal $A = 0$ e $K = 3$. Os estimadores para a média e variância são:

$$\hat{\mu} = \frac{1}{N}\sum_{i=1}^{N} X_i \tag{2.11}$$

$$\hat{\sigma^2} = \frac{1}{N}\sum_{i=1}^{N} (X_i - \hat{\mu})^2 \tag{2.12}$$

Os estimadores para a assimetria e curtose são:

$$\hat{A} = \frac{\frac{1}{N}\sum_{i=1}^{N} (X_i - \hat{\mu})^3}{\hat{\sigma}^3} \tag{2.13}$$

$$\hat{K} = \frac{\frac{1}{N}\sum_{i=1}^{N} (X_i - \hat{\mu})^4}{\hat{\sigma^4}} \tag{2.14}$$

Demonstra-se que $\hat{A} \sim N\left(0, \frac{6}{N}\right)$ e $\hat{K} \sim N\left(3, \frac{24}{N}\right)$. Bera e Jarque (1981) utilizaram este fato para identificar se uma amostra provém de uma distribuição normal. A estatística do teste é

$$JB = \left(\frac{\hat{A} - 0}{\sqrt{\frac{6}{N}}}\right)^2 + \left(\frac{\hat{K} - 3}{\sqrt{\frac{24}{N}}}\right)^2 \tag{2.15}$$

que sob a hipótese nula de normalidade distribui-se como uma qui-quadrado com dois graus de liberdade.

Teste Jarque-Bera

As hipóteses nula e alternativa para o teste Jarque-Bera são:

$$H_0 : X_t \text{ é normal}$$

$$H_A : X_t \text{ não é normal}$$

Calcule a estatística JB usando as equações (2.13), (2.14) e (2.15). Obtenha da tabela o valor crítico $\tau = \chi^2_\alpha(2)$ onde α é o nível de significância. Rejeite a hipótese nula caso $JB > \tau$.

2.6 Teste de estacionariedade

O teste mais usual para verificação da estacionariedade é o teste da raíz unitária. Considere inicalmente o modelo

$$y_t = y_{t-1} + \epsilon_t \tag{2.16}$$

onde $\epsilon_t \sim \mathrm{RB}(0, \sigma^2)$. Já vimos que o processo X_t é um passeio aleatório. Se o coeficiente de y_{t-1} do processo que se está investigando é de fato 1 tem-se o que se chama de raíz unitária e fica caracterizada a não estacionariedade. Seja então

$$y_t = \rho y_{t-1} + \epsilon_t \tag{2.17}$$

A equação (2.17) pode ser expressa de outra forma como

$$\Delta y_t = y_t - y_{t-1} = (\rho - 1)\, y_{t-1} + \epsilon_t$$

ou ainda por

$$\Delta y_t = \delta y_{t-1} + \epsilon_t \tag{2.18}$$

onde $\delta = \rho - 1$ e a equação (2.18) define a série y_t diferenciada. Na série diferenciada a hipótese da raíz unitária é $\delta = 0$ e neste caso $\Delta y_t = \epsilon_t$, ou seja, a primeira diferença de y_t é RB que é estacionário. A série y_t é dita integrada de ordem 1.

Testes DF e ADF

Os testes da raíz unitária são conhecidos na literatura por Dickey-Fuller, seus autores, e ADF (Augmented Dickey e Fuller). O teste DF requer a verificação das regressões descritas abaixo:

$$\Delta y_t = \delta y_{t-1} + \epsilon_t \tag{2.19}$$

$$\Delta y_t = \beta_1 + \delta y_{t-1} + \epsilon_t \tag{2.20}$$

$$\Delta y_t = \beta_1 + \beta_2 t + \delta y_{t-1} + \epsilon_t \tag{2.21}$$

As hipóteses nulas para as regressões acima são respectivamente:

(i) $H_0 : \delta = 0$, X_t é um passeio aleatório

(ii) $H_0 : \delta = 0$, X_t é um passeio aleatório com *drift*

(iii) $H_0 : \delta = 0$, X_t é um passeio aleatório com *drift* e tendência

A hipótese alternativa para cada um dos casos acima é bilateral, $H_A : \delta \neq 0$. O software EVIEWS usa a hipótese alternativa $H_A : \delta < 0$.

Se o ruído ϵ_t é autocorrelacionado os testes anteriores devem ser modificados para

$$\Delta y_t = \beta_1 + \beta_2 t + \delta y_{t-1} + \sum_{i=1}^{m} \alpha_i \Delta y_{t-i} + \epsilon_t \tag{2.22}$$

Este é o teste ADF e a hipótese nula é a mesma, ou seja, $H_0 : \delta = 0$ ou $\rho = 1$.

Dickey e Fuller (1979) provaram que a estatística de teste não é a estatística-t convencional. Eles definiram os valoes críticos com base na simulação de Monte-Carlo. Mais recentemente Mackinnon (1991) apresentou valores críticos e p-valores para um espectro maior de cenários.

Teste Phillips-Perron (PP)

O teste PP utiliza uma correção na estatística de teste baseado em um ajuste não paramétrico na forma desta estatística, a qual corrige a presença de heterocedasticidade e/ou autocorrelação nos resíduos. As regressões são as mesmas descritas acima sem a presença do somatório do teste ADF. Os valores críticos permanecem os mesmos.

Exercício 15: Considere a série de preços e a sua respectiva série de retornos. Realize os testes de normalidade JB e os teste de raíz unitária ADF e PP.

2.7 Teste para autocorrelação

Já vimos anteriormente que os testes Box-Pierce e Ljung-Box são utilizados para detectar autocorrelação. Iremos detalhar os procedimentos para realizá-los.

Testes Box-Pierce e Ljung-Box

Primeiramente escolha o lag k para verificar a autocorrelação. As hipósteses dos testes são:

$$H_0 : \rho_1 = \rho_2 = \ldots = \rho_k = 0$$

$$H_A : \text{ pelo menos um } \rho \text{ não é nulo}$$

Calcule as estatísticas:

$$Q\left(k\right) = N \sum_{j=1}^{k} \hat{\rho}_j^2\left(\hat{\epsilon}\right) \tag{2.23}$$

$$LB\left(k\right) = N\left(N+2\right) \sum_{j=1}^{k} \left(N-j\right)^{-1} \hat{\rho}_j^2\left(\hat{\epsilon}\right) \tag{2.24}$$

As estatísticas em (2.23) e (2.24) distribuem-se como uma qui-quadrado com k graus de liberdade. Escolha o nível de significância α. Rejeite H_0 se as estatísticas acima forem superiors que o valor crítico $\tau = \chi^2_\alpha\left(k\right)$. Os testes acima são plenamente válidos se a distribuição é normal e estacionária. Em caso de uma (ou as duas) premissa(s) não se verificar(em) a potência do teste fica reduzida.

2.8 Teste de independência

O teste de independência ou dependência não linear mais apropriado para dados econômico-financeiros, que possuem amostras com tamanhos limitados, é o teste de Brock, Dechert, Scheinkman e LeBaron (1986). O teste envolve o conceito de correlação espacial. Para examinar a correlaçãoe espacial, a série y_t deve ser inserida no espaço-m construído pelo vetor:

$$y_t^m = [y_t, \ldots, y_{t-m+1}] \quad t = 1, 2 \ldots, N - m + 1$$

Por exemplo, seja o caso em que $t = 1, 2 \ldots, 5$ e $m = 3$. Os vetores criados são:

$$y_1^3 = [y_1, y_2, y_3]$$
$$y_2^3 = [y_2, y_3, y_4]$$
$$y_3^3 = [y_3, y_4, y_5]$$

Este operação permite verificar a inserção em um espaço de dimensão maior. A escolha da dimensão do espaço é subjetiva. A motivação é examinar a correlação no contexto espacial. A dependência de y_t é examinada através do conceito de correlação integral, uma medida que examina a distância entre pontos no espaço de dimensão 3, como no exemplo. Para cada dimensão m e um valor de ϵ, a correlação integral é definida por

$$C(\epsilon, m, N) = \frac{1}{N_m (N_m - 1)} \sum_{t \neq s} I(y_t^m, y_s^m, \epsilon) \tag{2.25}$$

onde $N_m = N - m + 1$, t e s variam entre 1 e $N - m + 1$. A função indicadora em (2.25) é

$$I(y_t^m, y_s^m, \epsilon) = \begin{cases} 1 & \text{se } \|y_t^m - y_s^m\| \leq \epsilon \\ 0 & \text{caso contrário} \end{cases}$$

A correlação integral irá medir a fração dos pares (y_t^m, y_s^m) para os quais a distância entre y_t^m ey_s^m não é superior a ϵ. Por exemplo, a distância entre y_1^3 e y_3^3 é calculada por $y_1^3 - y_3^3 = [y_1 - y_3, y_2 - y_4, y_3 - y_5]$. Escolhe-se o maior elemento do vetor resultante. Este maior elemento é comparado com ϵ. Se for maior que ϵ o par é contado, caso contrário não. Variando-se t e s, contamos os pares que satisfazem a condição. Então divide-se o número total de pares que satisfazem a condição pelo número total de pares. Isto leva à fração de pares que estão dentro da distância ϵ. Se todos os pares satisfizerem a condição $C(\epsilon, m, N) = 1$. Se nenhum satisfizer $C(\epsilon, m, N) = 0$.

Sob a hipótese nula de independência a estatística BDS é assintóticamente $N(0,1)$:

$$BDS(\epsilon, m, N) = \frac{\sqrt{N}\left[C(\epsilon, m, N) - C(\epsilon, 1, N)^m\right]^{\frac{1}{2}}}{V^{\frac{1}{2}}} \tag{2.26}$$

onde V é a variância. Observe que a estatística é função de ϵ e m. A aproximação assintótica é válida para $\frac{T}{m} > 200$. Quando em amostras pequenas este fato não é observado pode-se usar a técnica do *bootstrap* para gerar diversos cenários de p-valores. O software EVIEWS disponibiliza esta alternativa. Na prática deve-se testar a estatística para vários valores de m e ϵ.

Teste BDS

Escolha uma matriz de valores para m e ϵ. Considere ϵ variando de $0,5$ a 2 vezes o desvio padrão da série. Para cada valor da matriz realize o teste considerando

$$H_0 : y_t \text{ é independente (RW1)}$$
$$H_A : \text{dependência}$$

A hipótese alternativa de dependência descrita acima significa efetivamente dependência linear, dependência não linear (na média/ variância) ou caos determinístico. Esta última está fora do escopo deste trabalho. Caso estejamos procedendo o teste nos resíduos após o ajuste de um modelo AR à série, a hipótese alternativa ficará limitada à condição de dependência não linear na média, na variância ou ambas.

Calcule a estatística BDS em (2.26) e compare com os valores críticos para o nível de significância α. Rejeite H_0 caso $|\text{BDS}| > |\tau_\alpha|$.

Rejeitar a independência significa alguma forma de dependência. Para descartar a possibilidade de dependência linear, enfatizamos que a mesma deve ser filtrada através de um modelo autoregressivo. Neste caso a rejeição de H_0 limitará à situação de dependência não linear.

Exercício 16: Considere o processo média móvel não linear

$$y_t = \epsilon_t + 8\epsilon_{t-1}\epsilon_{t-2}$$

Simule uma série de tamanho $N = 1500$. Realize os testes Box-Pierce para verificar a correlação. Em seguida realize o teste BDS para testar a independência da série.

2.9 Teste para dependência não linear

Um fato estilizado e bem conhecido das séries de retornos financeiros é o comportamento heterocedástico. Esta constatação empírica pode ser feita pelo teste de McLeod e Li (1983). Este teste não é restrito às séries financeiras mas pode ser realizado para as séries de tempo em geral. Esta estatística realiza uma investigação idêntica àquela de Ljung e Box, agora considerando a correlação dos quadrados dos resíduos.

Teste de McLeod e Li

O teste é baseado no fato de que se $\rho_k = \rho_k^2$, para todo o k, então a série é linear, já a desigualdade define uma série não linear. Para realizá-lo inicialmente calcule a autocorrelação dos quadrados dos resíduos dada por

$$\hat{\rho}_k\left(\hat{\epsilon}^2\right) = \frac{\sum_{j=1}^{N-k}\left(\hat{\epsilon}_j^2 - \hat{\sigma}^2\right)\left(\hat{\epsilon}_{j+k}^2 - \hat{\sigma}^2\right)}{\sum_{j=1}^{N}\left(\hat{\epsilon}_j^2 - \hat{\sigma}^2\right)^2} \tag{2.27}$$

onde $\hat{\sigma}^2 = \frac{1}{N}\sum_{j=1}^{N}\hat{\epsilon}_j^2$. Considere as hipóteses:

$$H_0 : \text{quadrado dos resíduos descorrelatados}$$

$$H_A : \text{autocorrelação nos quadrados dos resíduos}$$

A estatística de teste é dada por:

$$McL\left(k\right) = N\left(N+2\right)\sum_{j=1}^{k}\left(N-j\right)^{-1}\rho_j^2\left(\hat{\epsilon}^2\right) \tag{2.28}$$

Quando aplicado aos resíduos de um processo $ARMA\left(p,q\right)$, a estatística McL tem uma distribuição assintótica $\chi^2\left(k-p-q\right)$. Para um nível de significância de α verifique o valor crítico $\tau = \chi_\alpha^2\left(k-p-q\right)$. Rejeite H_0 para $McL > \tau$. Uma vez rejeitada H_0 no teste de McLeod e Li há evidência de não linearidade.

Da mesma forma, uma vez que foi rejeitada a hipótese nula no teste BDS, após o ajuste de um processo AR, resta a condição de não linearidade na média e/ou na variância. O teste de Hsieh permite a diferenciação entre as duas situações.

Definição: Diz-se que existe dependência não linear na média se o valor esperado condicional dos resíduos em relação às informações passadas é diferente de zero, isto é:

$$E\left(\epsilon_t | y_{t-1}, \ldots, y_{t-k}, \epsilon_{t-1}, \ldots, \epsilon_{t-k}\right) \neq 0 \tag{2.29}$$

Exemplos de dependência não linear na média são os modelos média móvel não lineares, como por exemplo:

$$y_t = \epsilon_t + \beta\epsilon_{t-1}\epsilon_{t-2}$$

Outros modelos que apresentam não linearidade na média são os modelos TAR (*Threshold Autoregressive*) propostos por Tong e Lim (1980):

$$y_t = \begin{cases} \phi_1 y_{t-1} + \epsilon_t & \text{se } \; y_{t-1} \leq \delta \\ \phi_2 y_{t-1} + \epsilon_t & \text{se } \; y_{t-1} > \delta \end{cases}$$

onde ϵ_t é RB e y_t segue dois regimes dependendo do valor de y_{t-1}.

Ainda dentro da classe dos modelos não lineares na média podemos citar os modelos bilineares propostos por Granger e Andersen (1978):

$$y_t = \phi y_{t-1} + \beta y_{t-1}\epsilon_{t-1} + \epsilon_t$$

onde ϵ_t é RB.

Definição: Diz-se que existe dependência não linear na variância para o caso em que o valor esperado de ϵ_t é zero:

$$E\left(\epsilon_t | y_{t-1}, \ldots, y_{t-k}, \epsilon_{t-1}, \ldots, \epsilon_{t-k}\right) = 0 \tag{2.30}$$

Os exemplos de não linearidade na variância são os modelos de volatilidade condicional heterocedástica, ou modelos da família GARCH mais apropriadamente.

O teste de Hsieh (1989) considera a hipótese nula de não linearidade na variância, ou seja, o coeficiente de autocorrelação de terceira ordem é zero. A hipóstese alternativa é que o mesmo seja zero. O teste para a autocorrelação de terceira ordem é dado por

$$V_{ij} = \frac{\sqrt{N}\rho_{ij}^{(3)}}{W_{ij}} \tag{2.31}$$

onde os termos da equação (2.31) são:

$$\rho_{ij}^{(3)} = \frac{E\left(y_t, y_{t-i}, y_{t-j}\right)}{\left[E\left(y_t^2\right)\right]^{\frac{3}{2}}}$$

$$W_{ij} = \frac{E\left(y_t^2 y_{t-i}^2 y_{t-j}^2\right)}{\left[E\left(y_t^2\right)\right]^3}$$

Uma vez ajustado um modelo linear aos dados, a não linearidade estará contida nos resíduos. Ou seja, o teste deve ser realizado nos resíduos. Os estimadores para $\rho_{ij}^{(3)}$ e W_{ij} são:

$$\hat{\rho}_{ij}^{(3)} = \frac{\frac{1}{N}\sum_{t=1}^{N} \epsilon_t \epsilon_{t-i} \epsilon_{t-j}}{\left[\frac{1}{N}\sum_{t=1}^{N} \epsilon_t^2\right]^{\frac{3}{2}}} \tag{2.32}$$

$$\hat{W}_{ij} = \frac{\frac{1}{N}\sum_{t=1}^{N} \epsilon_t^2 \epsilon_{t-i}^2 \epsilon_{t-j}^2}{\left[\frac{1}{N}\sum_{t=1}^{N} \epsilon_t^2\right]^{3}} \tag{2.33}$$

A estatística V_{ij}, sob a hipótese nula de que o coeficiente de autocorrelação de terceira ordem é nulo, é normal.

Teste de Hsieh

Após ajustar um modelo AR que remova a dependência linear da série y_t, proceda a análise dos resíduos da série como se segue.

Inicialmente considere $i = 1$ e $j = 1$. Calcule os valores de $\rho^{(3)}$ e W_{ij} conforme as equações (2.32) e (2.33). Avalie a estatística V_{ij} conforme equação (2.31). As hipóteses do teste são:

$$H_0 : \text{dependência não linear na variância}$$

$$H_A : \text{dependência não linear na média}$$

Rejeite H_0 se $|V_{ij}| > \tau_\alpha$, onde τ_α é o valor crítico em uma distribuição normal padrão com nível de significância α. Não há muita clareza para quantos valores de i, j o teste deve ser realizado. Sugerimos que o procedimento seja considerado para $i = 1, 2, 3$ e $j = 1, 2, 3$. Desta forma, serão obtidas as tabelas seguintes com os valores das estatísticas V_{ij} e os respectivos p-valores:

Tabela 2.1: Estatística V_{ij}

i/j	1	2	3
1	V_{11}	V_{12}	V_{13}
2	V_{21}	V_{21}	V_{23}
3	V_{31}	V_{32}	V_{33}

Tabela 2.2: p-valores de V_{ij}

i/j	1	2	3
1	p_{11}	p_{12}	p_{13}
2	p_{21}	p_{21}	p_{23}
3	p_{31}	p_{32}	p_{33}

Exercício 17: Verifique se os modelos

(i) $y_t = \epsilon_t + \beta\epsilon_{t-1}^2 \quad \epsilon_t \sim NID\,(0, \sigma^2)$

(ii) $y_t = \epsilon_t h_t^{\frac{1}{2}}$ com $h_t = \alpha_0 + \alpha_1 y_{t-1}^2$ e $\epsilon_t \sim NID\,(0, \sigma^2)$

apresentam não linearidade na média ou na variância.

Capítulo 3

Volatilidade condicional

Na teoria de Finanças a incerteza ocupa um espaço preponderante. O risco, que quantifica a incerteza, está presente em muitos modelos dentre os quais podemos mencionar o CAPM (*Capital Asset Pricing Model*). A volatilidade como medida de risco é um dos parâmetros de maior relevância no apreçamento de opções. É uma variável não observável diretamente. Além disso, está relacionada a algumas propriedades ou a alguns fatos estilizados que são bem conhecidos. Pode-se citar, por exemplo, que a volatilidade em séries financeiras não é constante ao longo do tempo, responsável portanto pelo seu comportamento heterocedástico. Períodos de alta volatilidade são seguidos por períodos de alta volatilidade. Já aos períodos de baixa volatilidade seguem-se períodos amenos. Isto confere a propriedade a que a litereatura se refere como aglomerados de volatilidade.

Estas características peculiares da volatilidade são capturadas pelos modelos heterocedásticos condicionais ARCH (*Autoregressive Conditional Heterocedasticity*) proposto por Engel (1982) e extendido por Bollerslev (1986) então denominado de GARCH (*Generalized Autoregressive Conditional Heterocedasticity*). Não obstante, algumas propriedades do comportamento da volatilidade ficaram ao largo dos modelos GARCH. Por exemplo, o seu comportamento assimétrico não é capturado pelos modelos GARCH. Esta assimetria refere-se ao comportamento da volatilidade frente aos diferentes efeitos de choques positivos ou negativos. Os choques negativos trazem maior impacto à volatilidade. Estas constatações trouxeram novos modelos dentro da categoria de modelos GARCH e foram denominados GARCH não lineares. Assim é que a pesquisa no final da década de 80 e início dos anos 90 foi profícua em tais modelos. Além dos modelos de volatilidade, acima mencionados, será apresentado o modelo de volatilidade estocástica.

Até o momento os modelos estuados eram da forma

$$y_t = E\left(y_t|I_{t-1}\right) + \nu_t \tag{3.1}$$

onde ν_t é tal que é homocedástico condicional e incondicionalmente, isto é

$$E\left(\nu_t^2\right) = E\left(\nu_t^2|I_{t-1}\right)$$

Agora admitiremos que a variância condicional varie com o tempo, considerando que

$$E\left(\nu_t^2|I_{t-1}\right) = h_t \tag{3.2}$$

onde h_t representa a variância do resíduo instante t. Assim o modelo em (3.1) torna-se

$$y_t = E\left(y_t|I_{t-1}\right) + h_t^{\frac{1}{2}}\epsilon_t \tag{3.3}$$

onde ϵ_t é iid com média zero e variância 1. Em princípio iremos adotar $\epsilon_t \sim NID\left(0,1\right)$. Assim podemos dizer que $\nu_t|I_{t-1} \sim N\left(0,h_t\right)$.

3.1 Modelos de volatilidade condicional lineares

Esta seção trata de alguns modelos da família GARCH que foram denominados GARCH lineares por não capturarem os efeitos assimétricos dos choques. Tratam-se pricipalemente dos modelos clássicos propostos por Engle (1982) e Bollerslev (1986).

3.1.1 Modelo ARCH

O modelo de Engle (1982) considera que a volatildade em t é uma função linear do quadrado do choque em $t-1$, assim denominado ARCH(1).

$$y_t = h_t^{\frac{1}{2}}\epsilon_t \tag{3.4}$$

$$h_t = \omega + \alpha y_{t-1}^2$$

onde $\omega > 0$ e $\alpha \geq 0$ são condições que garantem a positividade de h_t. Mais genericamente pode-se considerar o modelo ARCH(1) escrito como na equação (3.1)

$$y_t = a_0 + a_1 y_{t-1} + \nu_t \tag{3.5}$$

$$\nu_t = h_t^{\frac{1}{2}}\epsilon_t$$
$$h_t = \omega + \alpha\nu_{t-1}^2$$

onde ϵ_t e ν_{t-1} são independentes.

O modelo ARCH(1) pode ser analisado sob a ótica de um processo autoregressivo em ν_t^2. Somando e subtarindo ν_t^2 na variância h_t, temos:

$$h_t + \nu_t^2 = \omega + \alpha\nu_{t-1}^2 + \nu_t^2$$
$$\nu_t^2 = \omega + \alpha\nu_{t-1}^2 + u_t \tag{3.6}$$

onde $u_t = \nu_t^2 - h_t = h_t\epsilon_t^2 - h_t = h_t\left(\epsilon_t^2 - 1\right)$. Ainda $E\left(u_t|I_{t-1}\right) = E\left(h_t\right)E\left(\epsilon_t^2 - 1\right) = 0$. Em (3.6) o processo para ν_t^2 é estacinário de segunda ordem se $\alpha < 1$. Neste caso os momentos em t e $t-1$ são iguais, ou seja

$$E\left(\nu_t^2\right) = \omega + \alpha E\left(\nu_{t-1}^2\right)$$
$$E\left(\nu_t^2\right) = \frac{\omega}{1-\alpha} \tag{3.7}$$

Ainda note que $E\left(\nu_t\right) = 0$. Observe também em (3.4) que valores grandes de y_{t-1} (positivos ou negativos) são seguidos por valores grandes de y_t. O mesmo ocorre para valores pequenos. Ou seja, o modelo captura os aglomerados de volatilidade. Além disso, a curtose de y_t é dada por

$$K_y = \frac{3\left(1-\alpha^2\right)}{1-3\alpha^2} \tag{3.8}$$

onde $1 > 3\alpha^2$ e K_y será maior que 3 o que significa que o modelo captura o excesso de curtose, um dos fatos estilizados das séries de retornos financeiros.

Exercício 18: Mostre que para o modelo em (3.4) que a curtose é dada por

$$K_y = \frac{3\left(1-\alpha^2\right)}{1-3\alpha^2}$$

Exercício 19: Seja o modelo dado em (3.5):

(i) calcule a média condicional de y_t

(ii) calcule a variância condicional de y_t

(iii) calcule a média incondicional de y_t

(iv) calcule a variância incondicional de y_t

Exercício 20: Seja o modelo em (3.4).

(i) escreva o modelo AR(1) para y_t^2

(ii) calcule a FAC para y_t^2

Uma extensão natural do modelo ARCH(1) em (3.4) é o modelo ARCH(q) onde a variância é escrita como:

$$h_t = \omega + \alpha_1 y_{t-1}^2 + \alpha_2 y_{t-2}^2 + \ldots + \alpha_q y_{t-q}^2 \tag{3.9}$$

Da mesma forma que anteriormente o modelo pode ser escrito como um AR(q) para y_t^2. Neste caso a variância incondicional de y_t será:

$$E\left(y_t^2\right) = \frac{\omega}{1 - \alpha_1 - \ldots - \alpha_q} \tag{3.10}$$

uma vez que sejam atendidas as condições de estacionariedade de segunda ordem.

Exercício 21: Mostre que a variância incondicional de y_t é dada pela equação (3.10).

Para trabalhar com os modelos ARCH modelando a volatilidade condicional das séries de retorno deve-se utilizar grandes valores de q o que torna os modelos poucos parciomoniosos trazendo complexidade para a estimação dos parâmetros. Esta complexidade é oriunda das restrições que se deve impor aos parâmetros evitando a não negatividade da variância e buscando a estacionariedade do modelo.

3.1.2 Modelo GARCH

Bollerslev (1986) propôs o modelo GARCH (*Generalized Autoregressive Conditional Hetercedasticity*) através da inclusão da variância passada ao modelo ARCH. O objetivo foi o de obter um modelo mais parciomonioso e sem os problemas de estimação do modelo ARCH. Seja então o modelo na forma da equação (3.4) para o modelo GARCH(1,1):

$$y_t = h_t^{\frac{1}{2}} \epsilon_t \tag{3.11}$$

$$h_t = \omega + \alpha y_{t-1}^2 + \beta h_{t-1}$$

onde $\omega > 0$, $\alpha > 0$ e $\beta \geq 0$ garantindo que $h_t \geq 0$. Fazendo substituições recursivas do termo h_{t-1} em (3.11) mostra-se a equivalência deste modelo com o modelo ARCH(∞).

Acrescentando y_t^2 em ambos os lados da expressão de h_t temos que

$$y_t^2 = \omega + \alpha y_{t-1}^2 + \beta h_{t-1} - h_t + y_t^2$$

$$y_t^2 = \omega + \alpha y_{t-1}^2 + \beta h_{t-1} + u_t$$

onde $u_t = y_t^2 - h_t$ e $u_{t-1} = y_{t-1}^2 - h_{t-1}$, fazendo a substituição

$$y_t^2 = \omega + \alpha y_{t-1}^2 + \beta \left(y_{t-1}^2 - u_{t-1} \right) + u_t$$

$$y_t^2 = \omega + (\alpha + \beta) \, y_{t-1}^2 - \beta u_{t-1} + u_{t-1} \tag{3.12}$$

E o processo GARCH(1,1) pode ser escrito como um ARMA(1,1) que será estacionário de segunda ordem se $\alpha + \beta < 1$.

Exercício 22: Seja o modelo descrito em (3.11)

(i) calcule a média e a variância condicionais de y_t

(ii) calcule a média e a variância incondicionais de y_t

(iii) calcule a autocorrelação ρ_1 do modelo em (3.12)

Exercício 23: Mostre que a curtose de y_t por

$$K_y = \frac{3 \left[1 - (\alpha + \beta)^2 \right]}{1 - (\alpha + \beta)^2 - 2\alpha^2}$$

O modelo GARCH(p,q) é dado por

$$h_t = \omega + \sum_{i=1}^{q} \alpha_i y_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j} \tag{3.13}$$

Alternativamente o modelo pode ser escrito como:

$$h_t = \omega + \alpha\left(L\right) y_t^2 + \beta\left(L\right) h_t \tag{3.14}$$

onde

$$\alpha\left(L\right) = \alpha_1 L + \ldots + \alpha_q L^q$$

$$\beta\left(L\right) = \beta_1 L + \ldots + \beta_p L^p$$

O modelo em (3.13) ou (3.14) será estacionário de segunda ordem caso as raízes do polinômio $1 - \alpha\left(L\right) - \beta\left(L\right)$ estiverem fora do círculo unitário. A seleção da ordem p, q do modelo deve ser feita minimizando os critérios de informação tais como o AIC (Akaike Information Criteria) e BIC (Bayesian Information Criteria) nas equações (1.15) e (1.16). Na maior parte dos casos práticos o modelo GARCH(1,1) atende as necessidades de modelagem.

3.1.3 Estimação do modelo GARCH(1,1)

O modelo GARCH(1,1) na equação (3.11) está aqui reescrito

$$y_t = h_t^{\frac{1}{2}} \epsilon_t \tag{3.15}$$

$$h_t = \omega + \alpha y_{t-1}^2 + \beta h_{t-1}$$

onde são observadas as mesmas restrições dos parâmetros e $\epsilon_t \sim NID\left(0, 1\right)$. A função distribuição conjunta do modelo é dada por

$$f\left(y_1, y_2, \ldots, y_N\right) = f\left(y_1\right) f\left(y_2 | y_0, y_1\right) \ldots f\left(y_N | y_0, \ldots, y_{N-1}\right)$$

$$f\left(y_1, y_2, \ldots, y_N\right) = f\left(y_1\right) \prod_{t=2}^{N} f\left(y_t | y_0, \ldots, y_{t-1}\right) \tag{3.16}$$

A função verossimilhança do modelo é

$$L\left(\Theta; y\right) = \ln f\left(y_1, y_2, \ldots, y_N\right) = \ln\left(y_1\right) + \sum_{t=2}^{N} \ln f\left(y_t | y_0, \ldots, y_{N-1}\right)$$

onde Θ representa o vetor dos parâmetros $\Theta = [\omega, \alpha, \beta]$. E a função de verossimilhança dado y_1 será dada por:

$$L(\Theta; y) = \sum_{t=2}^{N} \ln f(y_t | y_0, \ldots, y_{N-1}) \tag{3.17}$$

Por outro lado temos que a função densidade de ϵ_t é

$$f(\epsilon_t | y_{t-1}) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\epsilon_t^2\right)$$

Da equação (3.15) temos que

$$g(y_t) = \frac{y_t}{h_t^{\frac{1}{2}}} \quad \text{e} \quad g\prime(y_t) = \frac{1}{h_t^{\frac{1}{2}}} \tag{3.18}$$

E a função $f(y_t | y_0, \ldots, y_{N-1})$ será escrita por

$$f(y_t | y_0, \ldots, y_{N-1}) = f(g(y_t))\, g\prime(y_t) \tag{3.19}$$

Usando a equação (3.18) na equação (3.19), teremos:

$$f(y_t | y_0, \ldots, y_{N-1}) = f\left(\frac{y_t}{h^{\frac{1}{2}}}\right) \frac{1}{h_t^{\frac{1}{2}}} \tag{3.20}$$

Levando a equação (3.20) na equação (3.17), temos

$$L(\Theta; y) = \sum_{t=2}^{N} \ln f\left(\frac{y_t}{h_t^{\frac{1}{2}}}\right) + \sum_{t=2}^{N} \ln h_t^{-\frac{1}{2}} \tag{3.21}$$

Mas $f\left(y_t / h_t^{1/2}\right) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\frac{y_t^2}{h_t}\right)$ e o seu logarítmo é

$$\ln f\left(\frac{y_t}{h_t^{\frac{1}{2}}}\right) = -\frac{1}{2}\ln 2\pi - \frac{1}{2}\frac{y_t^2}{h_t}$$

Levando este resultado em (3.21), finalmente teremos:

$$L(\Theta; y) = \sum_{t=2}^{N} \left(-\frac{1}{2}\ln 2\pi - \frac{1}{2}\left(\frac{y_t^2}{h_t}\right)\right) - \frac{1}{2}\sum_{t=2}^{N} \ln h_t$$

$$L(\Theta; y) = -\frac{N}{2}\ln 2\pi - \frac{1}{2}\sum_{t=2}^{N} \ln h_t - \frac{1}{2}\sum_{t=2}^{N} \frac{y_t^2}{h_t} \tag{3.22}$$

Os parâmetros em Θ são obtidos pela maximização da função $L(\Theta; y)$ em (3.22).

A suposição em (3.15) de que $\epsilon_t \sim NID(0, 1)$, pode ser modificada. Bollerslev (1987) sugere o uso de distribuição com caudas mais pesadas para capturar o excesso de curtose. Considerando uma variável aleatória x com distribuição t de Student com ν graus de liberdade e ainda $\epsilon_t = \frac{x}{\sqrt{\nu(\nu-2)}}$, a função densidade de ϵ_t é

$$f(\epsilon_t|\nu) = \frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2)\sqrt{(\nu-2)\pi}} \left(1 + \frac{\epsilon_t^2}{\nu-2}\right)^{-\frac{\nu+1}{2}} \tag{3.23}$$

para $\nu > 2$ onde $\Gamma(z) = \int_0^\infty y^{z-1} e^{-y} dy$. E seguindo as mesmas etapas anteriores chegaremos a

$$f(y_t|y_0, \ldots, y_{N-1}) = \prod_{t=2}^{N} \frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2)\sqrt{(\nu-2)\pi}} \frac{1}{h_t^{1/2}} \left[1 + \frac{y_t^2}{(\nu-2)h_t}\right] \tag{3.24}$$

Se o valor dos graus de liberdade fôr uma variável exógena tem-se para a função verossimilhança:

$$L(\Theta; y) = -\sum_{t=2}^{N} \left[\frac{\nu+1}{2} \ln\left(1 + \frac{y_t^2}{(\nu-2)h_t}\right) + \frac{1}{2}\ln h_t\right] \tag{3.25}$$

Se o número de graus de liberdade estiver sendo estimado, acrescente-se à função anterior a parcela que se segue ao produtório em (3.24).

3.1.4 Previsão da volatilidade condicional

Seja novamente o modelo em (3.15), a variância condicional em t é dada por

$$h_t = \omega + \alpha y_{t-1}^2 + \beta h_{t-1}$$

A variância um passo à frente é

$$h_{t+1} = \omega + \alpha y_t^2 + \beta h_t$$

O valor esperado $\hat{h}_{t+1}$ é

$$\hat{h}_{t+1} = \omega + \alpha \hat{y}_t^2 + \beta h_t = \omega + \alpha y_t^2 + \beta h_t = h_{t+1} \tag{3.26}$$

Para o instante $t + 2$, temos

$$h_{t+2} = \omega + \alpha y_{t+1}^2 + \beta h_{t+1}$$

Por conseguinte o valor esperado será

$$\hat{h}_{t+2} = \omega + \alpha \hat{y}_{t+1}^2 + \beta \hat{h}_{t+1} = \omega + \alpha \hat{h}_{t+1} + \beta h_{t+1} = \omega + (\alpha + \beta)\, h_{t+1}$$

onde foi usado o resulado da equação (3.26). Repetindo o procedimento em $t+3$, teremos

$$\hat{h}_{t+3} = \omega + \omega\,(\alpha + \beta) + (\alpha + \beta)^2\, h_{t+1}$$

E para $t+s$, $s \in \mathbb{Z}$, teremos

$$\hat{h}_{t+s} = \omega \sum_{i=0}^{s-2} (\alpha + \beta)^i + (\alpha + \beta)^{s-1}\, h_{t+1} \tag{3.27}$$

3.1.5 Modelo IGARCH

O modelo IGARCH (*Integrated* GARCH) é útil em se tratando de dados de alta freqüência. Neste modelo $\alpha + \beta = 1$. Retomando a equação (3.12), podemos escrever

$$(1 - L)\, y_t^2 = \omega - \beta u_{t-1} + u_t \tag{3.28}$$

O modelo não é estacionário de segunda ordem. Porém pode-se demonstrar que estritamente estacionário, veja Nelson (1990).

3.1.6 Modelo GARCH-M

O modelo GARCH-M (GARCH na média) foi proposto por Engle, Lilien e Robins (1997) com a finalidade de capturar a relação entre risco e retorno tão bem estabelecida dentro da teoria de Finanças. O modelo GARCH(1,1)-M pode ser assim descrito:

$$y_t = \mu + cg\,(h_t) + \nu_t \tag{3.29}$$

$$\nu_t = h_t^{1/2} \epsilon_t$$

$$h_t = \omega + \alpha \nu_{t-1}^2 + \beta h_{t-1}$$

onde $\mu,\ c,\ \omega,\ \alpha,$ e β são parâmetros. A função $g\,(h_t)$ em geral é $g\,(h_t) = h_t^{1/2}$ ou $g\,(h_t) = h_t$. O parâmetro c é o prêmio de risco. Para c positivo há indicação de que o retorno está positivamente correlacionado à volatilidade passada.

3.2 Modelos de volatilidade condicional não lineares

Os efeitos dos choques na volatilidade condicional são diferentes para choques positivos ou negativos. Este é um fato estilizado. O efeito na volatilidade condicional de um choque negativo é mais acentuado do que o de um choque positivo. Entretanto os modelos ARCH e GARCH nas equações (3.4) e (3.11), respectivamente consideram o choque em $t-1$ elevado ao quadrado. Neste caso, os modelos são indiferentes ao sinal no choque e o efeito constatado empiricamente não é capturado pelos modelos. Este efeito foi observado por Black (1976). Quando a ação de uma empresa cai a relação entre a dívida e o capital próprio aumenta sugerindo um aumento da alavancagem. Simultanearmente a ação fica mais volátil. Estes dois fatos ficaram associados e o fenômeno ficou conhecido como efeito alavanca. Os modelos GARCH não lineares que serão apresentados foram estabelecidos com a finalidade de capturar o efeito alavanca.

3.2.1 Modelo EGARCH

O modelo EGARCH (ou *exponential* GARCH) foi desenvolvido por Nelson (1991). O EGARCH(1,1) é descrito por:

$$\ln h_t = \omega + \alpha y_{t-1} + \gamma\left(|y_{t-1}| - E\left(|y_{t-1}|\right)\right) + \beta \ln h_{t-1} \tag{3.30}$$

que ainda pode ser escrito por

$$\ln h_t = \omega + g\left(y_{t-1}\right) + \beta \ln h_{t-1} \tag{3.31}$$

onde α, ω, β e γ são constantes e $g\left(y_t\right)$ é dada por

$$g\left(y_t\right) = \alpha\epsilon_t + \gamma\left(|y_t| - E\left(|y_t|\right)\right) \tag{3.32}$$

O uso do logarítmo da variância no modelo EGARCH flexibiliza as restrições de positividade imposta aos parâmetros. Vejamos os efeitos em h_t para choques positivos ou negativos em $t-1$:

$$g\left(y_{t-1}\right) = \begin{cases} \left(\alpha+\gamma\right) y_{t-1} - \gamma E\left(|y_{t-1}|\right) & \text{para } y_{t-1} > 0 \\ \left(\alpha-\gamma\right) y_{t-1} - \gamma E\left(|y_{t-1}|\right) & \text{para } y_{t-1} < 0 \end{cases}$$

Note que a assimetria dos choques ocorre se $\gamma \neq 0$ e presença do efeito alavanca ocorre quando $\gamma < 0$. A função $g\left(y_t\right)$ possui média zero pois tanto y_t como $|y_t| - E\left(|y_t|\right)$ possuem média zero. A assimetria dos choques é garantida pela especificação da função $g(y_t)$ em (3.32). Se $\epsilon_t \sim NID(0,1)$ então $E(|\epsilon|) = \sqrt{\frac{2}{\pi}}$. Se ϵ_t é uma distribuição t de Student padronizada dada em (3.23), então

$$E(|\epsilon_t|) = \frac{2\sqrt{\nu-2}\,\Gamma((\nu+1)/2)}{(\nu-1)\Gamma(\nu/2)\sqrt{\pi}}$$

3.2.2 Modelo TARCH

O modelo TARCH (*Threshold* ARCH) também foi concebido para considerar as diferenças na volatilidade condicional causadas por choques positivos e negativos. Na literatura por vezes este modelo ora aparece com o nome GJR devido a Glosten, Jagannathan e Runkle (1993) ora simplesmente TARCH devido a Zakoïan (1994). Essencialmente os dois modelos têm a mesma finalidade e aqui serão tratados indistintamente de TARCH. O modelo TARCH(1,1) é escrito por

$$h_t = \omega + \alpha y_{t-1}^2 + \beta h_{t-1} + \gamma y_{t-1}^2 (1 - I(y_{t-1} > 0)) \tag{3.33}$$

Se o choque em $t-1$ fôr positivo, $y_{t-1} > 0$, então $I(y_{t-1}) = 1$ e o impacto na variância será devido a α. Caso o choque seja negativo $I(y_{t-1} = 0)$ o impacto na variância será $\alpha + \gamma$. O efeito assimétrico fica caracterizado se $\gamma \neq 0$ e se $\gamma > 0$ fica constatado o efeito alavanca. O modelo mais geral como TARCH(p,q) pode ser escrito por

$$h_t = \omega + \sum_{i=1}^{q} \alpha_i y_{t-1}^2 + \sum_{j=1}^{p} \beta_j h_{t-j} + \sum_{k=1}^{r} \gamma_k y_{t-k}^2 (1 - I(y_{t-k} > 0)) \tag{3.34}$$

onde r representa a ordem do choque que impacta a volatilidade.

3.2.3 Modelo QGARCH

O modelo QGARCH (ou *quadratic* GARCH) também captura os efeitos de choques de diferentes sinais. Foi proposto por Sentana (1995) e pode ser escrito por

$$h_t = \omega + \gamma y_{t-1} + \alpha y_{t-1}^2 + \beta h_{t-1} \tag{3.35}$$

Deferencia do GARCH tradicional pela introdução do termo γy_{t-1}. O modelo em (3.35) pode ser escrito por

$$h_t = \omega + \left(\frac{\gamma}{y_{t-1}} + \alpha\right) y_{t-1}^2 + \beta h_{t-1} \tag{3.36}$$

Para $\gamma < 0$ os choques negativos causarão um impacto em h_t superior aos choques positivos. Neste modelo o tamanho do choque também é capturado como influenciando a variância.

Além dos modelos aqui apresentados existem muitos outros que tornam a família GARCH muito extensa. Citamos alguns outros modelos dentro dos GARCH não lineares:

(i) LSTGARCH (*Logistic Smooth Transition Garch*) - Enquanto no modelo TARCH o efeito do choque positivo para o negativo é devido à mudança abrupta de α para $\alpha + \gamma$, neste modelo há uma mudança suave de uma situação para outra através da função logística. Foi proposto por Hagerud (1997) e Gonzáles-Rivera (1998).

(ii) GARCH com mudança de regime - Outros modelos de volatilidade condicional levam em consideração a possibilidde de mudança de regime. Rabemananjara e Zakoïan (1993) argumetam que choques negativos aumentam a volatilidade condicional somente se o choque negativo (em valor absoluto) é grande em magnitude. Observaram que choques negativos e pequenos têm menor impacto sobre a volatilidade que choques positivos de magnitude igual. Nesta linha de trabalho podemos ainda citar Fornari e Melle (1996,1997) e Anderson, Nam e Vahid (1999).

3.3 Teste para GARCH linear

O teste para detectar heterocedasticidades condicional ou efeito ARCH nos resíduos de uma regressão foi proposto por Engle (1982). A volatilidade condicional será constante se todos os α_i do modelo ARCH(q) em (3.9), aqui reescrito

$$h_t = \omega + \sum_{i=1}^{q} \alpha_i y_{t-i}^2$$

forem nulos.

Teste ARCH-LM

Fazendo-se uma regressão em que

$$\hat{e}_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \hat{e}_{t-i}^2 + \nu_t \tag{3.37}$$

onde $\hat{e}_t$ são os resíduos estimados da regressão, podemos testar a hipótese nula:

$$H_0 : \alpha_1 = \alpha_2 = \ldots = \alpha_q = 0 \text{ (sem efeito ARCH)}$$

$$H_A : \text{presença do efeito ARCH}$$

O teste é baseado no princípio dos multiplicadores de Lagrange. A estatística LM é $LM = N \times R^2$ distribuindo assitoticamente como uma $\chi^2_\alpha(q)$ sendo α o nível de significância.

3.4 Teste para GARCH não linear

Os testes propostos por Engle e Ng (1993) verificam a presença do efeito assimétrico dos choques na volatilidade.

Teste do sinal do choque

Este teste verifica se magnitude do quadrado do choque em t é afetado pelo sinal do choque em $t-1$. Considere uma variável *dummy* N_{t-1} em que $N_{t-1} = 1$ se o choque em $t-1$ é negatio, isto é $y_{t-1} < 0$, e zero caso contrário. Faça a regressão

$$\hat{e}_t^2 = \alpha_0 + \alpha_1 N_{t-1} + \nu_t \tag{3.38}$$

Considere as hipóteses:

$$H_0 : \alpha_1 = 0$$

$$H_A : \alpha_1 \neq 0$$

Para $\alpha_1 = 0$ não existe assimetria ou efeito alavanca. A estatística de teste é a estatística t tradicional. Rejeite H_0 se estatística t $> t_\alpha(N-2)$ onde α é o nível de significância e N o tamanho da série.

Teste do tamanho do choque

Neste teste é verificado se além do sinal, o tamanho do choque em $t-1$ afeta o quadrado do choque em t. Faça a regressão

$$\hat{e}_t^2 = \alpha_0 + \alpha_1 N_{t-1}\hat{e}_{t-1} + \nu_t \tag{3.39}$$

Nesta regressão é investigado se o choque negativo e sua magnitude afetam $\hat{e}_t^2$ e consequentemente a variância condicional, o teste é analisado sob a estatística t.

Faça também a regressão

$$\hat{e}_t^2 = \alpha_0 + \alpha_1 P_{t-1}\hat{e}_{t-1} + \nu_t \tag{3.40}$$

onde $P_{t-1} = 1 - N_{t-1}$. Neste caso investiga-se se o sinal e a magnitude do choque positivo em $t-1$ afetam simultâneamente a variância condicional.

Também pode-se realizar o teste conjunto proposto em (3.38), (3.39) e (3.40). Faça a regressão

$$\hat{e}_t^2 = \alpha_0 + \alpha_1 N_{t-1} + \alpha_2 N_{t-1}\hat{e}_{t-1} + \alpha_3 P_{t-1}\hat{e}_{t-1} + \nu_t \tag{3.41}$$

As hipóteses são

$$H_0 : \alpha_1 = \alpha_2 = \alpha_3 = 0$$

$$H_A = \text{pelo menos um } \alpha_i \neq 0$$

A estatística de teste é $LM = N \times R^2$ e é assitoticamente distribuída sob uma distribuição $\chi^2\,(3)$. Rejeite H_0 caso $LM > \chi^2_\alpha\,(3)$ sendo α o nível de significância.

Variantes do teste de Engle e Ng (1993) podem ser facilmente consideradas para os diversos modelos GARCH não lineares.

3.5 Testes de adequação do modelo

A suposição de que os resíduos são independentes e identicamente distribuídos deve ser testada após o ajuste do modelo de variância condicional. Assim os resíduos padronizados $\hat{\epsilon}_t$ nos modelos (3.4) e (3.11) devem ser testados. Verifique a independência através do teste Brock, Dechert, Scheinkman e LeBaron (1986). Verifique também a presença de descorrelação serial entre os resíduos estimados ao quadrado $(\hat{\epsilon}_t^2)$ através dos tested de McLeod e Li (1993) ou através do teste ARCH-LM de Engle (1982). Estes testes apontam para presença do efeito ARCH remanescente no modelo ajustado. Obviamente o teste para identificar o tipo da distribuição de $\hat{\epsilon}_t$ deve ser analisado sob a natureza da distribuição adotada para os resíduos.

3.6 Volatilidade estocástica

Os modelos até então analisados consideram que a volatilidade em t é função dos choques e volatilidades passadas. Portanto, dadas as informações em $t-1$ a volatilidade condicional é determinística. Além deste fato, os choques na série y_t e na volatilidade h_t possuem a mesma natureza.

No modelo de volatilidade estocástica os choques simultâneos na média da série y_t e na volatilidade h_t são governados por processos descorrelatados. O modelo de volatilidade estocástica foi proposto por Taylor (1986). Este modelo recebeu pouca atenção devido às dificuldades de estimação. No entanto, com a evolução computacional e a redução do tempo de processamento, novas técnicas de estimação têm sido utilizadas. Desta forma, o modelo de volatilidade estocástica tem recebido especial atenção principalmente no que se refere a metodologias de estimação.

O modelo pode ser escrito como:

$$y_t = \sigma_t \epsilon_t \tag{3.42}$$

$$\sigma_t^2 = k\, e^{h_t}$$

$$h_t = \gamma h_{t-1} + \eta_t$$

onde $\epsilon_t \sim NID\,(0,1)$, $\eta_t \sim NID\left(0, \sigma_\eta^2\right)$, $|\gamma| < 1$ e $E\,(\epsilon_t \eta_{t-s}) = 0$ para $s \geq 0$. Isto significa que o processo seguido por h_t é estacionário.

Exercício 24: Seja o modelo de volatilidade estocástica formulado em (3.42).

(i) calcule a média e variância incondicionais de h_t

(ii) calcule a média e variância condicional de y_t

(iii) calcule a média e a variância incondicional de y_t

(iv) calcule o quarto momento de y_t e a curtose K_y

(v) calcule os demais momentos pares de y_t, isto é, o sexto, oitavo, ... e o $2m$-ésimo momento.

(vi) calcule a covariância de y_t

(vii) calcule a correlação de y_t^2

Comparando os resultados do exercício 22 (iii) e do exercício 24 (vii) observamos que os modelos GARCH(1,1) e volatilidade estocástica possuem FACś que são um decaimento exponencial para zero. Veja em Carnero et al. um estudo comparativo entre os modelos GARCH e volatilidade estocástica. A maior diferença entre os modelos GARCH e volatilidade estocástica recai sobre a estimação. No modelo GARCH a estimação é feita pela maximização da verossimilhança. A função de verossimilhança é construída a partir de informações passadas de y_t. No modelo de volatilidade estocástica $y_t|I_{t-1}$ não pode ser construída a partir das informações passadas de y_t uma vez que h_t está sujeita a um processo de choques diferentes de y_t.

Dentre as metodologias de estimação para o modelo de volatilidade estocástica podemos mencionar:

(i) método dos momentos

(ii) métodos de máxima verossimilhança, através de simulação numérica usando amostragem ponderada e Monte-Carlo cadeia de Markov (veja Shepard e Pitt (1997), Sandman e Koopman (1998), Jaquier et al (1994) e Kim et al (1998)).[1]

(iii) método de quase-máxima verossimilhança (QMLE) (veja Nelson (1998), Harvey et al (1994)). Em Ruiz (1994) veja que o QMLE é consistente e assintoticamente normal.

(iv) métodos de linearização

O método de quase máxima verossimilhança está implementado no software STAMP . O modelo realiza a estimação dos componentes não observáveis através do filtro de Kalman dentro da abordagem da metodologia espaço-estado. Veja Durbin e Koopman (2002) e Harvey (1989).

Reescrevemos o modelo da equação (3.42) tal qual foi implementado no software acima (veja Koopman et al (2000)). Combinando a primeira e a segunda equações temos:

$$y_t = \sigma \epsilon_t \exp(h_t/2) \tag{3.43}$$

$$h_t = \gamma h_{t-1} + \eta_t$$

[1]Broto e Ruiz (2002) apresentam uma resenha sobre as metodologias de estimação do modelo de volatilidade estocástica.

Elevando ao quadrado a equação e tomando o logarítmo, temos:

$$\ln y_t^2 = \ln \sigma^2 + \ln \epsilon_t^2 + h_t$$

Somando e subtraindo $E\left(\ln \epsilon_t^2\right)$

$$\ln y_t^2 = \ln \sigma^2 + E\left(\ln \epsilon_t^2\right) + \ln \epsilon_t^2 - E\left(\ln \epsilon_t^2\right) + h_t$$

$$\ln y_t^2 = k + h_t + \xi_t \tag{3.44}$$

onde $\xi_t = \ln \epsilon_t^2 - E\left(\ln \epsilon_t^2\right)$ e $k = \ln \sigma^2 + E\left(\ln \epsilon_t^2\right)$. Observe que não há necessidade de assumir uma distribuição particular de ϵ_t. O método de quase-máxima verossimilhança ignora qual a correta distribuição de $\ln y_t^2$ ou ξ_t e adota como sendo normalmente distribuída. Fuller (1996) propôs a seguinte transformação para y_t:

$$\ln y_t^2 = \ln\left(y_t^2 + cs_y^2\right) - \frac{cs_y^2}{y_t^2 + cs_y^2} \quad \text{para } t = 1, \ldots, N$$

onde s_y^2 é a variância amostral de y_t e c é uma constante adotada como $0,02$ em vários estudos, veja Breidt e Carriquiry (1996) e Bollerslev e Wright (2001).

Bibliografia

- Akaike, H. (1974). A new look at statistical model identification, IEEE Transaction on Automatic Control AC 19,716-23.

- Anderson, H. M., Nam K. e Vahid, F. (1999). Asymetric nonlinear smooth transition GARCH models, in P. Rothmam (ed.). Nonlinear Time Series Analysis of Economic and Financial Data, Boston: Klewer, 191-207

- Bera, A. K. e Jarque, C. M. (1981). An efficient large sample test for normality of observations and regression residuals. Working Paper in Econometrics, No. 40, Australian National University, Canberra

- Black, F. (1976). The pricing of commodity contracts, Journal of Financial Economics 3,167-79

- Bollerslev , T. (1986). Generalized autoregressive conditional heterocedasticity, Journal of Econometrics 31, 307-27

- (1987). A conditional heteroskedastic time series model for speculative prices and rates of return, Review of Economics and Statistics 69, 542-7

- Bollerslev, T. e Wrigth, J. H. (2001). High frequency data, frequency domain influence and volatility forecasting, Review of Economics and Statistics 83, 596-02

- Breidt, F. J. e Carriquiry, A. L. (1996). Improved quasi-maximum likelihood estimators for stochastic volatility models, in A. Zellner and J. S. Lee (eds.), Modelling and Prediction: Honoring Seymour Geisel, Springer-Verlag, New York

- Brock, W. A., Dechert, W. Scheinkman, J. A. e LeBaron, B. (1986). A test of independence based on correlation dimension, Econometric Reviews 15, 197-35

- Broto, C. e Ruiz E. (2002). Estimation methods for stochastic volatility models: a survey, Working Paper

- Campbell, J. Y., Lo, A. W. e MacKinlay, A. C. (1997). The econometrics of financial markets, Princeton: Princeton University Press

- Dickey, D. A. e Fuller, W. A. (1979). Distribution of the estimators of autoregressive time series with unit root, Journal of American Statistical Association 74, 427-31

- Durbin, J. e Koopman, S. J. (2002). Time Series Analysis by State Space Methods, Oxford University Press

- Enders, W. (1995). Applied Econometric Time Series. John Wiley & Sons

- Engle, R. F. (1982). Autoregressive conditional heterocedasticity with estimates of the variance of United Kingdom inflation, Econometrica 50, 987-007

- Engle, R. F., Lilien, D. M. e Robins, R. P. (1987). Estimating time varying risk premia in the term sturcture: The ARCH-M model, Econometrica 55, 391-07

- Engle, R. F. e Ng, V. K. (1993). Measuring and testing the impact of news on volatility, Journal of Finance 48, 1749-78

- Fornari, F. e Mele, A. (1996). Modelling the changing asymetry of conditional variances, Economic Letters 50, 197-03

- (1997). Sign-and volatility-switching ARCH models: theory and apllications to international stock markets, Journal of Applied Econometrics 12, 49-65

- Franses, P. H. e van Dijk, D. (2000). Non-Linear Time Series Models in Empirical Finance, Cambridge University Press

- Fuller, W. A. (1996). Introduction to Statistical Time Series, Wiley, New York

- Glosten, L. R., Jagannathan R. e Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess on stocks, Journal of Finance 48, 1779=801

- Gonzáles-Rivera, G. (1998). Smooth transition GARCH models, Studies in Nonlinear Dynamics and Econometrics 3, 61-78

- Granger, C. W. J. e Andersen A. (1978). An Introduction to Bilinear Time Series Models, Göttingem: Vandenhoeck & Ruprecht

- Hamilton, J. D. (1994). Time Series Analysis, Princeton University Press, Princeton, New Jersey
- Hagerud, G. E. (1997). A new nonlinear GARCH model, PhD Thesis, IFE, Stockholm School of Economics
- Harvey, A. C., Ruiz, E. e Shephard, N. G. (1994). Mutivariate variane models, Riview of Economic Studies
- Hsieh, D. A. (1983). A heterocedasticity-consistent covariance matrix estimator for time series regressiona, Journal of Econometrics 22, 281-90
- Jaquier, E., Polson, N. G. e Rossi, P. E. (1994). Bayesian analysis of stochastic volatility models, Journal of Business and Economics Statistics 12, 371-17
- Kim, S. e Shephard N. (1994). Stochastic volatility and comparinson with ARCH models, Unpublished Paper: Nuffield College, Oxford
- Koopman, S. J., Harvey, A. C., Doornik, J. A. e Shephard, N. (2000). STAMP - Structural Time Series Analyser, Modeller and Predictor, Manual do usuário, Timberlake Consultants
- Mackinnon, J. G. (1990). Critical values for co-integration tests, Working Paper, University of California, San Diego
- McLeod, A. I. e Li, W. K. (1983). Diagnostic cheking ARMA time series models using squared-residual autocorrelations, Journal of Time Series Analysis 4, 269-43
- Moretin, P. A. e Toloi, C. M. C. (2004). Análise de Séries Temporais, São Paulo, Edgard Blücher
- Nelson, D. B. (1988). The time series behavior of sotck market wolatility and returns, Unpublised PhD dissertation, MIT, Cambridge, MA
- (1990). Stationarity and persistence in the GARCH(1,1) model, Econometric Theory 6, 318-34
- (1991). Conditional heteroskedasticity in asset returns: a new approach, Econometrica 59, 347-70
- Rabemananjara, R. e Zakoïan, J. M. (1993). Threshold ARCH models and asymmetries in volatility, Journal of Applied Econometrics 8, 31-49

- Ruiz, E. (1994). Quasi-maximum likelihood estimation of Stochastic volatility models, Journal of Econometrics 63, 289-06
- Sandmann, G. e Koopman, S. J. (1998). Estimation of stochastic volatility models via Monte-Carlo maximum Likelihood, Journal of Econometrics 87, 271-301
- Schwarz, G. (1978). Estimation the dimension of a model, Annals of Statistics 6, 461-4
- Sentana, E. (1995). Quadratic GARCH models, Review of Economic Studies 62, 639-61
- Taylor, S. J. (1988). Moddeling Financial Time Series, New York, John Wiley
- Tong, H. e Lim, K. S. (1980). Threshold autoregressions, limit cycles, and data, Journal of Royal Statistical Society B42, 245-92
- Tsay, R. S. (2002) Analysis of Fenancial Series, John Wiley & Sons
- Zakoïan, J. M. (1994). Threshold heteroskedastic models, Journal of Ecnomic Dynamics and Control 18, 931-44