

Modelos para a Variância Condicional e Não Condicional Univariada: Especificação e Previsão*

Pedro Valls[†]

Sao Paulo School of Economics -FGV e CEQEF-FGV

7 de julho de 2024

Sumário

1	Introdução	5
2	Construção do Modelo	6
2.1	Teste para Heteroscedasticidade condicional	7
2.2	Testando Heteroscedasticidade Condicional para o Índice Bovespa	8
3	Modelo Amostral	9
3.1	Comando no OxMetrics	10
4	Suavizamento Exponencial	11
4.1	Introdução	11
4.2	Parâmetro Ótimo de Decaimento	11
4.3	Programa no R para obter o Parâmetro de Decaimento	12
4.4	Estimação do modelo EWMA no OxMetrics	13
5	Modelos GARCH	14
5.1	Introdução	14
5.2	Modelos ARCH	14
5.2.1	Construção do Modelo <i>ARCH</i>	16
5.2.2	Função de Verossimilhança para um <i>ARCH</i> (m)	16
5.2.3	Programa no EVIEWS para Estimar um ARCH(1) com Distribuição Normal	18
5.2.4	Programa no EVIEWS para Estimar um ARCH(1) com Distribuição t-Student	21
5.2.5	Programa no EVIEWS para Estimar um ARCH(1) com Distribuição GED	24

*©-Pedro Valls (2022)

[†]EESP-FGV Rua Dr. Plínio Barreto 365 sala 1319, 01313-020, São Paulo, SP, Brasil. E-mail: pedro.valls@fgv.br

5.2.6	Estimação do Modelo ARCH no OxMetrics	27
5.2.7	Estimação do Modelo ARCH no R	29
5.3	Modelos GARCH	30
5.3.1	Programa no EVIEWS para Estimar um GARCH(1,1)	32
5.3.2	Estimação do GARCH(1,1) no OxMetrics para todas as distribuições	35
5.3.3	Estimação do GARCH(1,1) no R para todas as distribuições . . .	37
5.4	Modelos IGARCH	38
5.4.1	Programa no EVIEWS para Estimar um IGARCH	40
5.4.2	Estimação do IGARCH no OxMetrics	40
5.4.3	Estimação do IGARCH no R	42
5.5	Outros Membros da família XARCH	43
5.5.1	Modelos EGARCH	44
5.5.2	Programa no EVIEWS para estimar um EGARCH	44
5.5.3	Modelo EGARCH no OxMetrics	47
5.5.4	Modelo GJR ou Limiar GARCH no EVIEWS	50
5.5.5	Modelo GJR no OxMetrics	52
5.6	Modelo APARCH	54
6	Previsão para Modelos GARCH	57
6.1	Comandos no EVIEWS para fazer Previsões	59
6.2	Previsões no OxMetrics	59
7	Modelo Fracionário para Volatilidade	60
7.1	Introdução	60
7.2	Teste R/S para Memória Longa	63
7.3	Teste GPH para Memória Longa	65
7.4	Estimando um Modelo ARFIMA	67
7.5	Modelo <i>FIGARCH</i> para Volatilidade	69
7.6	Modelo <i>HYGARCH</i> para Volatilidade	70
7.7	Modelo <i>FIEGARCH</i> para Volatilidade	71
7.8	Modelo <i>FIAPARCH</i> para Volatilidade	72
8	Modelos <i>Dynamic Conditional Score(DCS)</i> ou <i>Generalized Autoregressive Score(GAS)</i> para Volatilidade	74
8.1	Introdução	74
8.2	Filtro de Kalman e Extração de Sinal	77
8.3	Especificação dos Modelos <i>Dynamical Conditional Score (DCS)</i> e <i>Generalized Autoregressive Score (GAS)</i>	81
8.4	Comparação do modelos GARCH com os GAS para o índice BOVESPA .	85
9	Comparação do Modelo de Volatilidade Estocástica e do GARCH(1,1)	86
9.1	Introdução	86
9.2	Especificação do Modelo	86
9.2.1	Programa no EVIEWS para estimar um modelo VE	88

Lista de Figuras

1	FAC dos Quadrados dos Retornos	8
2	SD Móvel para $m = 22, 44, 66, 126, 252$	10
3	Variância estimada pelo EWMA usando amostra com pandemia	12
4	Variância estimada pelo EWMA usando amostra sem pandemia	13
5	Volatilidade Condicional estimada por EWMA com $\lambda = 0.9401$	14
6	Volatilidade BOVESPA usando N-GARCH	19
7	FAC e FACP para Resíduos Padronizados	20
8	FAC e FACP para Resíduos Padronizados ao Quadrado	20
9	Volatilidade BOVESPA usando t-GARCH	22
10	FAC e FACP para Resíduos Padronizados para t-GARCH	23
11	FAC e FACP para Resíduos Padronizados ao Quadrado para t-GARCH	23
12	Volatilidade BOVESPA usando GED-GARCH	25
13	FAC e FACP para Resíduos Padronizados para GED-ARCH	26
14	FAC e FACP para Resíduos Padronizados ao Quadrado para GED-GARCH	26
15	Cond Variance para os Modelos ARCH com todas as distribuições	28
16	Cond Volatilidade para os Modelos ARCH com todas as distribuições no R	30
17	Cond Variance para os Modelos N-GARCH(1,1) para IBOVESPA	33
18	FAC e FACP dos Resíduos Padronizados do modelo N-GARCH(1,1) para IBOVESPA	34
19	FAC e FACP dos Resíduos Padronizados ao quadrado do modelo N-GARCH(1,1) para IBOVESPA	34
20	Volatilidade para todos os GARCH(1,1) para IBOVESPA	36
21	Cond Volatilidade para os Modelos GARCH(1,1) com todas as distribuições no R	38
22	Volatilidade para todos os GARCH(1,1) para IBOVESPA	41
23	Volatilidade para todos os IGARCH(1,1) para IBOVESPA	43
24	Volatilidade para todos os N-EGARCH(1,1) para IBOVESPA	45
25	FAC e FACP para Resíduos Padronizados para o modelo N-EGARCH(1,1) para IBOVESPA	46
26	FAC e FACP para Resíduos Padronizados ao quadrado para o modelo N-EGARCH(1,1) para IBOVESPA	46
27	Volatilidade para todos os EGARCH(1,1) para IBOVESPA	49
28	Volatilidade do N- GJR(1,1) para IBOVESPA	52
29	Volatilidade para todos os GJR(1,1) para IBOVESPA	53
30	Volatilidade para todos os APARCH(1,1) para IBOVESPA	56
31	Previsão fora da amostra para N-GARCH para o IBOVESPA	59
32	Previsão fora da amostra para N-GARCH usando OxMetrics para o IBOVESPA	60
33	FAC para o Valor Absoluto para o retorno composto percentual do IBOVESPA	61
34	FAC para Ruído Branco Fracionário com $d = 1/3$ e um AR(1) com $\phi = 0.5$	63
35	AbsDLIBOVM X Ajustado melhor ARFIMA(0,d,1)	69
36	Desvio Padrão Condicional do modelo $SkSt - FIAPARCH(1, d, 1)$ para o retorno composto percentual do IBOVESPA	73
37	Preços e Retornos diários em % de BMY	75
38	Retornos Absolutos e o desvio padr ao condicional de BMY	76

39	Retornos Absolutos, $SDSkSt - GARCH$ e $SD - GAS - SkSt$ para BMY	77
40	Volatilidade Estocástica e N-GARCH(1,1) para IBOVESPA	89
41	Volatilidade Estocástica e N-GARCH(1,1) para IBOVESPA estimado no OxMetrics	90

Lista de Tabelas

1	Teste de Heteroscedasticidade Condicional - Teste de Engle com 12 defasagens	9
2	Estimação N-ARCH(1) para BOVESPA	19
3	Estimação t-ARCH(1) para BOVESPA	21
4	Estimação GED-ARCH(1) para BOVESPA	24
5	Comparação dos Modelos ARCH para as quatro distribuições	28
6	Teste de Normalidade para o melhor Modelo ARCH - SkSt	29
7	Teste de Box-Pierce para Resíduos Padronizados do melhor Modelo ARCH - SkSt	29
8	Teste de Box-Pierce para Resíduos Padronizados ao Quadrado do melhor Modelo ARCH - SkSt	29
9	Comparação dos Modelos ARCH para as quatro distribuições usando o R	29
10	Estimação N-GARCH(1,1) para Retorno Composto Percentual do Índice BOVESPA	33
11	Comparação dos Modelos GARCH(1,1) para as quatro distribuições	35
12	Teste de Normalidade para o melhor Modelo GARCH - SkSt	36
13	Teste de Box-Pierce para Resíduos Padronizados do melhor Modelo GARCH - SkSt	36
14	Teste de Box-Pierce para Resíduos Padronizados ao Quadrado do melhor Modelo GARCH - SkSt	37
15	Comparação dos Modelos GARCH(1,1) para as quatro distribuições no R	37
16	Comparação dos Modelos IGARCH(1,1) para as quatro distribuições	40
17	Teste de Normalidade para o melhor Modelo IGARCH - SkSt	41
18	Teste de Box-Pierce para Resíduos Padronizados do melhor Modelo IGARCH - SkSt	41
19	Teste de Box-Pierce para Resíduos Padronizados ao Quadrado do melhor Modelo IGARCH - SkSt	42
20	Comparação dos Modelos IGARCH(1,1) para as quatro distribuições	42
21	Estimativa do Modelo N-EGARCH no Eviews	45
22	Comparação dos Modelos EGARCH(1,1) para as quatro distribuições	48
23	Estimativa do Modelo SkSt-EGARCH no OxMetrics	48
24	Teste de Normalidade para o melhor Modelo EGARCH - SkSt	49
25	Teste de Box-Pierce para Resíduos Padronizados do melhor Modelo IGARCH - SkSt	49
26	Teste de Box-Pierce para Resíduos Padronizados ao Quadrado do melhor Modelo IGARCH - SkSt	50
27	Estimação do Modelo Limiar GARCH no EVIEWS	51
28	Comparação dos Modelos GJR(1,1) para as quatro distribuições	52
29	Estimativa do Modelo SkSt-GJR no OxMetrics	53
30	Teste de Normalidade para o melhor Modelo GJR - SkSt	54

31	Teste de Box-Pierce para Resíduos Padronizados do melhor Modelo GJR - SkSt	54
32	Teste de Box-Pierce para Resíduos Padronizados ao Quadrado do melhor Modelo GJR - SkSt	54
33	Comparação dos Modelos APARCH(1,1) para as quatro distribuições . .	55
34	Estimativa do SkSt-APARCH(1,1) para IBOVESPA	56
35	Teste de Normalidade para o melhor Modelo APARCH - SkSt	57
36	Teste de Box-Pierce para Resíduos Padronizados do melhor Modelo APARCH - SkSt	57
37	Teste de Box-Pierce para Resíduos Padronizados ao Quadrado do melhor Modelo GJR - SkSt	57
38	Teste R/S para o Retorno Composto Percentual do IBOVESPA	64
39	Teste R/S para o Valor Absoluto Retorno Composto Percentual do IBOVESPA	65
40	Teste de GPH para Valor absoluto do retorno composto percentual do Índice BOVESPA	67
41	Estimação do Modelo ARFIMA(1,d,1)	68
42	Estimação do Modelo ARFIMA(0,d,1)	68
43	Comparação entre os modelos ARFIMA(p,d,q)	68
44	Comparação entre os modelos FIGARCH(1,d,1)	70
45	Comparação entre os modelos HYGARCH(1,d,1)	71
46	Comparação entre os modelos FIEGARCH(1,d,1)	72
47	Comparação entre os modelos FIAPARCH(1,d,1)	72
48	Estimação do Modelo $SkSt - FIAPARCH(1, d, 1)$	73
49	Estimação do Modelo $AR(1) + GARCH(1, 1) + SkSt$	76
50	Comparação dos Modelos APARCH(1,1) para as quatro distribuições . .	85
51	Comparação entre os modelos GAS(1,1)	85
52	Comparação entre os modelos EGAS(1,1)	85
53	Comparação entre os modelos AEGAS(1,1)	85
54	Modelo de Volatilidade Estocástica para o Índice BOVESPA.	89

1 Introdução

A especificação e seleção de modelos é sempre guiada por fatos estilizados empíricos. A capacidade de um modelo de reproduzir estes fatos é uma característica desejável e por outro lado a sua incapacidade de reproduzir é um critério para se rejeitar tal modelo. Alguns dos fatos estilizados de séries de tempo financeiras são:

- (a) **caudas pesadas:** desde o trabalho de Mandelbrot [1963] e Fama [1963]-Fama [1965] é reconhecido que retornos de ativos são leptocúrticos. Como resultado, vários estudos tem proposto modelar retornos de ativos através de distribuições com caudas pesadas, por exemplo, Pareto-Lévy;
- (b) **agrupamento de volatilidade:** qualquer observação de séries temporais financeiras revela instantes de tempo de alta e baixa volatilidade que se agrupam. De fato, agrupamentos de volatilidade e caudas pesadas estão relacionados. Uma das características dos modelos ARCH é a ligação entre dinâmica (condicional) da volatilidade e caudas pesadas (não condicional);

- (c) **efeitos de alavancagem:** Black [1976] observou que os movimentos de preço são negativamente correlacionados com a volatilidade. Se o preço dos ativos está em queda ocorre um aumento da alavancagem da firma e, em geral, também um aumento da incerteza, fatos que tendem a gerar aumento da volatilidade.
- (d) **chegada de informações:** usualmente retornos são mensurados em frequências fixas, por exemplo: dia, semana, mês. Alguns autores sugeriram relacionar explicitamente os retornos aos fluxos de informações que chegam ao mercado. De fato Clark [1973] é um dos primeiros a propor um modelo de volatilidade estocástica (*Stochastic Volatility - SV*) onde a chegada de informações é não uniforme no tempo e em geral não observada. Este fato está relacionado ao conceito de deformação temporal que pode ser diretamente ligada a volatilidade através de uma função que depende de volume, tempo entre cotações, anúncios com informações específicas ao ativo, notícias macroeconômicas e fechamento do mercado;
- (e) **memória longa e persistência:** volatilidade é altamente persistente. Em particular para dados de alta frequência existe evidência de que o processo gerador da variância condicional tem raiz próxima de unitária. Esta evidência fez com que a modelagem da persistência fosse feita através de modelos com raízes unitárias ou através de modelos de memória longa;
- (f) **co-movimentos de volatilidade:** existe uma extensa literatura sobre co-movimentos dos mercados especulativos internacionais. Globalização dos mercados financeiros aumenta a volatilidade dos preços;
- (g) **correlação na volatilidade implícita:** como a volatilidade implícita é calculada através do modelo de Black and Scholes [1973], é, obviamente, baseada num modelo (model-based). Pelo fato da volatilidade ser calculada com base nos preços diários existe uma inconsistência, pois o modelo de Black-Scholes pressupõe que a volatilidade é constante. As propriedades estatísticas da série temporal de volatilidade implícita podem ser aproximadas por modelos de reversão a média do tipo autor-regressivo. Além disto, a volatilidade implícita deve conter informação a respeito da volatilidade futura e, portanto, deve ser um preditor desta última;
- (h) **estrutura a termo da volatilidade implícita:** o modelo de Black e Scholes pressupõe uma estrutura a termo constante para a volatilidade. Na realidade, a estrutura a termo da volatilidade implícita de opções *at-the-money* tem inclinação positiva quando a volatilidade de curto prazo é baixa e negativamente inclinada quando a volatilidade é alta;
- (i) **smiles:** é conhecido na literatura que a volatilidade implícita de Black e Scholes produz um vício conhecido na literatura por *smile effect* devido ao formato em U da volatilidade implícita para diferentes preços de exercício.

2 Construção do Modelo

O processo para identificar um modelo para um ativo financeiro é composto por quatro etapas:

1. Como existem feriados, os retornos de ativo financeiros não são igualmente espaçados. Através do Filtro de Kalman e usando um modelo de nível local é possível estimar as observações faltantes para o nível dos ativos financeiros.¹
2. Especificar um modelo para média condicional testando se existe correlação serial nos dados e se necessário especificar um modelo *ARMA* para remover a dependência linear.
3. Usar os resíduos da equação da média para testar se existe efeito de heteroscedasticidade condicional.
4. Especificar um modelo para a variância condicional e estimar conjuntamente média condicional e variância condicional
5. Verificar a validade do modelo

2.1 Teste para Heteroscedasticidade condicional

Seja $a_t = r_t - \mu_t$ o resíduo da equação da média condicional. A série dos resíduos ao quadrado, isto é, a_t^2 , será usada para verificar a existência de heteroscedasticidade condicional.

Dois testes estão disponíveis na literatura

O primeiro usa a estatística de Ljung-Box (Ljung and Box [1978]), $Q(m)$, na série $\{a_t^2\}$, onde a hipótese nula é que as primeiras m defasagens na FAC de a_t^2 são zero.

O segundo teste é o teste de Multiplicadores de Lagrange apresentado em Engle [1982]. Este teste é equivalente a um teste F para se testar $\alpha_i = 0$ ($i = 1, \dots, m$) na regressão linear

$$a_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \dots + \alpha_m a_{t-m}^2 + e_t \quad t = m+1, \dots, T \quad (1)$$

onde e_t denota o erro, m um inteiro positivo pre-especificado e T o tamanho da amostra.

A hipótese nula é

$$H_0 : \alpha_1 = \dots = \alpha_m = 0 \quad (2)$$

e a soma de quadrado sob H_0 será dada por $SQR_0 = \sum_{t=m+1}^T (a_t^2 - \varpi)^2$ onde $\varpi =$

$(1/T) \sum_{t=1}^T a_t^2$ é a média amostral de a_t^2 , e a soma de quadrados sob H_1 será dada por

$.SQR_1 = \sum_{t=m+1}^T \hat{e}_t^2$ onde \hat{e}_t^2 é o resíduo de (1). Temos então a seguinte estatística de teste

$$F = \frac{(SQR_0 - SQR_1)/m}{SQR_1/(T - 2m - 1)} \quad (3)$$

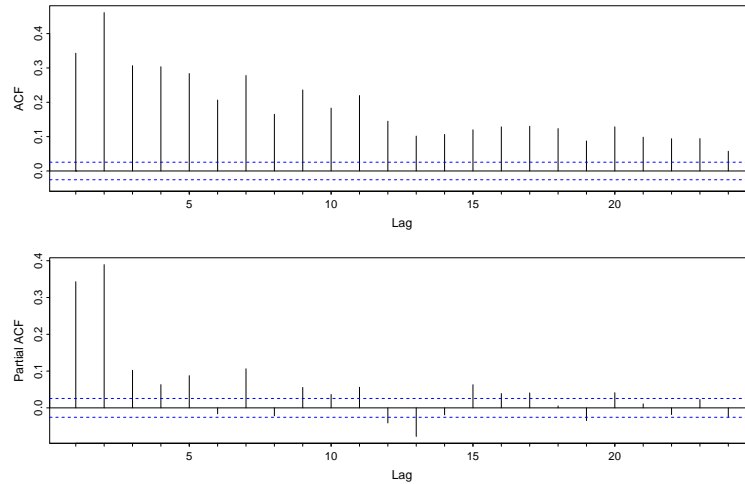
que tem distribuição qui-quadrada com m graus de liberdade.

¹Um exemplo de como isto pode ser feito no R é o script Estimando Observações faltantes no IBOVC. No caso do OxMetrics podemos usar o seguinte *batch* Estimando Observações faltantes no IBOVC usando o STAMP do OxMetrics.

2.2 Testando Heteroscedasticidade Condicional para o Índice Bovespa

O FAC e FACP para o quadrado dos retornos do Ibovespa é apresentado abaixo

Figura 1: FAC dos Quadrados dos Retornos



cujo teste de Ljung-Box para as primeiras 50 autocorrelações é dado por:

$$LB(50) \sim \chi^2(50) = 6659.7[0.0000] \quad (4)$$

Como o p-valor é menor do que 10^{-4} temos evidências de heteroscedasticidade condicional.

E o teste de Engle é obtido usando os seguintes comandos do OxMetrics em Github-Ox-Garch e no R Teste de Engle para DLIbov_sq

Temos os seguintes resultados para o programa do R:

Tabela 1: Teste de Heteroscedasticidade Condicional - Teste de Engle com 12 defasagens

EQ(1) Teste de Engle com 12 defasagens				
The estimation sample is: 13 - 5880				
	Coefficient	Std. Error	t-Statitistics	p-value
<i>const</i>	0.7457	0.1131	6.591	0.0000
<i>DLIbov_sq_{t-1}</i>	0.1628	0.0131	12.466	0.0000
<i>DLIbov_sq_{t-2}</i>	0.3234	0.0132	24.489	0.0000
<i>DLIbov_sq_{t-3}</i>	0.0591	0.0138	4.265	0.0000
<i>DLIbov_sq_{t-4}</i>	0.03676	0.01387	2.651	0.00806
<i>DLIbov_sq_{t-5}</i>	0.05671	0.0139	4.090	0.0000
<i>DLIbov_sq_{t-6}</i>	-0.0335	0.0138	-2.420	0.0155
<i>DLIbov_sq_{t-7}</i>	0.0875	0.0138	6.323	0.0000
<i>DLIbov_sq_{t-8}</i>	-0.0443	0.0139	-3.198	0.0014
<i>DLIbov_sq_{t-9}</i>	0.0345	0.0139	2.491	0.0128
<i>DLIbov_sq_{t-10}</i>	0.0412	0.0138	2.981	0.0029
<i>DLIbov_sq_{t-11}</i>	0.0626	0.0132	4.746	0.0000
<i>DLIbov_sq_{t-12}</i>	-0.04139	0.0130	-3.179	0.0015
no of observations: 5879				
no of parameters: 13				
$\sigma_\varepsilon = 7.605$				
$R^2 = 0.2849$		$\overline{R}^2 = 0.2835$		
$F(12, 5854) = 194.4[0.0000]$				
$AIC = 6.8982$				
$BIC = 6.8988$				
log-likelihood = -20222.73				

Como o p-valor é menor do que 10^{-4} existe evidência de heteroscedasticidade condicional.

3 Modelo Amostral

Inicialmente, considere o problema da estimação das volatilidades (desvios padrões) dos retornos dos ativos. Assumindo média igual a zero, a volatilidade amostral dos retornos do ativo i utilizando uma amostra de T observações é definido como:

$$S_i = \sqrt{\frac{1}{T} \sum_{t=1}^T r_{ti}^2} \quad (5)$$

O cálculo do estimador utilizando toda amostra permite pouca adaptabilidade às informações mais recentes. Isto decorre do fato de que todas observações da amostra recebem o mesmo peso.

Para contornar este problema, utiliza-se, ao invés de toda amostra, uma janela móvel com um número fixo de observações. Apesar de ainda manter peso igual para todas observações utilizadas na janela, consegue-se alguma flexibilidade, pois pode-se controlar a importância da observação mais recente através da escolha do tamanho da janela.

3.1 Comando no OxMetrics

O OxMetrics tem um comando para calcular o desvio padrão móvel e é dado por:

```
> movingSD(VAR, LAG, LEAD, MEAN)
```

onde **VAR** representa a variável, **LAG** representa o número de defasagens, **LEAD** representa o número de períodos posteriores a observação vigente e **MEAN** representa a média da variável **VAR**.

Como o OxMetrics pode calcular a média usando também uma janela móvel através do seguinte comando:

```
> movingavg(VAR, LAG, LEAD)
```

podemos então usar **movingavg** como argumento em **movingSD** obtendo-se o seguinte comando:

```
> movingavg(VAR, LAG, LEAD)
> movingSD(VAR, LAG, LEAD, movingavg)
```

No R o comando é:

```
> runsd(x,k,center=runmean(x,k))
```

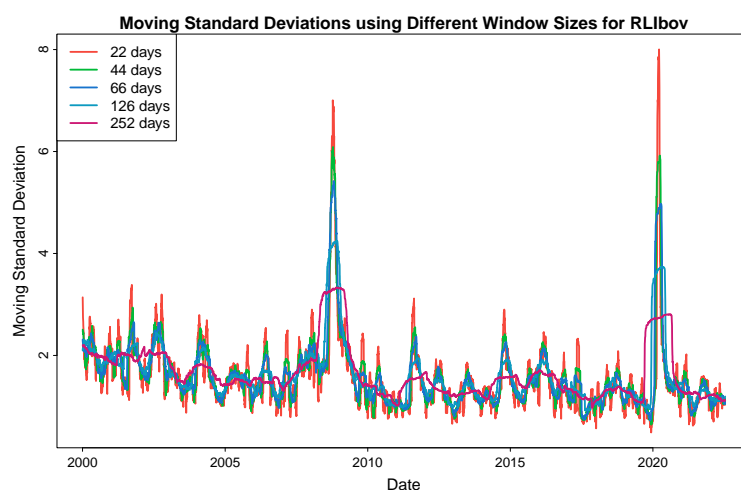
onde x é a variável que se deseja calcular o desvio padrão móvel, k é a janela e **runmean**(x,n) é a média móvel da série x com janela de k períodos.

O programa no R é apresentado abaixo:

Desvio Padrão Móvel para DLIbov

Usando as janelas com 22, 44, 66, 126 e 252 dias, temos o seguinte gráfico:

Figura 2: SD Móvel para $m = 22, 44, 66, 126, 252$



No entanto, o uso deste estimador apresenta outro sério inconveniente. Em geral, a ocorrência de eventos extremos nas séries financeiras se dá com relativa frequência. Como este estimador utiliza o mesmo peso para todas as observações da amostra, a

volatilidade estimada dá um salto para cima após um retorno extremo permanecendo neste nível enquanto a observação permanecer na amostra. Quando a observação extrema sai da amostra a volatilidade salta para baixo novamente. Desta forma, em períodos subsequentes a grandes variações de preços dos ativos, volatilidade e, por consequência, o risco tende a ser superestimado.

Esta forma de cálculo é usada para a determinação de uma medida de dispersão de um conjunto de dados de diversas outras áreas de estudo, além da financeira. Com efeito, não existe uma preocupação com a reprodução dos fatos estilizados citados na seção anterior. Isto faz com que sua aplicabilidade no cálculo da volatilidade dos retornos de ativos financeiros seja bastante limitada.

4 Suavizamento Exponencial

4.1 Introdução

A técnica de suavizamento exponencial é uma tentativa de contornar a limitação do método amostral. Neste caso, o estimador da variância dos retornos do ativo i na data t é dado por:

$$\sigma_{t,i}^2 = \lambda \sigma_{t-1,i}^2 + (1 - \lambda) r_{t-1,i}^2 \quad \text{para } 0 \leq \lambda \leq 1 \quad (6)$$

A variância do retorno num dado instante é composto por dois termos. O primeiro, um termo autorregressivo expressando a dependência temporal da variância dos retornos, fato estilizado encontrado em séries financeiras. O segundo, representando a contribuição da observação mais recente para a variância estimada. A expressão (6) pode ser rescrita da seguinte forma:

$$\sigma_{T,i}^2 = \lambda^T \sigma_{0,i}^2 + (1 - \lambda) \sum_{t=1}^T \lambda^t r_{T-t,i}^2 \quad (7)$$

Por (7) a estimativa da variância dos retornos é igual a da variância inicial mais uma soma com pesos geometricamente declinantes dos quadrados dos retornos. A influência da variância inicial sobre a variância presente tende a zero com o número de observações. Um candidato natural a estimador deste termo é o estimador (5) com uma janela arbitrária. O segundo termo faz com que o efeito de choques nas séries de retornos sejam dissipados suavemente com o tempo. Note que (5) é o caso particular de (7) com $\lambda = \frac{T-1}{T}$.

4.2 Parâmetro Ótimo de Decaimento

A escolha do parâmetro λ é ad hoc no sentido de que não envolve nenhum procedimento estatístico que gere uma estimativa a partir das observações passadas. No Riskmetrics é sugerido um procedimento para a escolha de um λ ótimo baseado no erro de previsão um passo a frente.

O previsor da variância do retorno um passo a frente é definido com $E(r_{t+1}^2) = \sigma_{t+1|t}^2$. Definindo o erro de predição um passo a frente como sendo $\varepsilon_{t+1|t} = r_{t+1}^2 - \sigma_{t+1|t}^2$ Segue-se

que $E(\varepsilon_{t+1|t}) = 0$. Partindo destes conceitos, um critério razoável para a escolha do λ ótimo de cada ativo seria tomar o λ que minimizasse a soma dos erros de predição um passo a frente:

$$RMSE_t = \sqrt{\frac{1}{T} \sum_{t=1}^T (r_{t+1}^2 - \sigma_{t+1|t}^2(\lambda))^2} \quad (8)$$

O modelo de suavizamento exponencial tem maior capacidade de reproduzir os fatos estilizados citados na introdução. A principal característica do modelo é que um valor de λ próximo de 1 reproduz o fato estilizado (e), segundo o qual a volatilidade é altamente persistente. Contudo, o fato da soma dos parâmetros do modelo (λ e $1 - \lambda$) ser igual a um gera um inconveniente, a volatilidade não condicional dos retornos, como veremos a seguir, será dada pela constante do modelo 6 que é igual a zero, dividido por $1 - \lambda - (1 - \lambda) = 0$, logo é indeterminada. Como resultado, a distribuição não condicional dos retornos tem média igual a zero e variância não definida, o que não faz sentido, uma vez que este fato sem não tem nenhum apoio empírico.

4.3 Programa no R para obter o Parâmetro de Decaimento

Em

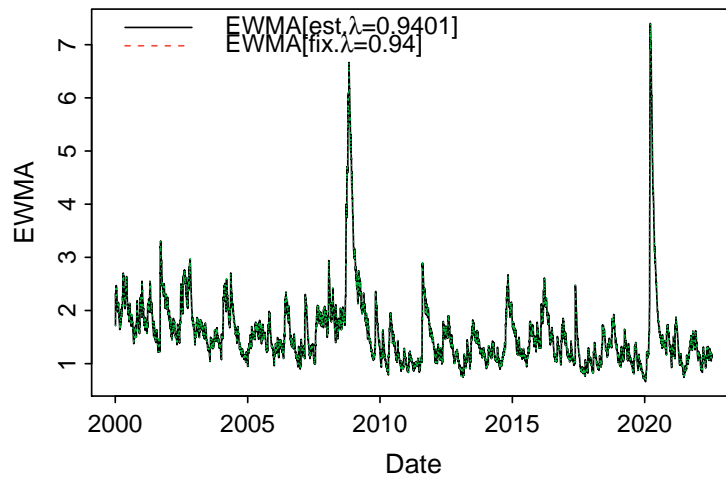
EWMA no R usando pacote rugarch

é apresentado um programa para estimar o parâmetro de decaimento ótimo usando a amostra toda que inclui a pandemia.

Para a série do log retorno do Índice Bovespa temos que o λ ótimo é $\lambda = 0.9401$.

O gráfico abaixo apresenta a volatilidade estimada usando-se este λ ótimo.

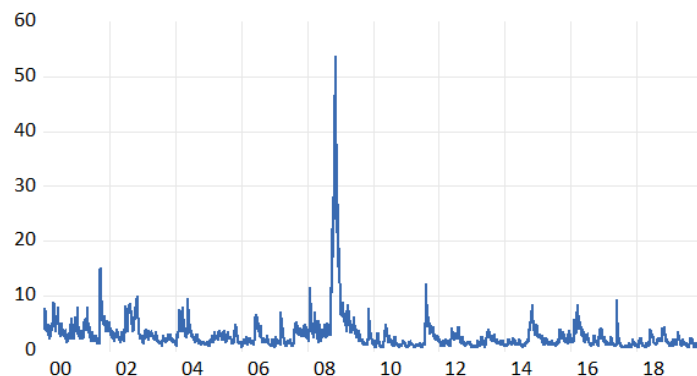
Figura 3: Variância estimada pelo EWMA usando amostra com pandemia



Agora se excluirmos o período da pandemia o programe em Github-EWMA-sem-pandemia e neste caso temos que o λ ótimo é $1 - \alpha = 1 - 0.098 = 0.902$.

O gráfico abaixo apresenta a volatilidade estimada usando-se este λ ótimo.

Figura 4: Variância estimada pelo EWMA usando amostra sem pandemia



Em Engle and Patton [2001] Meia Vida (*Half Life* denotada por HL) é definida "*as the time taken for the volatility to move halfway back towards its unconditional mean following a deviation from it*" e é dada por:

$$HL = k : |h_{t+k|t} - \sigma^2| = \frac{1}{2}|h_{t+1|t} - \sigma^2| \quad (9)$$

Para o período total, isto é, que inclui a pandemia este valoré $HL_{all-sample} = 11$ e para o período sem a pandemia, $HL_{Jan-2000-Jan-2020} = 6.7$.

4.4 Estimação do modelo EWMA no OxMetrics

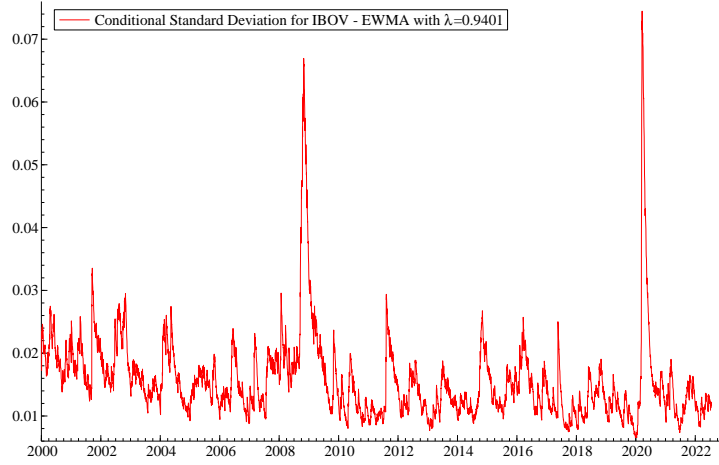
É possível estimar um modelo EWMA no OxMetrics mas a desvantagem é que o parâmetro ótimo não é estimado sendo fixado a um certo valor indicado pelo usuário. Usando $\lambda = 0.9401$.

O programa no Ox para estimar um EWMA com $\lambda = 0.9401$ é dado por:

EWMA no Ox com $\lambda = 0.9401$.

temos o gráfico da volatilidade condicional:

Figura 5: Volatilidade Condicional estimada por EWMA com $\lambda = 0.9401$



5 Modelos GARCH

5.1 Introdução

Como dito na introdução, para retornos de séries financeiras é comum o fato de que grandes valores num determinado instante do tempo sejam seguidos por valores também elevados nos períodos subsequentes, não necessariamente na mesma direção, fato estilizado (b). Estatisticamente, esta característica pode ser descrita pela presença de elevada autocorrelação no quadrado dos retornos. A autocorrelação presente no quadrado dos retornos das séries financeiras faz com que a variância condicional dos retornos apresente uma dependência temporal dos choques passados.

O modelo de suavizamento exponencial apresentado na seção anterior captura esta característica destas séries, pois, como demonstrado, a estimativa da variância dos retornos é igual a da variância inicial mais uma soma com pesos geometricamente declinantes dos quadrados dos retornos. O problema com este modelo é que não existe um critério estatístico para estimação do parâmetro λ que leve em consideração as propriedades específicas de cada série de retorno. Desta forma, não é possível realizar inferência sobre as estimativas do modelo.

5.2 Modelos ARCH

Um modelo mais genérico para a estimação da variância condicional dos retornos é o ARCH proposto por Engle [1982]. Este modelo expressa a variância condicional como uma defasagem distribuída do quadrado dos retornos passados. Seja

$$r_t = a_t = \sigma_t \varepsilon_t \quad \text{com } \varepsilon_t \sim NI(0, 1) \quad (10)$$

$$E(r_t^2 \mid \mathbf{I}_{t-1}) = \sigma_t^2 = \omega + \sum_{i=1}^m \alpha_i a_{t-i}^2 = \omega + \alpha(L) a_t^2 \quad (11)$$

sendo $\alpha(L)$ um polinômio no operador defasagem do tipo $\alpha(L) = \alpha_1 L + \alpha_2 L^2 + \dots + \alpha_m L^m$ e, para garantir a não negatividade da variância condicional: $\omega, \alpha_i > 0$ para $i = 1, \dots, m$.

Para o caso particular em que $m = 1$, temos as seguintes propriedades:

$$\begin{aligned} E(a_t \mid \mathbf{I}_{t-1}) &= E(\varepsilon_t \sqrt{\omega + \alpha_1 a_{t-1}^2} \mid \mathbf{I}_{t-1}) \\ &= \sqrt{\omega + \alpha_1 a_{t-1}^2} E(\varepsilon_t \mid \mathbf{I}_{t-1}) = 0 \implies E(a_t) = E(E(a_t \mid \mathbf{I}_{t-1})) = 0 \end{aligned} \quad (12)$$

$$\begin{aligned} E(a_t^2 \mid \mathbf{I}_{t-1}) &= E(\varepsilon_t^2 (\omega + \alpha_1 a_{t-1}^2) \mid \mathbf{I}_{t-1}) \\ &= (\omega + \alpha_1 a_{t-1}^2) E(\varepsilon_t^2 \mid \mathbf{I}_{t-1}) \\ &= (\omega + \alpha_1 a_{t-1}^2) \end{aligned} \quad (13)$$

$$E(a_t a_{t-j} \mid \mathbf{I}_{t-j}) = a_{t-j} E(a_t \mid \mathbf{I}_{t-j}) = 0 \quad (14)$$

$$\begin{aligned} Var(a_t) &= E(a_t^2) \\ &= E(E(a_t^2 \mid \mathbf{I}_{t-1})) \\ &= E(\omega + \alpha_1 a_{t-1}^2) \\ &= \omega + \alpha_1 E(a_{t-1}^2) \\ &= \omega + \alpha_1 Var(a_{t-1}^2) \\ &\quad \Downarrow \\ Var(a_t) &= \frac{\omega}{1 - \alpha_1} \text{ se } \omega > 0 \text{ e } 0 \leq \alpha_1 < 1 \end{aligned} \quad (15)$$

Como estamos assumindo distribuição normal para ε_t temos que

$$E(a_t^4 \mid \mathbf{I}_{t-1}) = 3[E(a_t^2 \mid \mathbf{I}_{t-1})]^2 = 3(\omega + \alpha_1 a_{t-1}^2)^2 \quad (16)$$

Logo

$$E(a_t^4) = E[E(a_t^4 \mid \mathbf{I}_{t-1})] = 3E(\omega + \alpha_1 a_{t-1}^2)^2 = 3E(\omega^2 + 2\omega\alpha_1 a_{t-1}^2 + \alpha_1^2 a_{t-1}^4) \quad (17)$$

Agora se a_t for estacionário e tiver quarto momento finito, isto é, $E(a_t^4) = m_4$, temos que (17) pode ser escrito da seguinte forma

$$\begin{aligned} m_4 &= 3[\omega^2 + 2\omega\alpha_1 Var(a_t) + \alpha_1^2 m_4] \\ &= 3\omega^2 + 6\omega\alpha_1 \frac{\omega}{1 - \alpha_1} + 3\alpha_1^2 m_4 \\ &= 3\omega^2 \left(1 + 2\frac{\alpha_1}{1 - \alpha_1}\right) + 3\alpha_1^2 m_4 \\ \implies m_4 &= \frac{3\omega^2(1 + \alpha_1)}{(1 - \alpha_1)(1 - 3\alpha_1^2)} \end{aligned} \quad (18)$$

Por (18) temos as seguintes implicações:

- a) como o quarto momento é positivo temos que $1 - 3\alpha_1^2 > 0 \implies 0 \leq \alpha_1^2 < \frac{1}{3}$
- b) o curtosis de a_t será dado por:

$$\frac{E(a_t^4)}{[Var(a_t)]^2} = \frac{3\omega^2(1 + \alpha_1)}{(1 - \alpha_1)(1 - 3\alpha_1^2)} \times \frac{(1 - \alpha_1)^2}{\omega^2} = 3 \frac{1 - \alpha_1^2}{(1 - 3\alpha_1^2)} > 3 \quad (19)$$

Observe que (95), (15) e (14) implicam que a_t tem média zero, variância constante e são não correlacionados, logo parece um ruído branco. Mas, por (19), temos que o excesso de curtosis é positivo, implicando que a distribuição não condicional de a_t tem caudas mais pesadas do que a distribuição normal e portanto o modelo *ARCH* tem mais chance de produzir outliers do que um ruído branco Gaussiano e reproduz o fato estilizado (a). Por outro lado (13) implica que a variância condicional varia com o tempo.

Observe também que se definirmos $v_t = r_t^2 - \sigma_t^2$ podemos escrever (11) quando $q = 1$ da seguinte forma:

$$r_t^2 = \omega + \alpha_1 r_{t-1}^2 + v_t \quad (20)$$

que é um modelo *AR*(1) para o quadrado das observações, mas com os erros sendo uma diferença martingale, isto é, $E(v_t | \mathbf{I}_{t-1}) = 0$. Este modelo é chamado de *ARCH*(1) - Condicional em ser Heteroscedastico é um *AR*(1).

O modelo acima possui algumas propriedades indesejáveis.

1. Choques positivos ou negativos tem o mesmo efeito na volatilidade, violando o efeito alavancagem, fato estilizado (c).
2. O modelo *ARCH* é bem restritivo uma vez que para o modelo de primeira ordem o coeficiente $\alpha_1^2 \in [0, \frac{1}{3}]$ que limita a capacidade do modelo de capturar o excesso de curtosis.
3. O modelo *ARCH* tende a superprever a volatilidade porque eles respondem vagarosamente a grande choques.

5.2.1 Construção do Modelo *ARCH*

A determinação da ordem do modelo *ARCH* pode ser feita através da FAC e FACP de a_t^2 . Mas como na representação autorregressiva de a_t^2 as inovações não são i.i.d. os estimadores de Mínimos Quadrados Ordinários dos coeficientes da parte autorregressiva são consistentes mas ineficientes.

5.2.2 Função de Verossimilhança para um *ARCH*(m)

Através da técnica de decomposição de erros de predição, é possível construir a função de verossimilhança o que permite a estimação dos parâmetros pelo método de máxima verossimilhança.

$$\begin{aligned}
f(r_1, \dots, r_T | \theta) &= \prod_{t=m+1}^T f(r_t | \mathbf{I}_{t-1}) \times f(r_1, \dots, r_m | \theta) \\
&= \prod_{t=m+1}^T \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp \left\{ -\frac{r_t^2}{2\sigma_t^2} \right\} \times f(r_1, \dots, r_m | \theta)
\end{aligned} \tag{21}$$

onde $\theta = (\omega, \alpha_1, \dots, \alpha_m)$ e $f(r_1, \dots, r_m | \theta)$ representa a densidade conjunta das condições iniciais. Em geral esta densidade é complicada e em vez de obter o estimador exato obtemos o estimador aproximado ao desconsiderar esta densidade, temos então a seguinte função de verossimilhança condicional

$$l(\theta | r_1, \dots, r_m) = \prod_{t=m+1}^T \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp \left\{ -\frac{r_t^2}{2\sigma_t^2} \right\} \tag{22}$$

e aplicando a transformação logarítmica temos a log verossimilhança condicional que é dada por:

$$\ln l(\theta | r_1, \dots, r_m) = \sum_{t=m+1}^T \left(-\frac{1}{2} \ln(2\pi) - \frac{1}{2} \ln(\sigma_t^2) - \frac{r_t^2}{2\sigma_t^2} \right) \tag{23}$$

onde $\sigma_t^2 = \omega + \alpha_1 a_{t-1}^2 + \dots + \alpha_m a_{t-m}^2$.

Maximizando, numericamente, (23) em relação ao parâmetros θ obtém-se os estimadores de máxima verossimilhança condicional.

Em algumas situações a distribuição de ε_t tem caudas mais pesadas, por exemplo, uma t – *Student* com v graus de liberdade. Em geral esta distribuição é reparametrizada para ter variância unitária, uma vez que a variância da distribuição t – *Student* é $\frac{v}{v-2}$. Então definimos a densidade para $\tilde{\varepsilon}_t = \varepsilon_t / \sqrt{v/(v-2)}$ da seguinte forma:

$$f(\tilde{\varepsilon}_t | v) = \frac{\Gamma((v+1)/2)}{\Gamma(v/2)\sqrt{(v-2)\pi}} \left(1 + \frac{\tilde{\varepsilon}_t^2}{v-2} \right)^{-(v+1)/2} \quad v > 2 \tag{24}$$

onde $\Gamma(x)$ é a função gamma definida por $\Gamma(x) = \int_0^\infty y^{x-1} e^{-y} dy$. Usando que $r_t = a_t = \sigma_t \tilde{\varepsilon}_t$ temos a função de densidade conjunta condicional nas condições iniciais, dada por:

$$f(r_{m+1}, \dots, r_T | \theta, v) = \prod_{t=m+1}^T \frac{\Gamma((v+1)/2)}{\Gamma(v/2)\sqrt{(v-2)\pi}\sigma_t} \left(1 + \frac{r_t^2}{(v-2)\sigma_t^2} \right)^{-(v+1)/2} \tag{25}$$

Agora caso os graus de liberdade sejam conhecidos a priori, o logaritmo da função de verossimilhança será dado por:

$$\ln(l(\theta | v, r_1, \dots, r_m)) = \sum_{t=m+1}^T \left[-\frac{v+1}{2} \ln \left(1 + \frac{r_t^2}{(v-2)\sigma_t^2} \right) - \frac{1}{2} \ln(\sigma_t^2) \right] \tag{26}$$

e caso se deseje estimar os graus de liberdade também (26) é modificado para:

$$\begin{aligned} \ln(l(\theta, v|r_1, \dots, r_m)) &= (T - m)[\ln(\Gamma((v + 1)/2)) - \ln(\Gamma(v/2)) - 0.5 \ln((v - 2)\pi)] + \\ &\quad + \ln(l(\theta|v, r_1, \dots, r_m)) \end{aligned} \quad (27)$$

onde o último termo é dado por (26).

Uma outra distribuição que é usada para estimar modelos *ARCH* é a distribuição de erro generalizada (generalized error distribution - GED) cuja densidade é dada por

$$f(x) = \frac{v \exp\left(-\frac{1}{2} \left| x/\lambda \right|^v\right)}{\lambda 2^{(1+1/v)} \Gamma(1/v)} \quad -\infty < x < \infty, \quad 0 < v < \infty$$

onde $\Gamma(\cdot)$ é a função gamma e

$$\lambda = \left[2^{(-2/v)} \Gamma(1/v) \Gamma(3/v) \right]^{1/2} \quad (28)$$

Esta distribuição reduz-se a Normal se $v = 2$ e terá caudas pesadas se $v < 2$.

PAREI AQUI E INCLUIR OS PROGRAMAS EM R

5.2.3 Programa no EVIEWS para Estimar um ARCH(1) com Distribuição Normal

As primeiras quinze linhas do programa Github-GARCH-EVIEWS estima um $N - ARCH(1)$, modelo *ARCH*(1) com distribuição Normal, para o retorno composto percentual do Índice BOVESPA.

As estimativas dos parâmetros são apresentadas abaixo

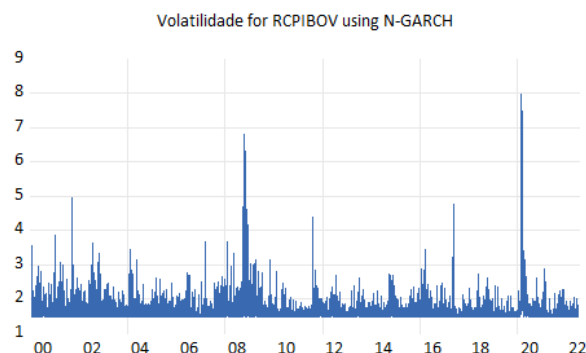
Tabela 2: Estimação N-ARCH(1) para BOVESPA

Dependent Variable: DLIBOVM
Method: ML - ARCH
Date: 08/22/22 Time: 09:12
Sample (adjusted): 1/04/2000 7/15/2022
Included observations: 5879 after adjustments
Convergence achieved after 33 iterations
Coefficient covariance computed using outer product of gradients
Presample variance: backcast (parameter = 0.7)
GARCH = C(2) + C(3)*RESID(-1)^2

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	0.017511	0.019822	0.883383	0.3770
Variance Equation				
C	2.171782	0.034276	63.36224	0.0000
RESID(-1)^2	0.237792	0.011483	20.70847	0.0000
R-squared	−0.000104	Mean dependent var	−2.52E − 16	
Adjusted R-squared	−0.000104	S.D. dependent var	1.715912	
S.E. of regression	1.716001	Akaike info criterion	3.820093	
Sum squared resid	17308.72	Schwarz criterion	3.823501	
Log likelihood	−11226.16	Hannan-Quinn criter.	3.821278	
Durbin-Watson stat	2.055721			

O gráfico abaixo apresenta o volatilidade para este modelo

Figura 6: Volatilidade BOVESPA usando N-GARCH



Para verificar a adequabilidade do modelo, devemos verificar se a FAC e FACP para os resíduos padronizados não apresenta correlação serial indicando que a equação para a média condicional está bem especificada. Verificando se a FAC e FACP para os resíduos padronizados ao quadrado não tem estrutura de cor-

relação é uma indicação de que a equação para a variância condicional está bem especificada.

As linhas 18 a 27 do programa Github-GARCH-EVIEWS apresenta a FAC e FACP dos resíduos padronizados e dos resíduos padronizados ao quadrado.

Figura 7: FAC e FACP para Resíduos Padronizados

Correlogram of RES_N_ARCH_ST						
A	B	C	D	E	F	G
Date: 08/22/22 Time: 10:05						
Sample (adjusted): 1/04/2000 7/15/2022						
Included observations: 5879 after adjustments						
Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.001	0.001	0.0040	0.949
		2	0.002	0.002	0.0283	0.986
		3	-0.025	-0.025	3.7963	0.284
		4	-0.007	-0.007	4.0906	0.394
		5	0.000	0.000	4.0912	0.536
		6	-0.005	-0.006	4.2518	0.643
		7	-0.012	-0.012	5.1106	0.646
		8	-0.004	-0.004	5.2113	0.735
		9	-0.002	-0.002	5.2336	0.813
		10	0.031	0.030	10.846	0.370
		11	0.008	0.008	11.261	0.422
		12	0.003	0.003	11.312	0.502

Figura 8: FAC e FACP para Resíduos Padronizados ao Quadrado

Correlogram of RES_N_ARCH_ST_SQ						
A	B	C	D	E	F	G
Date: 08/22/22 Time: 10:05						
Sample (adjusted): 1/04/2000 7/15/2022						
Included observations: 5879 after adjustments						
Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	-0.035	-0.035	7.4062	0.006
		2	0.230	0.229	318.12	0.000
		3	0.152	0.175	453.48	0.000
		4	0.113	0.083	529.00	0.000
		5	0.185	0.135	729.95	0.000
		6	0.097	0.059	785.17	0.000
		7	0.163	0.090	941.70	0.000
		8	0.075	0.012	974.42	0.000
		9	0.150	0.068	1106.3	0.000
		10	0.140	0.085	1221.5	0.000
		11	0.094	0.028	1273.8	0.000
		12	0.138	0.045	1386.1	0.000

Pela FAC e FACP dos Resíduos Padronizados a equação para média condicional está bem especificada, mas pela FAC e FACP dos Resíduos ao Quadrado a equação da variância condicional não

está bem especificada, devemos usar uma ordem superior a um para o modelo $ARCH$

5.2.4 Programa no EVIEWS para Estimar um ARCH(1) com Distribuição t-Student

As linhas de 30 até 39 do programa Github-GARCH-EVIEWS estima um $t - ARCH(1)$, modelo $ARCH(1)$ com distribuição t-Student, para o retorno composto percentual do Índice BOVESPA.

As estimativas dos parâmetros são apresentadas abaixo

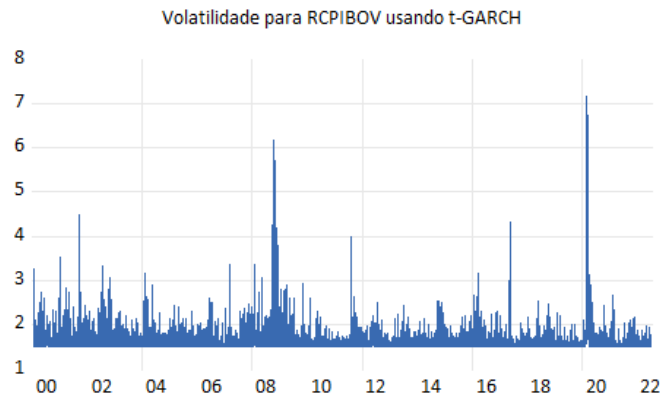
Tabela 3: Estimação t-ARCH(1) para BOVESPA

Dependent Variable: DLIBOVM
Method: ML ARCH - Student's t distribution (BFGS / Marquardt steps)
Date: 08/22/22 Time: 10:20
Sample (adjusted): 1/04/2000 7/15/2022
Included observations: 5879 after adjustments
Failure to improve likelihood (non-zero gradients) after 13 iterations
Coefficient covariance computed using outer product of gradients
Presample variance: backcast (parameter = 0.7)
GARCH = C(2) + C(3)*RESID(-1)^2

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	4.56E - 15	0.018859	2.42E - 13	1.0000
Variance Equation				
C	2.276148	0.086178	26.41221	0.0000
RESID(-1)^2	0.191270	0.022013	8.688950	0.0000
T-DIST. DOF	5.142679	0.349414	14.71801	0.0000
R-squared	0.000000	Mean dependent var	-2.52E - 16	
Adjusted R-squared	0.000000	S.D. dependent var	1.715912	
S.E. of regression	1.715912	Akaike info criterion	3.742094	
Sum squared resid	17306.91	Schwarz criterion	3.746638	
Log likelihood	-10995.88	Hannan-Quinn criter.	3.743673	
Durbin-Watson stat	2.055935			

O gráfico abaixo apresenta o volatilidade para este modelo

Figura 9: Volatilidade BOVESPA usando t-GARCH



Observe que o critério de Schwarz quando se usa a distribuição normal é de 3.8235 e quando se usa a distribuição $t - Student$ é de 3.7466, implicando que o modelo com distribuição $t - Student$ é melhor. Uma outra forma de verifica que o modelo $t - ARCH$ é mais adequado do que o $N - GARCH$ é através dos graus de liberdade, estimado, que neste caso foi 5.12 com desvio padrão de 0.35. Como este valor é pequeno a distribuição tem caudas mais pesadas do que a distribuição Normal.

Assim como fizemos no caso anterior a adequabilidade do modelo será verificada através da FAC e FACP para os resíduos padronizados, caso não apresentem correlação serial há indicação de que a equação para a média condicional está bem especificada e se a FAC e FACP para os resíduos padronizados ao quadrado não tem estrutura de correlação é uma indicação de que a equação para a variância condicional está bem especificada.

As linhas 41 a 50 do programa Github-GARCH-EVIEWS apresenta a FAC e FACP dos resíduos padronizados e dos resíduos padronizados ao quadrado.

Figura 10: FAC e FACP para Resíduos Padronizados para t-GARCH

















































Correlogram of RES_T_ARCH_ST						
A	B	C	D	E	F	G
Date: 08/22/22 Time: 11:51						
Sample (adjusted): 1/04/2000 7/15/2022						
Included observations: 5879 after adjustments						
Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.000	0.000	0.0004	0.984
		2	0.002	0.002	0.0219	0.989
		3	-0.025	-0.025	3.5927	0.309
		4	-0.007	-0.007	3.8825	0.422
		5	0.001	0.001	3.8930	0.565
		6	-0.006	-0.007	4.1329	0.659
		7	-0.011	-0.011	4.8012	0.684
		8	-0.004	-0.004	4.9009	0.768
		9	-0.002	-0.002	4.9244	0.841
		10	0.031	0.030	10.568	0.392
		11	0.008	0.008	10.992	0.444
		12	0.004	0.004	11.085	0.522

Figura 11: FAC e FACP para Resíduos Padronizados ao Quadrado para t-GARCH

Correlogram of RES_T_ARCH_ST_SQ						
A	B	C	D	E	F	G
Date: 08/22/22 Time: 11:51						
Sample (adjusted): 1/04/2000 7/15/2022						
Included observations: 5879 after adjustments						
Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	-0.018	-0.018	1.9818	0.159
		2	0.247	0.246	359.96	0.000
		3	0.160	0.179	511.53	0.000
		4	0.129	0.089	610.07	0.000
		5	0.196	0.139	836.65	0.000
		6	0.107	0.058	903.43	0.000
		7	0.183	0.099	1100.5	0.000
		8	0.084	0.013	1141.6	0.000
		9	0.161	0.066	1294.7	0.000
		10	0.153	0.087	1431.9	0.000
		11	0.108	0.032	1501.1	0.000
		12	0.144	0.039	1623.2	0.000

Semelhante ao obtido no caso em que a distribuição dos erros era normal, a FAC e FACP dos resíduos padronizados não apresenta correlação serial indicando que a equação da média condicional está bem especificada, mas a FAC e FACP para os resíduos padronizados ao quadrado apresenta correlação serial indicando que é necessário um *ARCH* de ordem superior a um para captar a dinâmica da volatilidade.

5.2.5 Programa no EVIEWS para Estimar um ARCH(1) com Distribuição GED

As linhas de 55 até 65 do programa Github-GARCH-EVIEWS estima um $GED - ARCH(1)$, modelo $ARCH(1)$ com distribuição GED , para o retorno composto percentual do Índice BOVESPA.

As estimativas dos parâmetros são apresentadas abaixo

Tabela 4: Estimação GED-ARCH(1) para BOVESPA

Dependent Variable: DLBOVM

Method: ML ARCH - Generalized error distribution (GED) (BFGS / Marquardt steps)

Date: 08/22/22 Time: 12:11

Sample (adjusted): 1/04/2000 7/15/2022

Included observations: 5879 after adjustments

Failure to improve likelihood (non-zero gradients) after 11 iterations

Coefficient covariance computed using outer product of gradients

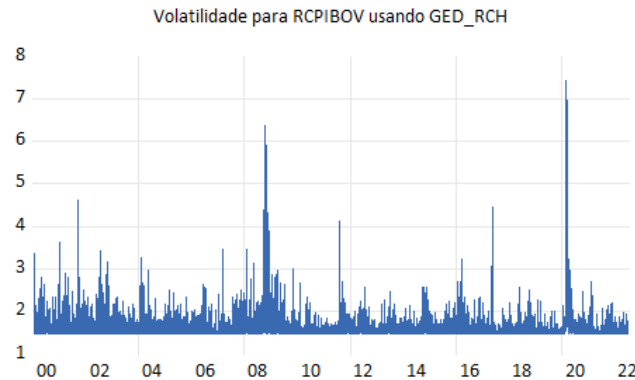
Presample variance: backcast (parameter = 0.7)

GARCH = C(2) + C(3)*RESID(-1)^2

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	$-2.14E - 15$	0.017973	$-1.19E - 13$	1.0000
Variance Equation				
C	2.191406	0.063958	34.26329	0.0000
RESID(-1)^2	0.205801	0.020508	10.03534	0.0000
GED PARAMETER	1.250002	0.025571	48.88350	0.0000
R-squared	0.000000	Mean dependent var	$-2.52E - 16$	
Adjusted R-squared	0.000000	S.D. dependent var	1.715912	
S.E. of regression	1.715912	Akaike info criterion	3.748102	
Sum squared resid	17306.91	Schwarz criterion	3.752646	
Log likelihood	-11013.54	Hannan-Quinn criter.	3.749681	
Durbin-Watson stat	2.055935			

O gráfico abaixo apresenta o volatilidade para este modelo

Figura 12: Volatilidade BOVESPA usando GED-GARCH



Observe que o critério de Schwarz quando se usa a distribuição GED é de 3.7526 que é maior do que quando se usa a distribuição $t - Student$ que é de 3.7466, implicando que o modelo com distribuição GED é pior que o modelo que usa a distribuição $t - Student$.

Assim como fizemos nos casos anteriores a adequabilidade do modelo será verificada através da FAC e FACP para os resíduos padronizados, caso não apresentem correlação serial há indicação de que a equação para a média condicional está bem especificada e se a FAC e FACP para os resíduos padronizados ao quadrado não tem estrutura de correlação é uma indicação de que a equação para a variância condicional está bem especificada.

As linhas 67 a 77 do programa Github-GARCH-EVIEWS apresenta a FAC e FACP dos resíduos padronizados e dos resíduos padronizados ao quadrado.

Figura 13: FAC e FACP para Resíduos Padronizados para GED-ARCH

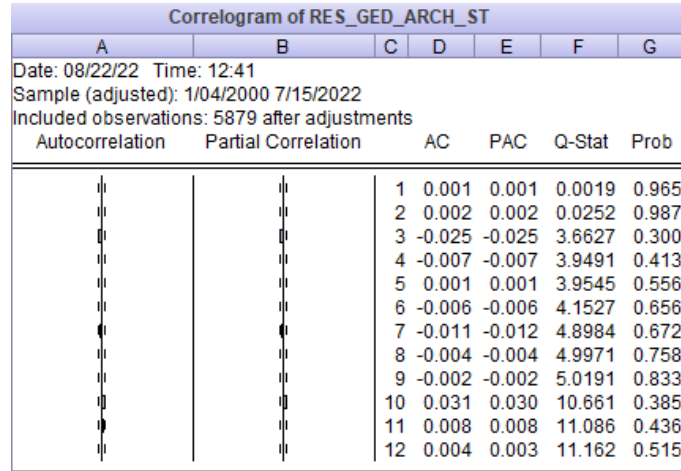
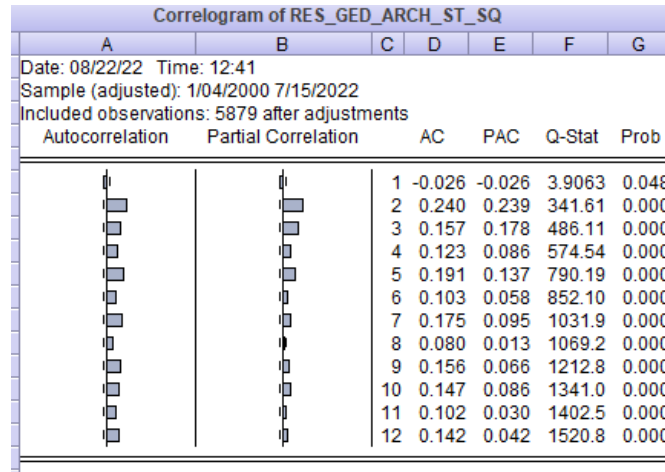


Figura 14: FAC e FACP para Resíduos Padronizados ao Quadrado para GED-GARCH



Semelhante ao obtido no caso em que a distribuição dos erros era normal, a FAC e FACP dos resíduos padronizados não apresenta correlação serial indicando que a equação da média condicional está bem especificada, mas a FAC e FACP para os resíduos padronizados ao quadrado apresenta correlação serial indicando que é necessário um *ARCH* de ordem superior a um para captar a dinâmica da volatilidade.

Semelhante aos casos anteriores não há má especificação na média condicional mas sim na variância condicional.

5.2.6 Estimação do Modelo ARCH no OxMetrics

A estimação do ARCH no OxMetrics é feita usando-se a o seguinte programa:

Github-Ox-GARCH

Uma opção no OxMetrics que não está implementada no EVIEWS é a estimação usando a distribuição t –*Student* assimétrica.

A log-verossimilhança para uma distribuição t-Student assimétrica padronizada (com média zero e variância unitária) é dada por:

$$\ln(l(\theta, v, s, m|r_1, \dots, r_m)) = (T - m) \left[\begin{array}{c} \ln(\Gamma((v + 1)/2)) - \ln(\Gamma(v/2)) \\ -0.5 \ln((v - 2)\pi) + \ln\left(\frac{2}{\xi + \frac{1}{\xi}}\right) \\ + \ln(s) \end{array} \right] \\ + \left[\begin{array}{c} \sum_{t=1}^T -\frac{v+1}{2} \ln\left(1 + \frac{(sr_t+m)^2}{(v-2)\sigma_t^2} \xi^{-2I_t}\right) \\ -\frac{1}{2} \ln(\sigma_t^2) \end{array} \right]$$

onde

$$I_t = \begin{cases} 1 & \text{se } \frac{r_t}{\sigma_t} \geq -\frac{m}{s} \\ -1 & \text{se } \frac{r_t}{\sigma_t} < -\frac{m}{s} \end{cases}$$

ξ é o parâmetro de assimetria, v são os graus de liberdade e m e s são definidos abaixo

$$m = \frac{\Gamma((v + 1)/2)\sqrt{v - 2}}{\sqrt{\pi}\Gamma(v/2)} \left(\xi - \frac{1}{\xi} \right) \quad (29)$$

$$s = \sqrt{\left(\xi^2 + \frac{1}{\xi^2} - 1 \right) - m^2} \quad (30)$$

Observe que em vez de estimar ξ é estimado $\ln(\xi)$ para facilitar a inferencia da hipótese nula de simetria, uma vez que quando $\xi = 1$ a distribuição simplifica para a $t - Student$ simétrica.

Temos então os seguintes resultados na comparação dos quatro modelos $ARCH(1)$, usando as distribuições: Normal, t-Student, GED e SkSt.

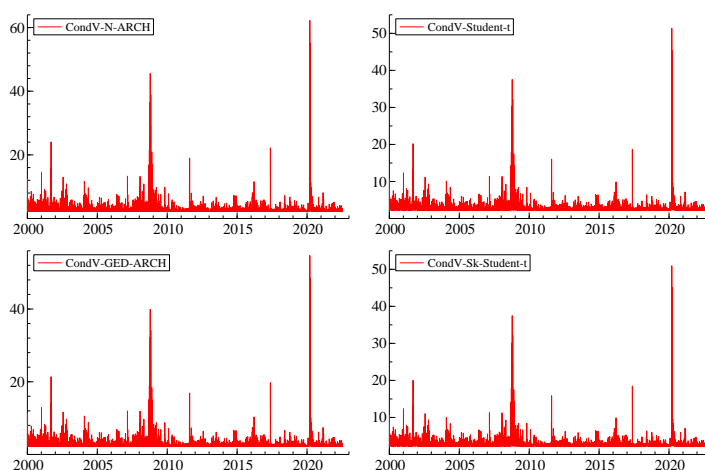
Tabela 5: Comparação dos Modelos ARCH para as quatro distribuições

Progress to date							
Model	T	p		log-likelihood	SC	HQ	AIC
N-ARCH(1)	5879	3	BFGS	-11232.170	3.8255	3.8233	3.8221
t-ARCH(1)	5879	4	BFGS	-10995.439	3.7465	3.7435	3.7419
GED-ARCH(1)	5879	4	BFGS	-11014.020	3.7528	3.7498	3.7483
SkSt-ARCH(1)	5879	5	BFGS	-10990.559	3.7463	3.7426	3.7406

Temos que o melhor modelo é exatamente quando se utiliza a distribuição $SkStudent - t$.

As estimativas da Variâncias Condicionais são apresentadas no gráfico abaixo.

Figura 15: Cond Variance para os Modelos ARCH com todas as distribuições



Pelos resultados das tabelas (6, 7 e 8) temos que normalidade é rejeitada, modelo para média condicional é adequado, só até defasagem de ordem 10 e modelo para a variância condicional é inadequado.

Tabela 6: Teste de Normalidade para o melhor Modelo ARCH - SkSt

Normality Test			
	Statistic	t-Test	P-Value
Skewness	-0.38286	11.987	$4.1340e-33$
Excess Kurtosis	3.2101	50.263	0.00000
Jarque-Bera	2667.8.	<i>NaN</i>	0.00000

Tabela 7: Teste de Box-Pierce para Resíduos Padronizados do melhor Modelo ARCH - SkSt

Q-Statistics on Standardized Residuals			
$Q(5)$	=	4.03315	[0.5446525]
$Q(10)$	=	10.8434	[0.3698449]
$Q(20)$	=	40.6082	[0.0041813] **
$Q(50)$	=	75.7601	[0.0108254]*
H_0 : No serial correlation => Accept H_0 when prob. is High			

Tabela 8: Teste de Box-Pierce para Resíduos Padronizados ao Quadrado do melhor Modelo ARCH - SkSt

Q-Statistics on Squared Standardized Residuals			
=> P-values adjusted by 1 degree(s) of freedom			
$Q(5)$	=	828.858	[0.0000000] **
$Q(10)$	=	1421.57	[0.0000000] **
$Q(20)$	=	2133.46	[0.0000000] **
$Q(50)$	=	3014.72	[0.0000000] **
H_0 : No serial correlation => Accept H_0 when prob. is High			

5.2.7 Estimação do Modelo ARCH no R

A estimação do ARCH no R é feita usando-se a o seguinte programa:

Programa no R para estimar Modelo ARCH com diversas distribuições

Temos então os seguintes resultados na comparação dos quatro modelos $ARCH(1)$, usando as distribuições: Normal, t-Student, GED e SkSt.

Tabela 9: Comparação dos Modelos ARCH para as quatro distribuições usando o R

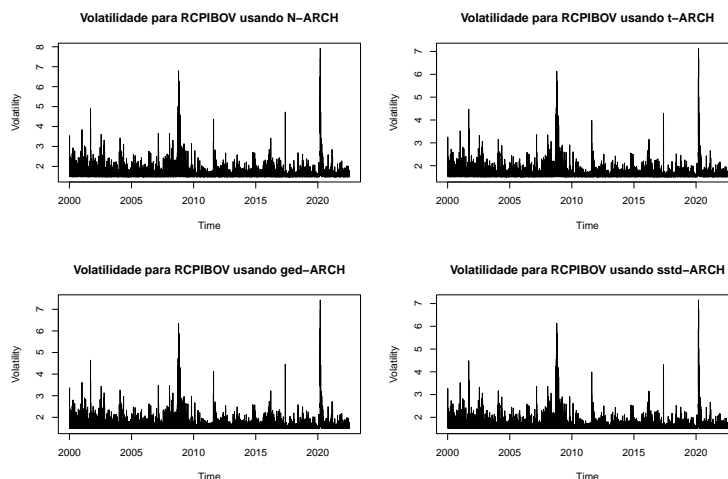
Progress to date							
Model	T	p		log-likelihood	SC	HQ	AIC
N-ARCH(1)	5879	3	BFGS	-11230.22	3.8215	3.8227	3.8215
t-ARCH(1)	5879	4	LBFGSB+NM	-10997.13	3.7425	3.7441	3.7425
GED-ARCH(1)	5879	4	LBFGSB+NM	-11013.19	3.7480	3.7496	3.7480
SkSt-ARCH(1)	5879	5	LBFGSB+NM	-10990.11	3.7405	3.7424	3.7405

Note: LBFGSB is the BFGS with lower and upper bound for the parameters. NM is a vector of step sizes for the finite-difference approximation to the gradient.

Temos que o melhor modelo é exatamente quando se utiliza a distribuição *SkStudent* – *t*.

As estimativas das Volatilidades (desvios padrões) Condicionais são apresentadas no gráfico abaixo:

Figura 16: Cond Volatilidade para os Modelos ARCH com todas as distribuições no R



Obviamente os resultados corroboram os obtidos com o OxMetrics. Resultados semelhantes as tabelas (6, 7 e 8) temos que normalidade é rejeitada, modelo para média condicional é adequado, só até defasagem de ordens 10 e 15, mas não para ordem 20 e modelo para a variância condicional é inadequado, indicando que é necessária uma ordem superior a um.

5.3 Modelos GARCH

Em geral, existe uma alta persistência na volatilidade das séries de retornos, fato estilizado (e), o que faz com que o valor de m no modelo *ARCH* seja elevado implicando a estimação de um grande número de parâmetros. O modelo *GARCH*, proposto por Bollerslev [1986], constitui-se numa tentativa de expressar de forma mais parcimoniosa a dependência temporal da variância condicional. Neste modelo a variância condicional, além de depender

do quadrado dos retornos passados como no modelo *ARCH*, depende também do passado das próprias variâncias condicionais. A variância condicional num modelo *GARCH*(s, m) é expressa por:

$$\sigma_t^2 = \omega + \sum_{i=1}^m \alpha_i a_{t-i}^2 + \sum_{j=1}^s \beta_j \sigma_{t-j}^2 = \omega + \alpha(L) a_t^2 + \beta(L) \sigma_t^2 \quad (31)$$

sendo $\alpha(L)$ e $\beta(L)$ polinômios no operador defasagem L .

A condição de não negatividade da variância condicional neste modelo é dada por: $\omega > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$ para $i = 1, \dots, m$ e $j = 1, \dots, s$. Esta condição é muito forte porque bastaria que o lado direito de 31 fosse positivo.

Usando a mesma mudança de variáveis que foi feita no modelo *ARCH*(1), é fácil mostrar que (31), quando $m = s = 1$, pode ser escrito da seguinte forma:

$$r_t^2 = \omega + (\alpha_1 + \beta_1) r_{t-1}^2 + v_t - \beta_1 v_{t-1} \quad (32)$$

onde $v_t = a_t^2 - \sigma_t^2 = r_t^2 - \sigma_t^2$. Logo o quadrado das observações é um modelo *ARMA*(1, 1) com as inovações sendo uma diferença martingale.

Para garantir que este processo *ARMA* para o quadrado dos retornos seja covariância estacionário, as raízes de: $1 - \alpha(L) - \beta(L)$ tem que estar fora do círculo unitário. Com a condição de não negatividade satisfeita estas condição de estacionaridade será dada por $0 < \sum_{i=1}^m \alpha_i + \sum_{j=1}^s \beta_j < 1$. Valendo esta condição, a variância não condicional que é dada por:

$$E(r_t^2) = E(\sigma_t^2) = \frac{\omega}{1 - \sum_{i=1}^m \alpha_i - \sum_{j=1}^s \beta_j} \quad (33)$$

e a esperança condicional da variância n passos a frente é igual a:

$$E(\sigma_{t+n}^2 \mid \mathbf{I}_t) = \left(\sum_{i=1}^m \alpha_i - \sum_{j=1}^s \beta_j \right)^{n-1} \left(\sigma_{t+1}^2 - \frac{\omega}{1 - \sum_{i=1}^m \alpha_i - \sum_{j=1}^s \beta_j} \right) + \frac{\omega}{1 - \sum_{i=1}^m \alpha_i - \sum_{j=1}^s \beta_j} \quad (34)$$

o que significa que existe uma tendência para a variância condicional retornar ao valor da variância não condicional

O modelo $GARCH(s, m)$ apresentado em (31) pode ser reescrito da seguinte forma:

$$\begin{aligned} \sigma_t^2 &= \omega + \alpha(L)a_t^2 + \beta(L)\sigma_t^2 \\ [1 - \beta(L)]\sigma_t^2 &= \omega + \alpha(L)a_t^2 \\ \sigma_t^2 &= \omega[1 - \beta(L)]^{-1} + \alpha(L)[1 - \beta(L)]^{-1}a_t^2 \end{aligned} \quad (35)$$

que corresponde a uma representação $ARCH(\infty)$.

Observe que pela representação (35) a variância condicional será estritamente positiva se $\omega[1 - \beta(L)]^{-1}$ for positiva e todos os coeficientes do polinômio $\alpha(L)[1 - \beta(L)]^{-1}$ são não negativos.

5.3.1 Programa no EViews para Estimar um GARCH(1,1)

As linhas 81 até 91 do programa em:

Github-GARCH-EViews

estima um GARCH(1,1) com distribuição Normal para o Retorno Composto Percentual do Índice Bovespa.

Temos a seguinte estimativa dos parâmetros:

Tabela 10: Estimação N-GARCH(1,1) para Retorno Composto Percentual do Índice BOVESPA

Dependent Variable: DLIBOVM				
Method: ML ARCH - Normal distribution (BFGS / Marquardt steps)				
Date: 08/24/22 Time: 13:35				
Sample (adjusted): 1/04/2000 7/15/2022				
Included observations: 5879 after adjustments				
Failure to improve likelihood (non-zero gradients) after 28 iterations				
Coefficient covariance computed using outer product of gradients				
Presample variance: backcast (parameter = 0.7)				
GARCH = C(2) + C(3)*RESID(-1)^2 + C(4)*GARCH(-1)				
Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	2.52E - 14	0.018418	1.37E - 12	1.0000
Variance Equation				
C	0.074380	0.008365	8.892044	0.0000
RESID(-1)^2	0.081577	0.004852	16.81465	0.0000
GARCH(-1)	0.890134	0.006693	133.0020	0.0000
R-squared	0.000000	Mean dependent var	-2.52E - 16	
Adjusted R-squared	0.000000	S.D. dependent var	1.715912	
S.E. of regression	1.715912	Akaike info criterion	3.681788	
Sum squared resid	17306.91	Schwarz criterion	3.686332	
Log likelihood	-10818.62	Hannan-Quinn criter.	3.683368	
Durbin-Watson stat	2.055935			

Observe que a persistência neste caso é dada por $\alpha + \beta = 0.9717$, que é menor do que um.

Abaixo é apresentado o gráfico da volatilidade

Figura 17: Cond Variance para os Modelos N-GARCH(1,1) para IBOVESPA



A adequabilidade do modelo será verificada através da FAC e FACP para os resíduos padronizados, caso não apresentem correlação serial há indicação de que a equação para a média condicional está bem especificada e se a FAC e FACP para os resíduos

padronizados ao quadrado não tem estrutura de correlação é uma indicação de que a equação para a variância condicional está bem especificada.

As linhas 93 até 103 do programa em Github-GARCH-EVIEWS apresenta a FAC-FACP dos resíduos padronizados e, também a FAC e FACP dos resíduos padronizados ao quadrado.

Temos então

Figura 18: FAC e FACP dos Resíduos Padronizados do modelo N-GARCH(1,1) para IBOVESPA

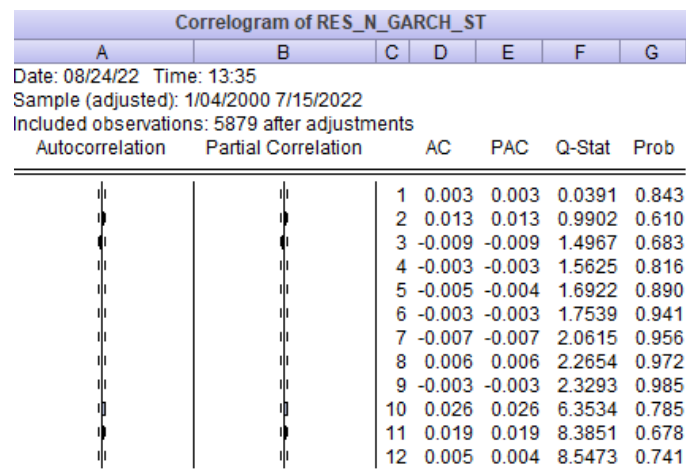
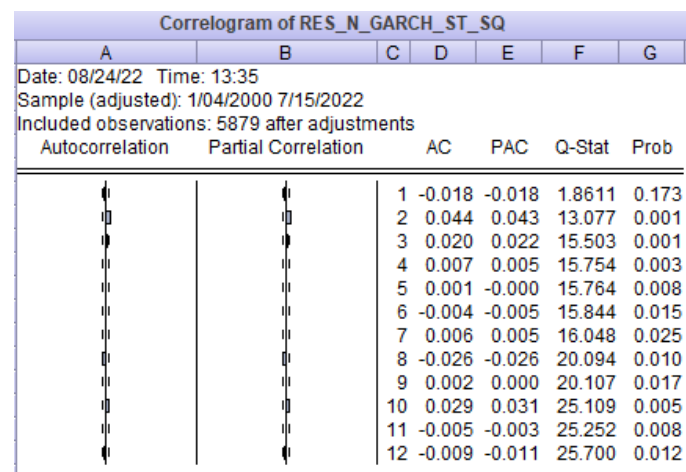


Figura 19: FAC e FACP dos Resíduos Padronizados ao quadrado do modelo N-GARCH(1,1) para IBOVESPA



Observe que neste caso tanto a FAC e FACP para os resíduos não apresentam correlação serial indicam que o modelo capta bem a estrutura na média condicional. Mas quanto a FAC e FACP dos resíduos ao quadrado sómente a primeira autocorrelação é não significativa mas as outras são indicando que o modelo N-GARCH(1,1) não capta toda a estrutura da variância condicional.

Observe que o critério de Schwarz para o modelo *GARCH* teve por valor 3.686332 que é inferior a todas as especificação para o modelo *ARCH* implicando que temos, agora, não só um modelo que é superior aos anteriores, mas ainda apresenta problemas de má especificação na variância condicional.

5.3.2 Estimação do GARCH(1,1) no OxMetrics para todas as distribuições

Estimando GARCH(1,1) no OxMetrics usando as distribuições: Normal, t-Student, GED e SkSt, usando o programa:

GARCH Models in OxMetrics

obtém-se os seguintes resultados:

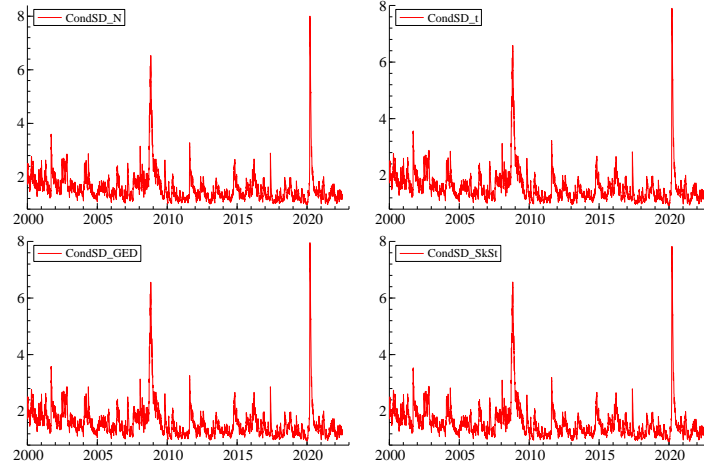
Tabela 11: Comparação dos Modelos GARCH(1,1) para as quatro distribuições

Progress to date							
Model	T	p		log-likelihood	SC	HQ	AIC
N-GARCH(1,1)	5879	3	BFGS	-10820.209	3.6869	3.6839	3.6823
t-GARCH(1,1)	5879	4	BFGS	-10731.961	3.6583	3.6546	3.6526
GED-GARCH(1,1)	5879	4	BFGS	-10744.309	3.6625	3.6588	3.6568
SkSt-GARCH(1,1)	5879	5	BFGS	-10722.442	3.6566	3.6521	3.6498

Temos que o melhor modelo é exatamente quando se utiliza a distribuição *SkStudent* – *t*.

E a estimativa da Variância Condiconal é apresentada no gráfico abaixo

Figura 20: Volatilidade para todos os GARCH(1,1) para IBOVESPA



Pelos resultados das tabelas (12, 13 e 14) temos que normalidade é rejeitada, modelo para média condicional é adequado, e modelo para a variância condicional é inadequado para defasagens 5, 10 e 20 mas não apresenta correlação serial para defasagens até 50.

Tabela 12: Teste de Normalidade para o melhor Modelo GARCH - SkSt

Normality Test			
	Statistic	t-Test	P-Value
Skewness	-0.37448	11.725	0.00000
Excess Kurtosis	1.8065	28.286	0.00000
Jarque-Bera	936.80	<i>NaN</i>	0.00000

Tabela 13: Teste de Box-Pierce para Resíduos Padronizados do melhor Modelo GARCH - SkSt

Q-Statistics on Standardized Residuals			
$Q(5)$	=	1.62902	[0.8977179]
$Q(10)$	=	6.19838	[0.7983297]
$Q(20)$	=	20.2859	[0.4401757]
$Q(50)$	=	40.1254	[0.8397099]
H_0 : No serial correlation => Accept H_0 when prob. is High			

Tabela 14: Teste de Box-Pierce para Resíduos Padronizados ao Quadrado do melhor Modelo GARCH - SkSt

Q-Statistics on Squared Standardized Residuals			
=> P-values adjusted by 1 degree(s) of freedom			
$Q(5)$	=	17.2073	[0.0006406] **
$Q(10)$	=	25.6674	[0.0011971] **
$Q(20)$	=	35.0966	[0.0091913] **
$Q(50)$	=	58.5717	[0.1410657]
H_0 : No serial correlation => Accept H_0 when prob. is High			

5.3.3 Estimação do GARCH(1,1) no R para todas as distribuições

Estimando GARCH(1,1) no R usando as distribuições: Normal, t-Student, GED e SkSt, usando o programa RUGARCH:

GARCH Models using RUGARCH in R

obtém-se os seguintes resultados:

Tabela 15: Comparação dos Modelos GARCH(1,1) para as quatro distribuições no R

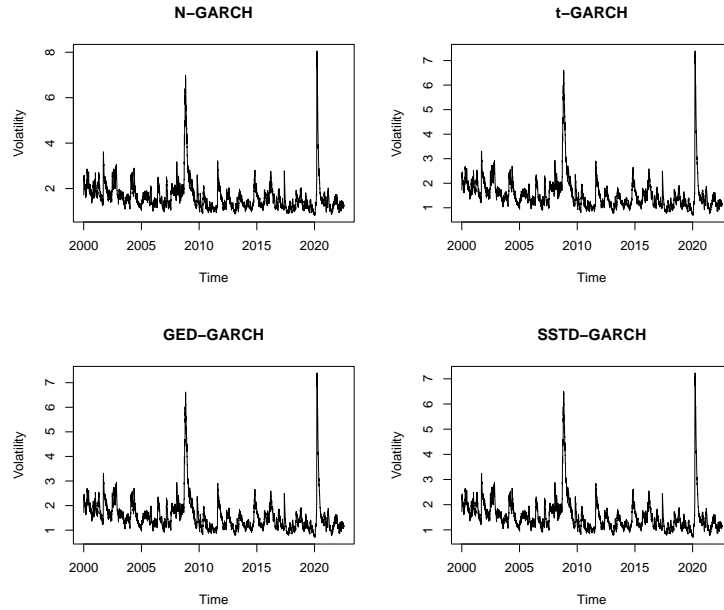
Progress to date							
Model	T	p		log-likelihood	SC	HQ	AIC
N-GARCH(1,1)	5879	3	HYBRID	-10841.04	3.6940	3.6910	3.6894
t-GARCH(1,1)	5879	4	HYBRID	-10767.68	3.6705	3.6667	3.6648
GED-GARCH(1,1)	5879	4	HYBRID	-10777.81	3.6739	3.6702	3.6682
SkSt-GARCH(1,1)	5879	5	HYBRID	-10751.43	3.6664	3.6620	3.6596

Note: HYBRID is the solver in rugarch that uses many solvers from optimization routines in R.

Temos que o melhor modelo é exatamente quando se utiliza a distribuição *SkStudent* – *t*.

As estimativas das Volatilidades (desvios padrões) Condicionais são apresentadas no gráfico abaixo:

Figura 21: Cond Volatilidade para os Modelos GARCH(1,1) com todas as distribuições no R



Obviamente os resultados corroboram os obtidos com o OxMetrics. Resultados semelhantes as tabelas (6, 7 e 8) temos que normalidade é rejeitada, modelo para média condicional é adequado todas as defasagem de ordens 1, 5 e 9, e modelo para a variância condicional é inadequado para defasagens 5 e 9, mas não apresenta correlação serial para defasagens até 1 com p-valor de 0,76.

5.4 Modelos IGARCH

Se $\sum_{i=1}^m \alpha_i + \sum_{j=1}^s \beta_j = 1$ então o processo $ARMA$ para r_t^2 possui uma raiz unitária. Nelson [1991] chamou o modelo em que esta condição se verifica de $GARCH$ integrado ou $IGARCH$. Se r_t segue um $IGARCH$, então sua variância não condicional é infinita e os processos r_t e r_t^2 não satisfazem a definição de processos covariância estacionários. No entanto, ainda é possível que r_t atende as condições de estacionaridade estrita no sentido de que sua den-

sidade não condicional não varia no tempo. Esta propriedade é demonstrada em Nelson [1990].

O processo EWMA definido anteriormente é, portanto, um IGARCH com $\omega = 0$ com os parâmetros fixados arbitrariamente. Neste modelo, os retornos teriam distribuição degenerada, isto é, sua distribuição não condicional teria média zero e variância indeterminada.

Desta forma, ao assumir que os retornos de determinado ativo tem variância condicional definida por um EWMA está-se, na verdade, fazendo uma restrição que pode não ser justificada pelos dados.

Metodologicamente, o procedimento mais indicado seria estimar um GARCH e então testar a hipótese de que os parâmetros aceitam uma restrição do tipo $\omega = 0$, $\alpha + \beta = 1$. Isto garantiria a estimação de um modelo coerente com a realização do processo estocástico dos retornos. Entretanto, testar estas duas restrições, conjuntamente, é complicado, pois, sob a hipótese nula, a distribuição seria degenerada.

Em (32) o $GARCH(1, 1)$ foi escrito como um $ARMA(1, 1)$ para o caso geral de um $GARCH(s, m)$ teríamos:

$$\begin{aligned} r_t^2 &= \omega + (\alpha_1(L) + \beta_1(L))r_t^2 + [1 - \beta_1(L)](r_t^2 - \sigma_t^2) \\ [1 - \alpha_1(L) - \beta_1(L)]r_t^2 &= \omega + [1 - \beta_1(L)](r_t^2 - \sigma_t^2) \end{aligned} \quad (36)$$

e quando $[1 - \alpha_1(L) - \beta_1(L)]$ tem uma raiz unitária este polinômio pode ser escrito como $\phi(L)(1 - L)$ onde $\phi(L) = [1 - \alpha_1(L) - \beta_1(L)](1 - L)^{-1}$ cuja ordem é $\max\{s, m\} - 1$. Podemos re-escrever (36) da seguinte forma:

$$\phi(L)(1 - L)r_t^2 = \omega + [1 - \beta_1(L)](r_t^2 - \sigma_t^2) \quad (37)$$

E (37) pode ser re-escrito para expressar a variância condicional em função dos resíduos ao quadrado, uma vez que $r_t = a_t$, temos que

$$\begin{aligned}
[1 - \beta_1(L)]\sigma_t^2 &= \omega + [1 - \beta_1(L) - \phi(L)(1 - L)]a_t^2 \\
\sigma_t^2 &= \omega[1 - \beta_1(L)]^{-1} \\
&\quad + [1 - \phi(L)(1 - L)][1 - \beta_1(L)]^{-1}a_t^2 \quad (38)
\end{aligned}$$

que é a representação $ARCH(\infty)$.

Este tipo de especificação do $IGARCH$ é útil quando for apresentado o modelo de memória longa para a volatilidade.

5.4.1 Programa no EVIEWS para Estimar um IGARCH

O EVIEWS não estima um modelo $IGARCH$ mas sim um $EWMA$. Deste modo iremos pular esta parte no EVIEWS.

As linhas 106 até 116 do programa

Github-GARCH-EVIEWS

estima este modelo (o $EWMA$) mas os resultados não fazem sentido.

5.4.2 Estimação do IGARCH no OxMetrics

O programa

Ox-IGARCH

estima os modelos $IGARCH$ no OxMetrics usando as mesmas quatro distribuições que no caso do $GARCH$.

Estimando $IGARCH(1,1)$ no OxMetrics usando as distribuições: Normal, t-Student, GED e SkSt, usando o programa obtém-se os seguintes resultados:

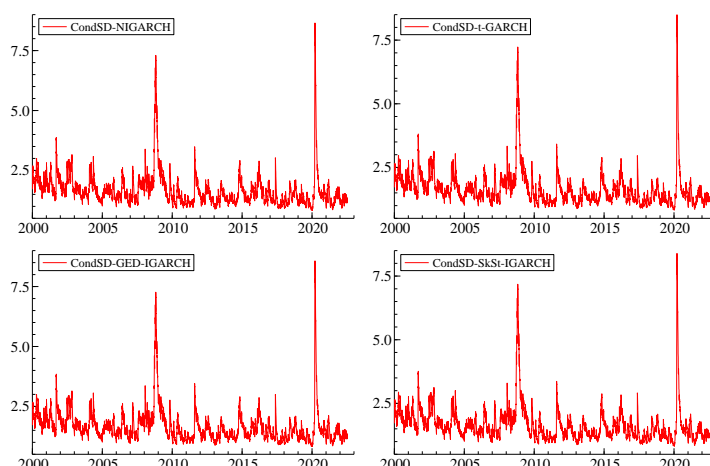
Tabela 16: Comparação dos Modelos $IGARCH(1,1)$ para as quatro distribuições

Progress to date							
Model	T	p		log-likelihood	SC	HQ	AIC
N-GARCH(1,1)	5879	3	MaxSA	-10841.314	3.6926	3.6904	3.6892
t-GARCH(1,1)	5879	4	MaxSA	-10742.063	3.6603	3.6573	3.6557
GED-GARCH(1,1)	5879	4	MaxSA	-10757.481	3.6655	3.6626	3.6610
SkSt-GARCH(1,1)	5879	5	MaxSA	-10732.353	3.6585	3.6548	3.6528

Temos que o melhor modelo é exatamente quando se utiliza a distribuição *SkStudent* – *t*.

E a estimativa da Variância Condiciona é apresentada no gráfico abaixo

Figura 22: Volatilidade para todos os GARCH(1,1) para IBOVESPA



Pelos resultados das tabelas (17, 18 e 19) temos que normalidade é rejeitada, modelo para média condicional é adequado, e modelo para a variância condicional é inadequado para defasagens 5, 10 e 20 mas não apresenta correlação serial para defasagens até 50.

Tabela 17: Teste de Normalidade para o melhor Modelo IGARCH - SkSt

Normality Test			
	Statistic	t-Test	P-Value
Skewness	-0.40045	12.538	0.00000
Excess Kurtosis	2.0019	31.345	0.00000
Jarque-Bera	1138.8	<i>NaN</i>	0.00000

Tabela 18: Teste de Box-Pierce para Resíduos Padronizados do melhor Modelo IGARCH - SkSt

Q-Statistics on Standardized Residuals			
$Q(5)$	=	1.38913	[0.9254986]
$Q(10)$	=	5.67290	[0.8419554]
$Q(20)$	=	19.2354	[0.5065754]
$Q(50)$	=	38.8643	[0.8730499]
H_0 : No serial correlation => Accept H_0 when prob. is High			

Tabela 19: Teste de Box-Pierce para Resíduos Padronizados ao Quadrado do melhor Modelo IGARCH - SkSt

Q-Statistics on Squared Standardized Residuals			
=> P-values adjusted by 1 degree(s) of freedom			
$Q(5)$	=	12.4842	[0.0058958] **
$Q(10)$	=	20.3796	[0.0089913] **
$Q(20)$	=	33.9550	[0.0127577] *
$Q(50)$	=	60.8866	[0.1003065]
H_0 : No serial correlation => Accept H_0 when prob. is High			

Observe que os mesmos comentários feitos sobre a adequabilidade dos modelos na classe *GARCH*, valem aqui também.

Comparando os modelos *GARCH* e *IGARCH* através dos critérios de informação, temos que os modelos *GARCH* são os preferíveis e até aqui o melhor é o modelo *GARCH*(1,1) com distribuição SkSt.

5.4.3 Estimação do IGARCH no R

O pacote RUGARCH estima o modelo IGARCH, tornando fixo o parâmetro $\omega = 0$. Temos o seguinte programa:

IGARCH using RUGARCH in R

estima os modelos *IGARCH* no R usando as mesmas quatro distribuições que no caso do *GARCH*.

Estimando IGARCH(1,1) no R usando as distribuições: Normal, t-Student, GED e SkSt, usando o programa obtém-se os seguintes resultados:

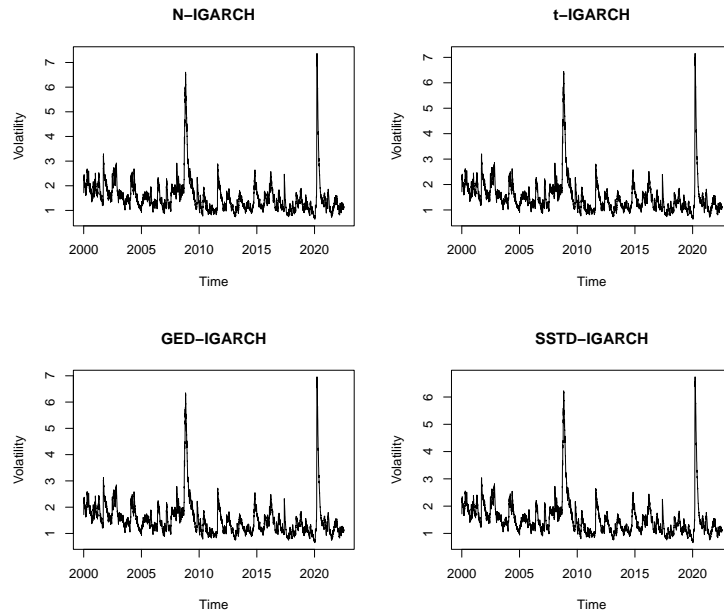
Tabela 20: Comparação dos Modelos IGARCH(1,1) para as quatro distribuições

Progress to date							
Model	T	p		log-likelihood	SC	HQ	AIC
N-GARCH(1,1)	5879	3	HYBRID	-10898.77	3.7121	3.7099	3.7087
t-GARCH(1,1)	5879	4	HYBRID	-10786.99	3.6756	3.6726	3.6710
GED-GARCH(1,1)	5879	4	HYBRID	-10802.45	3.6808	3.6779	3.6763
SkSt-GARCH(1,1)	5879	5	HYBRID	-10770.17	3.6713	3.6676	3.6656

Temos que o melhor modelo é exatamente quando se utiliza a distribuição *SkStudent* – *t*.

E a estimativa da Variância Condicional é apresentada no gráfico abaixo

Figura 23: Volatilidade para todos os IGARCH(1,1) para IBOVESPA



Pelos resultados das tabelas (17, 18 e 19) temos que normalidade é rejeitada, modelo para média condicional é adequado, e modelo para a variância condicional é inadequado para defasagens 5, 10 e 20 mas não apresenta correlação serial para defasagens até 50.

5.5 Outros Membros da família XARCH

Um dos problemas com o modelo GARCH é a resposta simétrica, na volatilidade, de grandes retornos positivos ou negativos. Essa simetria não permite acomodar um dos fatos estilizados presentes em dados financeiros: o mercado tem baixa volatilidade quando está subindo e alta volatilidade quando está em queda.

5.5.1 Modelos EGARCH

Nelson [1990] sugeriu o modelo EGARCH – GARCH exponencial –, em que o logaritmo da variância condicional é dado por:

$$\ln(\sigma_t^2) = \omega + \beta \ln(\sigma_{t-1}^2) + \alpha \left| \frac{a_{t-1}}{\sigma_{t-1}} \right| + \gamma \frac{a_{t-1}}{\sigma_{t-1}} \quad (39)$$

Nesse modelo, os parâmetros são irrestritos, sendo uma vantagem em relação à especificação GARCH. O parâmetro γ , por ser negativo, faz com que a volatilidade aumente quando os retornos são negativos.

Existem outros modelos que também incorporam efeitos de assimetria, como, por exemplo, os modelos de mudança markoviana de regime. Uma referência ao seu uso e aplicado a dados brasileiros pode ser encontrado em Valls Pereira et al. [1999]. Outro modelo que mantém a estrutura GARCH mas permite assimetria será apresentada na próxima seção e é conhecido por Limiar GARCH.

5.5.2 Programa no EVIEWS para estimar um EGARCH

As linhas 119 até 129 apresentam o programa para a estimação do EGARCH no EVIEWS.

Temos então os seguintes resultados para a estimativa dos parâmetros

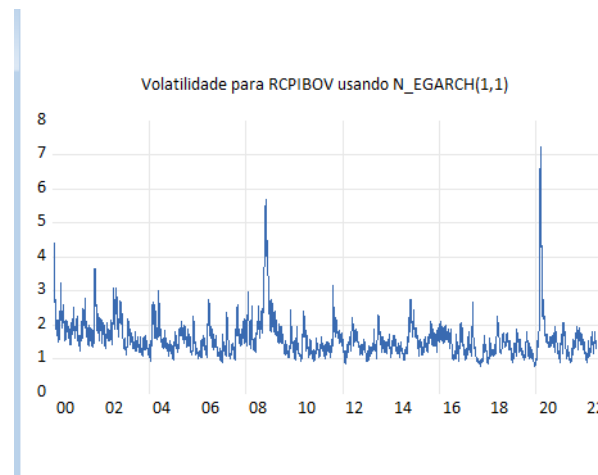
Tabela 21: Estimativa do Modelo N-EGARCH no Eviews

Dependent Variable: DLIBOVM				
Method: ML ARCH - Normal distribution (BFGS / Marquardt steps)				
Date: 08/24/22 Time: 23:49				
Sample (adjusted): 1/04/2000 7/15/2022				
Included observations: 5879 after adjustments				
Failure to improve likelihood (non-zero gradients) after 29 iterations				
Coefficient covariance computed using outer product of gradients				
Presample variance: backcast (parameter = 0.7)				
LOG(GARCH) = C(2) + C(3)*ABS(RESID(-1)/@SQRT(GARCH(-1))) + C(4)				
*RESID(-1)/@SQRT(GARCH(-1)) + C(5)*LOG(GARCH(-1))				
Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	1.85E - 14	0.018055	1.03E - 12	1.0000
Variance Equation				
C(2)	-0.082186	0.006600	-12.45321	0.0000
C(3)	0.138319	0.009225	14.99388	0.0000
C(4)	-0.080870	0.005686	-14.22266	0.0000
C(5)	0.970200	0.002771	350.1622	0.0000
R-squared	0.000000	Mean dependent var	-2.52E - 16	
Adjusted R-squared	0.000000	S.D. dependent var	1.715912	
S.E. of regression	1.715912	Akaike info criterion	3.665913	
Sum squared resid	17306.91	Schwarz criterion	3.671594	
Log likelihood	-10770.95	Hannan-Quinn criter.	3.667888	
Durbin-Watson stat	2.055935			

Observe que o parâmetro de assimetria é negativo e significativo, indicando uma volatilidade maior para retornos positivos.

Abaixo é apresentado o gráfico da volatilidade estimada por este modelo

Figura 24: Volatilidade para todos os N-EGARCH(1,1) para IBOVESPA



A adequabilidade do modelo será verificada através da FAC e FACP para os resíduos padronizados, caso não apresentem cor-

relação serial há indicação de que a equação para a média condicional está bem especificada e se a FAC e FACP para os resíduos padronizados ao quadrado não tem estrutura de correlação é uma indicação de que a equação para a variância condicional está bem especificada.

Nas linhas 132 até 142 do programa Github-GARCH-EVIEWS temos os comandos do EVIEWS

Figura 25: FAC e FACP para Resíduos Padronizados para o modelo N-EGARCH(1,1) para IBOVESPA

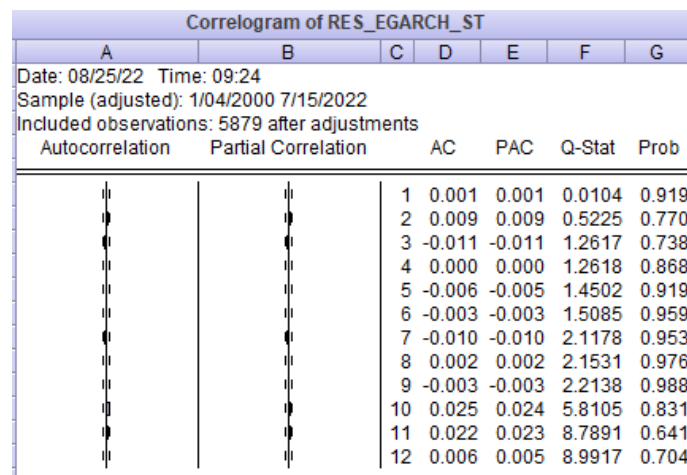
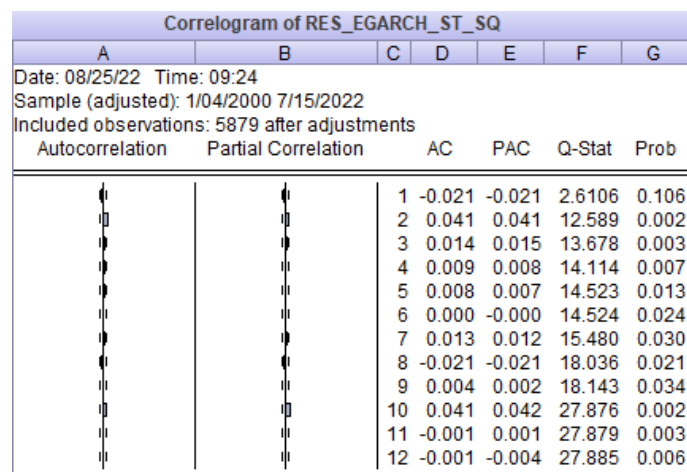


Figura 26: FAC e FACP para Resíduos Padronizados ao quadrado para o modelo N-EGARCH(1,1) para IBOVESPA



Observe que tanto o modelo para a média está bem especificados, mas quanto para a variância condicional não. Mas pelo critério de informação, 3.671594, este modelo é melhor do que o modelo *GARCH* que tinha um critério de Schwartz de 3.686332, superior ao do *EGARCH*. Portanto o modelo com assimetria é melhor do que os anteriores.

5.5.3 Modelo EGARCH no OxMetrics

A especificação utilizada no OxMetrics é diferente da original dada por (39) e é devido a Bollerslev & Mikkelsen (1996) sendo dada por:

$$\ln(\sigma_t^2) = \omega + [1 - \beta(L)]^{-1}[1 + \alpha(L)]g(\varepsilon_{t-1}) \quad (40)$$

onde o valor de $g(\varepsilon_t) = g(\frac{a_t}{\sigma_t})$ depende de vários fatores. Primeiro o valor de $g(\varepsilon_t)$ deve ser função tanto da magnitude quanto do sinal de ε_t . Podemos então expressar $g(\varepsilon_t)$ da seguinte forma:

$$g(\varepsilon_t) = \underbrace{\theta_1 \varepsilon_t}_{\text{efeito sinal}} + \underbrace{\theta_2[|\varepsilon_t| - E|\varepsilon_t|]}_{\text{efeito magnitude}} \quad (41)$$

onde $E|\varepsilon_t|$ depende das hipóteses feitas sobre a distribuição não condicional de ε_t . Por exemplo se a distribuição for normal temos que

$$E[|\varepsilon_t|] = \sqrt{\frac{2}{\pi}} \quad (42)$$

Se a distribuição for t - *Student* assimétrica temos que

$$E[|\varepsilon_t|] = \frac{4\xi^2}{\xi + \frac{1}{\xi}} \frac{\Gamma\left(\frac{1+v}{2}\right) \sqrt{v-2}}{\sqrt{\pi} \Gamma\left(\frac{v}{2}\right)} \quad (43)$$

e quando $\xi = 1$ temos a t - *Student* simétrica.

Se a distribuição for GED, temos que

$$E[|\varepsilon_t|] = 2^{(1/v)} \lambda \frac{\Gamma\left(\frac{2}{v}\right)}{\Gamma\left(\frac{1}{v}\right)} \quad (44)$$

onde ξ , v e λ foram definidos anteriormente

O programa Ox-EGARCH estima os modelos *EGARCH* no OxMetrics usando as mesmas quatro distribuições que no caso do *GARCH*.

Estimando EGARCH(1,1) no OxMetrics usando as distribuições: Normal, t-Student, GED e SkSt, usando o programa obtém-se os seguintes resultados:

Tabela 22: Comparação dos Modelos EGARCH(1,1) para as quatro distribuições

Progress to date							
Model	T	p		log-likelihood	SC	HQ	AIC
N-EGARCH(1,1)	5879	6	MaxSA	-10794.946	3.6812	3.6768	3.6744
t-EGARCH(1,1)	5879	7	MaxSA	-10704.342	3.6519	3.6467	3.6439
GED-EGARCH(1,1)	5879	7	MaxSA	-10721.025	3.6576	3.6524	3.6496
SkSt-EGARCH(1,1)	5879	8	MaxSA	-10694.469	3.6500	3.6441	3.6409

Temos que o melhor modelo é exatamente quando se utiliza a distribuição *SkStudent* – *t*. As estimativas são apresentadas abaixo.

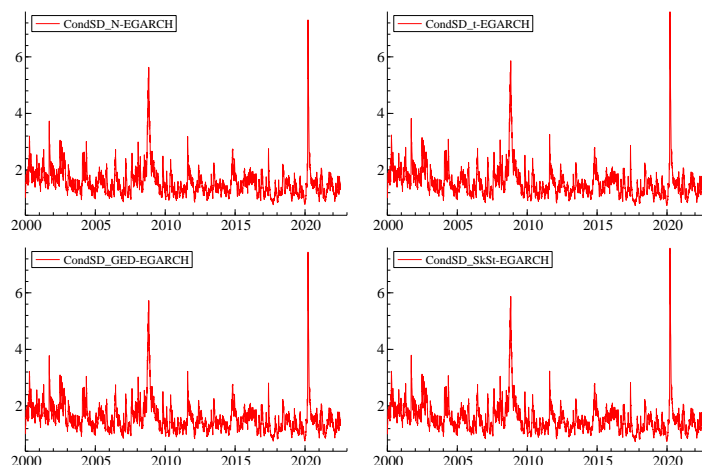
Tabela 23: Estimativa do Modelo SkSt-EGARCH no OxMetrics

Dependent Variable: DLIBOVM				
Method: MaxSA algorithm				
Sample (adjusted): 1/04/2000 7/15/2022				
Included observations: 5879 after adjustments				
Coefficient covariance computed using Robust Standard Errors (Sandwich formula)				
Variable	Coefficient	Std. Error	t-value	p-value.
Cst(M)	0.0282	0.0127	2.230	0.0230
Variance Equation				
Cst(V)	0.1538	0.2022	0.7608	0.4468
α	0.4644	0.3597	1.291	0.1968
β	0.9728	0.0044	125.8	0.0000
θ_1	-0.0547	0.0151	-3.622	0.0003
θ_2	0.1055	0.0233	4.524	0.0000
$\log(\xi)$	-0.0820	0.0178	-4.605	0.0000
v	9.1510	1.0858	8.428	0.0000
Mean dependent var	-2.52E - 16	S.D. dependent var		1.715912
S.E. of regression	1.715912	Akaike info criterion		3.6409
Sum squared resid	17306.91	Schwarz criterion		3.67500
Log likelihood	-10694.469	Hannan-Quinn criter.		3.6441

Observe que θ_1 é significativo assim como θ_2 , implicando a existência de ambos os efeitos sinal e magnitude na assimetria.

E a estimativa da Variância Condiconal para todas as distriuições são apresentadas no gráfico abaixo.

Figura 27: Volatilidade para todos os EGARCH(1,1) para IBOVESPA



Pelos resultados das tabelas (24, 25 e 26) temos que normalidade é rejeitada, modelo para média condicional é adequado, e modelo para a variância condicional é inadequado para defasagem 10, mas não apresenta correlação serial para defasagens até 5, 20 e 50.

Tabela 24: Teste de Normalidade para o melhor Modelo EGARCH - SkSt

Normality Test			
	Statistic	t-Test	P-Value
Skewness	-0.343335	10.750	0.00000
Excess Kurtosis	1.9672	30.802	0.00000
Jarque-Bera	1063.4	<i>NaN</i>	0.00000

Tabela 25: Teste de Box-Pierce para Resíduos Padronizados do melhor Modelo IGARCH - SkSt

Q-Statistics on Standardized Residuals		
$Q(5)$	= 1.21955	[0.9429879]
$Q(10)$	= 5.33293	[0.8678581]
$Q(20)$	= 19.1103	[0.5146647]
$Q(50)$	= 40.6203	[0.8254108]
H_0 : No serial correlation => Accept H_0 when prob. is High		

Tabela 26: Teste de Box-Pierce para Resíduos Padronizados ao Quadrado do melhor Modelo IGARCH - SkSt

Q-Statistics on Squared Standardized Residuals		
=> P-values adjusted by 1 degree(s) of freedom		
$Q(5)$	= 7.16531	[0.0668114]
$Q(10)$	= 16.9962	[0.0301492]*
$Q(20)$	= 21.7301	[0.244190]
$Q(50)$	= 46.0319	[0.5538293]
H_0 : No serial correlation => Accept H_0 when prob. is High		

5.5.4 Modelo GJR ou Limiar GARCH no EVIEWS

Este modelo foi proposto por Glosten et al. [1993] e a sua especificação

$$\sigma_t^2 = \omega + \sum_{i=1}^q (\alpha_i a_{t-i}^2 + \gamma_i S_{t-i}^- a_{t-i}^2) + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 \quad (45)$$

onde $S_{t-i}^- = \begin{cases} 1 & \text{se } a_{t-i} < 0 \\ 0 & \text{se } a_{t-i} \geq 0 \end{cases}$ fazendo com que a volatilidade

aumente mais para retornos negativos do que para retornos positivos.

Nas linhas 145 até 168 do programa Github-GARCH-EVIEWS temos os comandos do EVIEWS e obtém-se os seguintes resultados:

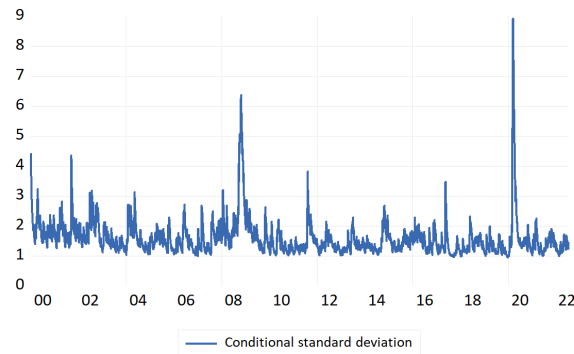
Tabela 27: Estimação do Modelo Limiar GARCH no EVIEWS

Dependent Variable: DLIBOVM				
Method: ML ARCH - Normal distribution (BFGS / Marquardt steps)				
Date: 08/25/22 Time: 09:38				
Sample (adjusted): 1/04/2000 7/15/2022				
Included observations: 5879 after adjustments				
Failure to improve likelihood (non-zero gradients) after 29 iterations				
Coefficient covariance computed using outer product of gradients				
Presample variance: backcast (parameter = 0.7)				
GARCH = C(2) + C(3)*RESID(-1)^2 + C(4)*RESID(-1)^2*(RESID(-1) < 0) +				
C(5)*GARCH(-1)				
Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-6.47E - 14	0.018791	-3.44E - 12	1.0000
Variance Equation				
C	0.080012	0.007743	10.33368	0.0000
RESID(-1)^2	0.020855	0.006155	3.388345	0.0007
RESID(-1)^2*(RESID(-1)>0)	0.103922	0.008523	12.19299	0.0000
GARCH(-1)	0.894772	0.007049	126.9341	0.0000
R-squared	0.000000	Mean dependent var	-2.52E - 16	
Adjusted R-squared	0.000000	S.D. dependent var	1.715912	
S.E. of regression	1.715912	Akaike info criterion	3.665294	
Sum squared resid	17306.91	Schwarz criterion	3.670975	
Log likelihood	-10769.13	Hannan-Quinn criter.	3.667269	
Durbin-Watson stat	2.055935			

Observe que o parâmetro de assimetria é positivo e significativo, indicando que esta série apresenta assimetria na volatilidade. Observe também que a volatilidade para retornos positivos é dada por $0.020855 + 0.894772 = 0.915627$ e a volatilidade para retornos negativos é dada por $0.020855 + 0.894772 + 0.5 * (0.103922) = 0.9967588$, indicando que para retornos positivos temos um GARCH e para retornos negativos um GARCH com maior persistência.

A volatilidade extraída deste modelo é apresentada no gráfico abaixo

Figura 28: Volatilidade do N- GJR(1,1) para IBOVESPA



5.5.5 Modelo GJR no OxMetrics

O programa Github-Ox-GJR estimando os modelos GJR no Ox-Metrics com as mesmas quatro distribuições como anteriormente. Obtém-se os seguintes resultados:

Tabela 28: Comparação dos Modelos GJR(1,1) para as quatro distribuições

Progress to date							
Model	T	p		log-likelihood	SC	HQ	AIC
N-EGARCH(1,1)	5879	5	MaxSA	-10772.305	3.6721	3.6683	3.6664
t-EGARCH(1,1)	5879	6	MaxSA	-10698.513	3.6484	3.6440	3.6416
GED-EGARCH(1,1)	5879	6	MaxSA	-10709.373	3.6521	3.6477	3.6453
SkSt-EGARCH(1,1)	5879	7	MaxSA	-10687.541	3.6462	3.6410	3.6382

Temos que o melhor modelo é exatamente quando se utiliza a distribuição *SkStudent* - *t*. As estimativas são apresentadas abaixo.

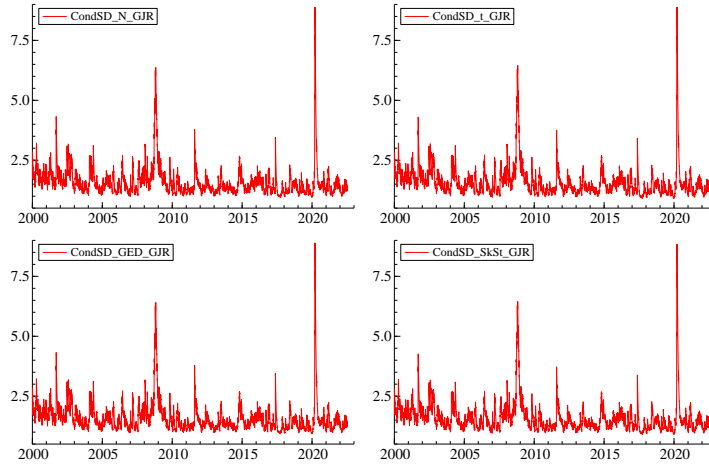
Tabela 29: Estimativa do Modelo SkSt-GJR no OxMetrics

Dependent Variable: DLIBOVM				
Method: MaxSA algorithm				
Sample (adjusted): 1/04/2000 7/15/2022				
Included observations: 5879 after adjustments				
Coefficient covariance computed using Robust Standard Errors (Sandwich formula)				
Variable	Coefficient	Std. Error	t-value	p-value.
Cst(M)	0.0292	0.0184	1.592	0.1115
Variance Equation				
Cst(V)	0.0597	0.0142	4.200	0.000
α	0.0222	0.0066	3.354	0.0008
β	0.9065	0.0132	68.71	0.0000
γ	0.0951	0.0159	5.979	0.0000
$\log(\xi)$	-0.0866	0.0179	-4.830	0.0000
v	9.7808	1.2229	7.998	0.0000
Mean dependent var	0.02961	Variance. dependent var		2.9439
S.E. of regression	1.715912	Akaike info criterion		3.6382
Sum squared resid	17306.91	Schwarz criterion		3.6462
Log likelihood	-10687.541	Hannan-Quinn criter.		3.6382

Observe que γ_1 é significativo, implicando a existência de assimetria.

E a estimativa da Variância Condicional para todas as distriuições são apresentadas no gráfico abaixo.

Figura 29: Volatilidade para todos os GJR(1,1) para IBOVESPA



Pelos resultados das tabelas (30, 31 e 32) temos que normalidade é rejeitada, modelo para média condicional é adequado, e modelo para a variância condicional é inadequado para defasagem 10, mas não apresenta correlação serial para defasagens até 5, 20 e 50.

Tabela 30: Teste de Normalidade para o melhor Modelo GJR - SkSt

Normality Test			
	Statistic	t-Test	P-Value
Skewness	-0.33092	10.361	0.00000
Excess Kurtosis	1.5595	24.419	0.00000
Jarque-Bera	706.06	<i>NaN</i>	0.00000

Tabela 31: Teste de Box-Pierce para Resíduos Padronizados do melhor Modelo GJR - SkSt

Q-Statistics on Standardized Residuals			
$Q(5)$	=	1.43250	[0.9207253]
$Q(10)$	=	5.66828	[0.8423206]
$Q(20)$	=	19.4846	[0.4905568]
$Q(50)$	=	40.8837	[0.8175311]
H_0 : No serial correlation => Accept H_0 when prob. is High			

Tabela 32: Teste de Box-Pierce para Resíduos Padronizados ao Quadrado do melhor Modelo GJR - SkSt

Q-Statistics on Squared Standardized Residuals			
=> P-values adjusted by 1 degree(s) of freedom			
$Q(5)$	=	9.60522	[0.022379]*
$Q(10)$	=	19.4983	[0.0124510]*
$Q(20)$	=	25.7886	[0.1046916]
$Q(50)$	=	52.3600	[0.3086095]
H_0 : No serial correlation => Accept H_0 when prob. is High			

Observe que o parâmetro de assimetria é positivo e significativo, indicando que esta série apresenta assimetria na volatilidade. Observe também que a volatilidade para retornos positivos é dada por $0.022217 + 0.906480 = 0.928697$ e a volatilidade para retornos negativos é dada por $0.022217 + 0.906480 + 0.5 * (0.095064) = 0.976229$, indicando que para retornos positivos a volatilidade segue um GARCH e para retornos negativos segue também um GARCH com uma maior persistência.

5.6 Modelo APARCH

O modelo $APARCH(p, q)$ introduzido por Ding et al. [1993] tem a seguinte especificação:

$$\sigma_t^\delta = \omega + \sum_{i=1}^q \alpha_i (|\varepsilon_{t-i}| - \gamma_i \varepsilon_{t-i})^\delta + \sum_{j=1}^p \beta_j \sigma_{t-j}^\delta \quad (46)$$

onde $\delta > 0$ e $-1 < \gamma_i < 1$ ($i = 1, \dots, q$).

O parâmetro δ é semelhante a transformação de Box-Cox aplicada a σ_t , e γ_i representam os efeitos alavancagens.

Ding et al. [1993] mostram que este modelo inclui sete extensões do modelo *ARCH* como casos especiais, a saber:

- (i) O modelo *ARCH* de Engle [1982] quando $\delta = 2$, $\gamma_i = 0$ ($i = 1, \dots, p$) e $\beta_j = 0$ ($j = 1, \dots, p$).
- (ii) O modelo *GARCH* de Bollerslev [1986] quando $\delta = 2$ e $\gamma_i = 0$ ($i = 1, \dots, p$).
- (iii) *GARCH* de Taylor [1986] quando $\delta = 1$ e $\gamma_i = 0$ ($i = 1, \dots, p$).
- (iv) O *TARCH* de Zakoian [1994] quando $\delta = 2$.
- (v) O *GJR* de Glosten et al. [1993] quando $\delta = 2$.
- (vi) O *NARCH* de Higgins and Bera [1992] quando $\gamma_i = 0$ ($i = 1, \dots, p$) e $\beta_j = 0$ ($j = 1, \dots, p$).
- (vii) O *Log-ARCH* de Geweke [1986] quando $\delta \rightarrow 0$.

Se $\omega > 0$ e $\sum_{i=1}^q \alpha_i E(|z_t| - \gamma_i z_t)^\delta + \sum_{j=1}^p \beta_j < 1$ a solução estacionária de 46 é dada por:

$$E(\sigma_t^\delta) = \frac{\omega}{1 - \sum_{i=1}^q \alpha_i E(|z_t| - \gamma_i z_t)^\delta + \sum_{j=1}^p \beta_j} \quad (47)$$

O programa Ox-GARCH-APARCH estima o modelo APARCH para as quatro distribuições. Temos os seguintes critérios de informações:

Tabela 33: Comparação dos Modelos APARCH(1,1) para as quatro distribuições

Progress to date							
Model	T	p		log-likelihood	SC	HQ	AIC
N-APARCH(1,1)	5879	6	BFGS	-10767.014	3.6717	3.6673	3.6649
t-APARCH(1,1)	5879	6	BFGS	-10704.370	3.6519	3.6467	3.6439
GED-APARCH(1,1)	5879	6	BFGS	-10704.370	3.6519	3.6467	3.6439
SkSt-APARCH(1,1)	5879	7	BFGS	-10681.404	3.6456	3.6396	3.6365

Temos que o melhor modelo é exatamente quando se utiliza a distribuição *SkStudent* – *t*. As estimativas são apresentadas na tabela 34 abaixo.

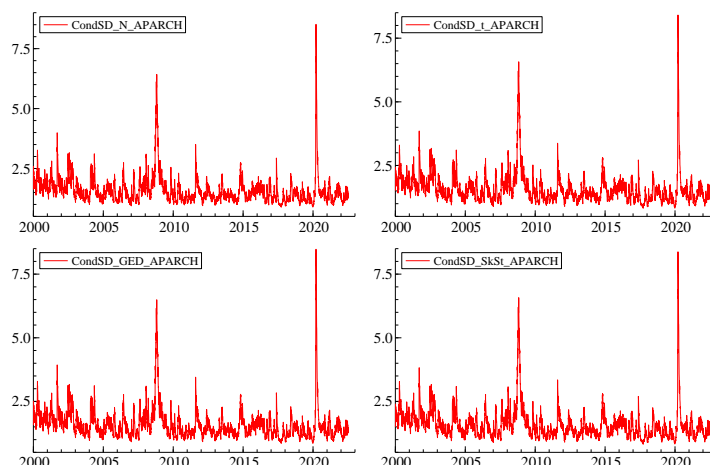
Tabela 34: Estimativa do SkSt-APARCH(1,1) para IBOVESPA

The dataset is: C:\Users\Pedro\Dropbox\Topicos_em_Financas_2022\Dados\Dados-BR.xlsx				
The estimation sample is : 2000-01-04 - 2022-07-15				
The dependent variable is; absDLIBOVM				
Mean Equation: ARMA(0,0) Model				
No regressor in the conditional mean				
Variance Equation: APARCH(1,1) Model				
No regressor in the conditional variance				
Skewed Student distribution with 9.74016 degrees of freedom.				
and Asymetric doefficient ($\log(\xi_i)$) = -0.0867677.				
Robust Standard Errors				
	Coefficient	Std.Error	t-value	t-prob
Const(M)	-0.001296	0.017892	-0.07245	0.9422
Const(V)	0.045722	0.011569	3.952	0.0001
ARCH(α_1)	0.068178	0.0096455	7.068	0.000
GARCH(β_1)	0.915181	0.012614	72.55	0.0000
APARCH(γ_1)	0.545480	0.087648	6.224	0.0000
APARCH(δ)	1.387327	0.16033	8.653	0.0000
Asymmetry	-0.086768	0.017974	-4.827	0.0000
Tail	9.740161	1.2151	8.016	0.0000
loglikelihood -10681.404				
n ^o of observations	5879	n ^o of parameters:		8
AIC	3.636470	BIC		3.636466
mean(DLIBOVM)	-0.0000	var(DLIBOVM)		2.94385
Skewness(DLIBOVM)	-0.39646	Kurtosis(DLIBOVM)		10.36097
The condition for existence of $E(\sigma^\delta)$ and $E(\varepsilon^\delta)$ is observed.				
The constraint equals 0.985118 and should be ≥ 1 .				

Observe que γ_1 é significativo, implicando a existência de assimetria. Observe também que $\delta < 2$.

E a estimativa da Variância Condicional para todas as distriuições são apresentadas no gráfico abaixo.

Figura 30: Volatilidade para todos os APARCH(1,1) para IBOVESPA



Pelos resultados das tabelas (35, 36 e 37) temos que normalidade é rejeitada, modelo para média condicional é adequado, e modelo para a variância condicional é inadequado para defasagem 10, mas não apresenta correlação serial para defasagens até 5, 20 e 50.

Tabela 35: Teste de Normalidade para o melhor Modelo APARCH - SkSt

Normality Test			
	Statistic	t-Test	P-Value
Skewness	-0.32161	10.070	0.00000
Excess Kurtosis	1.5937	24.953	0.00000
Jarque-Bera	723.48	<i>NaN</i>	0.00000

Tabela 36: Teste de Box-Pierce para Resíduos Padronizados do melhor Modelo APARCH - SkSt

Q-Statistics on Standardized Residuals			
$Q(5)$	=	1.36024	[0.9286127]
$Q(10)$	=	5.78167	[0.8332609]
$Q(20)$	=	19.5805	[0.4844340]
$Q(50)$	=	41.2048	[0.8076777]
H_0 : No serial correlation => Accept H_0 when prob. is High			

Tabela 37: Teste de Box-Pierce para Resíduos Padronizados ao Quadrado do melhor Modelo GJR - SkSt

Q-Statistics on Squared Standardized Residuals			
=> P-values adjusted by 1 degree(s) of freedom			
$Q(5)$	=	10.7953	[0.0128857]
$Q(10)$	=	21.5966	[0.0057205] **
$Q(20)$	=	26.5663	[0.0874965]
$Q(50)$	=	54.2555	[0.2482358]
H_0 : No serial correlation => Accept H_0 when prob. is High			

Observe que o parâmetro de assimetria é positivo e significativo, indicando que esta série apresenta assimetria na volatilidade. Observe também que pelos critérios de informações estes modelos são melhores do que os modelos *GJR*.

6 Previsão para Modelos GARCH

É relativamente fácil obter as previsões h-passos à frente para modelos na classe GARCH. Por exemplo para o GARCH(p,q) a pre-

visão ótima h-passos à frente para a variância condicional, $\hat{\sigma}_{t+h|t}^2$ é dada por:

$$\hat{\sigma}_{t+h|t}^2 = \hat{\omega} + \sum_{i=1}^q \hat{\alpha}_i a_{t+h-i|t}^2 + \sum_{j=1}^p \hat{\beta}_j \hat{\sigma}_{t+h|t}^2 \quad (48)$$

onde $a_{t+i|t}^2 = \sigma_{t+i|t}^2$ para $i > 0$, $a_{t+i|t}^2 = u_{t+i}^2$ e $\sigma_{t+i|t}^2 = \sigma_{t+i}^2$ quando $i \leq 0$.

A equação (48) é calculada recursivamente mesmo que seja possível obter uma solução fechada para $\hat{\sigma}_{t+h|t}^2$.

Para os modelos Limiar GARCH e EGARCH as previsões fora da amostra são um pouco mais complicada já que dependem das hipóteses feitas sobre a distribuição das inovações.

Por exemplo para o GJR(p,q) temos que a função de previsão para a variância condicional é dada por:

$$\hat{\sigma}_{t+h|t}^2 = \hat{\omega} + \sum_{i=1}^q (\hat{\alpha}_i a_{t+h-i|t}^2 + \hat{\gamma}_i S_{t-i+h|t}^- a_{t+h-i|t}^2) + \sum_{j=1}^p \hat{\beta}_j \hat{\sigma}_{t+h|t}^2 \quad (49)$$

Observe que $S_{t+i|t}^- = S_{t+i}^-$ para $i \leq 0$ e para $i > 0$, $S_{t+i|t}^-$ depende da escolha da distribuição de ε_t . Por exemplo se a distribuição for simétrica em trono do zero, como as distribuições Normal, t-Student e Distribuição de Valores Extremos, temos que a probabilidade de ε_t ser negativa é $S_{t+i|t}^- = 0.5$. Então nestes casos usamos $S_{t+i|t}^- = 0.5$ na expressão (49).

A função de previsão para o EGARCH é dada por:

$$\begin{aligned} \ln(\hat{\sigma}_{t+h|t}^2) &= E(\ln(\hat{\sigma}_{t+h|t}^2) \mid \mathbf{I}_t) \\ &= E \left(\hat{\omega} + [1 - \hat{\beta}(L)]^{-1} [1 + \hat{\alpha}(L)] \hat{g}(\varepsilon_{t+h-1}) \mid \mathbf{I}_t \right) \\ &= [1 - \hat{\beta}(L)] \hat{\omega} + \hat{\beta}(L) \ln(\hat{\sigma}_{t+h|t}^2) + [1 + \hat{\alpha}(L)] \hat{g}(\varepsilon_{t+h-1|t}) \end{aligned} \quad (50)$$

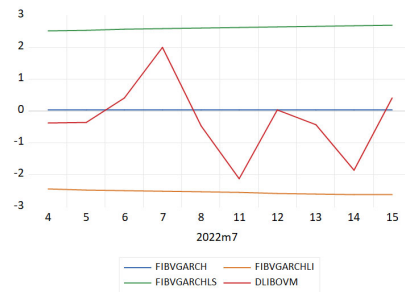
onde $\hat{g}(\varepsilon_{t+k|t}) = \hat{g}(\varepsilon_{t+k})$ para $k \leq 0$ e 0 para $k > 0$.

6.1 Comandos no EVIEWS para fazer Previsões

Nas linhas 172 até 183 do programa Github-GARCH-EVIEWS são apresentados os comandos para fazer previsões dez passos à frente, fora da amostra, para o modelo N-GARCH.

Temos o seguinte gráfico que apresenta estas previsões

Figura 31: Previsão fora da amostra para N-GARCH para o IBOVESPA

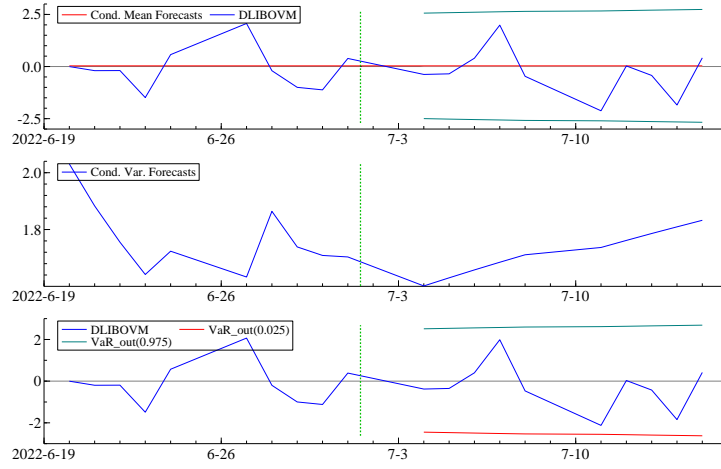


6.2 Previsões no OxMetrics

As previsões para o modelo GARCH no OxMetrics são obtidas através da programa Github-Ox-Garch

Obtém-se os seguintes resultados:

Figura 32: Previsão fora da amostra para N-GARCH usando OxMetrics para o IBOVESPA



7 Modelo Fracionário para Volatilidade

7.1 Introdução

Assim como existe uma distinção entre estacionariade e processos com raíz unitária, podemos ter raíz fracionárias para contempla a memória longa que pode existir na média condicional dos processos.

No caso de modelo estacionário, por exemplo um *ARMA*, temos que a FAC satisfaz a desigualdade:

$$|\rho_k| \leq Cr^k, \quad k = 1, 2, \dots \quad (51)$$

onde $C > 0$ e $r \in (0, 1)$ são constantes.

Logo $\sum_{k=0}^{\infty} |\rho_k| < \infty$ que implica que o processo tem memória curta.

Existe um outro processo estacionário para o qual a FAC decae a zero numa taxa mais lenta, por exemplo, o comportamento assintotica da FAC é dado por:

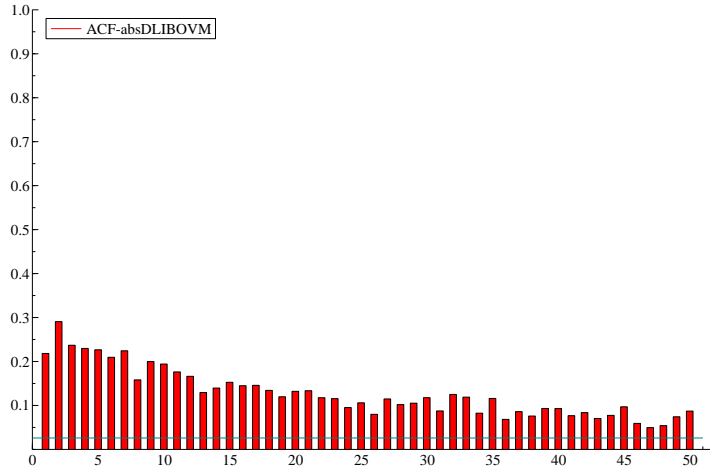
$$\rho_k \sim Ck^{2d-1}, \quad \text{com } k \longrightarrow \infty \quad (52)$$

onde $C \neq 0$ e $d < 0.5$ que é conhecido por processo de memória longa.

A FAC de um processo de memória longa decae para zero a uma taxa mais lenta k^{2d-1} e $\sum_{k=0}^{\infty} |\rho_k| = \infty$ quando $d \in (0, 0.5)$.

O gráfico a seguir apresenta a FAC para o valor absoluto do retorno composto percentual do Índice BOVESPA.

Figura 33: FAC para o Valor Absoluto para o retorno composto percentual do IBOVESPA



Observe que a FAC decae muito lentamente indicando existência de memória longa na série.

Um ruído branco fracionário é definido por:

$$(1 - L)^d y_t = \varepsilon_t \quad (53)$$

onde $\{\varepsilon_t\}$ é um ruído branco e $-0.5 < d < 0.5$.

Temos as seguintes propriedades:

1. Se $d < 0.5$ então y_t é estacionário fraco e tem uma representação $MA(\infty)$ dada por

$$y_t = \varepsilon_t + \sum_{k=1}^{\infty} \psi_k \varepsilon_{t-k}$$

$$\text{com } \psi_k = \frac{d(1+d)\dots(k-1+d)}{k!} = \frac{(k+d-1)!}{k!(d-1)!} \quad (54)$$

2. Se $d > -0.5$ então y_t é invertível e tem uma representação $AR(\infty)$ dada por

$$y_t = \sum_{k=1}^{\infty} \pi_k y_{t-k} + \varepsilon_t$$

$$\text{com } \pi_k = \frac{-d(1-d)\dots(k-1-d)}{k!} = \frac{(k-d-1)!}{k!(-d-1)!} \quad (55)$$

3. Para $-0.5 < d < 0.5$ a FAC de y_t é dada por

$$\rho_k = \frac{d(1+d)\dots(k-1+d)}{(1-d)(2-d)\dots(k-d)}, \quad k = 1, 2, \dots \quad (56)$$

em particular $\rho_1 = d/(1-d)$ e

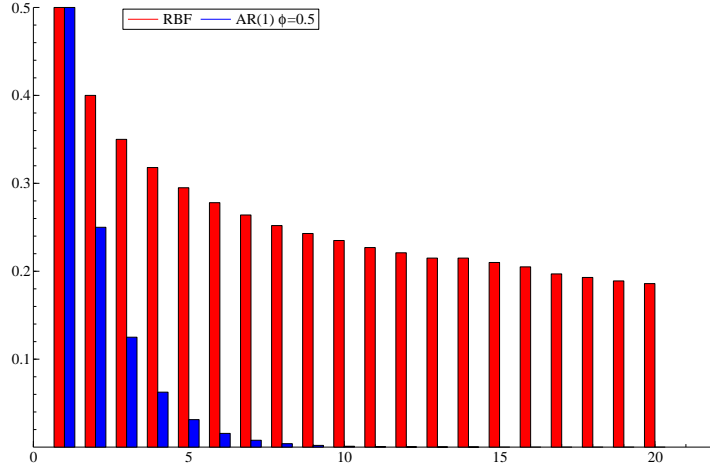
$$\rho_k \approx \frac{(-d)!}{(d-1)!} k^{2d-1} \quad \text{quando } k \longrightarrow \infty \quad (57)$$

4. Para $-0.5 < d < 0.5$ a FACP de y_t é dada por

$$\phi_{k,k} = \frac{d}{k-d} \quad \text{para } k = 1, 2, \dots \quad (58)$$

Abaixo é apresentada a FAC com ruído branco fracionário com $d = 1/3$. Como a primeira autocorrelação é igual a 0.5 devemos comparar esta FAC com a de um modelo $AR(1)$ com $\phi = 0.5$

Figura 34: FAC para Ruído Branco Fracionário com $d = 1/3$ e um $AR(1)$ com $\phi = 0.5$



Observe que o decaimento é bem mais lento do que no $AR(1)$ com $\phi = 0.5$.

Se a série fracionalmente diferenciada $(1-L)^d y_t$ segue um $ARMA(p, q)$ então y_t é chamado de $ARFIMA(p, d, q)$.

Assim como é feito no caso de d interior, temos que verificar se a série tem memória longa.

7.2 Teste R/S para Memório Longa

O primeiro teste de memória longa é o amplitude sobre desvio padrão (*range over standard deviation* R/S) proposto por Hurst [1951], e generalizado por Mandelbrot [1975], que é definido pela amplitude de somas parciais dos desvios de uma série temporal da sua média padronizada pelo desvio padrão que é definida por:

$$Q_T = \frac{1}{s_T} \left[\max_{1 \leq k \leq T} \sum_{t=1}^k (y_t - \bar{y}) - \min_{1 \leq k \leq T} \sum_{t=1}^k (y_t - \bar{y}) \right] \quad (59)$$

onde $\bar{y} = \frac{1}{T} \sum_{t=1}^T y_t$ e $s_T = \sqrt{\frac{1}{T} \sum_{t=1}^T (y_t - \bar{y})^2}$.

Se y_t é i.i.d. e normal temos que (59) converge em distribuição para um movimento Browniano, V , isto é

$$\frac{1}{\sqrt{T}}Q_T \Longrightarrow V \quad (60)$$

Lo mostrou que a estatística R/S não é robusta a curta dependência. Para permitir curta dependência em y_t , Lo propôs a seguinte modificação na estatística R/S

$$Q_T = \frac{1}{\hat{\sigma}_T(q)} \left[\max_{1 \leq k \leq T} \sum_{t=1}^k (y_t - \bar{y}) - \min_{1 \leq k \leq T} \sum_{t=1}^k (y_t - \bar{y}) \right] \quad (61)$$

onde $\hat{\sigma}_T(q)$ é a raiz quadrada do estimador de Newey-West para a variância de longo prazo com tamanho de janela igual a q , isto é

$$\hat{\sigma}_T^2(q) = \hat{\gamma}_0 + 2 \sum_{j=1}^q \omega_{j,T} \cdot \hat{\gamma}_j \quad (62)$$

onde $\omega_{j,T}$ são pesos que somam um, por exemplo, $\omega_{j,T} = 1 - \frac{j}{q+1}$

Aplicando o teste R/S ao retornos composto percentual e ao valor absoluto do retorno composto percentual do Índice BOVESPA usando o programa Ox-GARCH-TesteR/S temos:

Tabela 38: Teste R/S para o Retorno Composto Percentual do IBOVESPA

Hurst-Mandelbrot R/S test statistics: 1.32312	Conf. Int.
Lo R/S test statistics (q=1): 1.3429	
Critical Values	
90% :	[0.861, 1.747]
95% :	[0.809, 1.862]
99% :	[0.721, 2.098]
H0: Hurst-Mandelbrot = no autocorrelation and Lo = no long-term dependence	

Tabela 39: Teste R/S para o Valor Absoluto Retorno Composto Percentual do IBOVESPA

Hurst-Mandelbrot R/S test statistics: 6.38193	Conf. Int.
Lo R/S test statistics (q=1): 5.78184	
Critical Values	
90% :	[0.861 , 1.747]
95% :	[0.809 , 1.862]
99% :	[0.721 , 2.098]
H0: Hurst-Mandelbrot = no autocorrelation and Lo = no long-term dependence	

Observe que pela tabela 38 a hipótese nula é rejeitada pelos dois testes, logo o retorno composto percentual do IBOVESPA não tem memória longa.

Por outro lado, pela tabela 39 a hipótese nula não é rejeitada pelos dois testes, logo o retorno composto percentual do IBOVESPA tem memória longa.

7.3 Teste GPH para Memória Longa

Antes de definir o teste precisamos do conceito de densidade espectral que é uma função suave da autocorrelação amostral e pode ser definida por:

$$f(\omega) = \frac{1}{2\pi} \sum_{j=-(T-1)}^{T-1} K(j) \hat{\rho}_{|j|} \cos(j\omega) \quad \text{para } 0 \leq \omega \leq \pi \quad (63)$$

onde $|\cdot|$ representa o valor absoluto, tal que, $\hat{\rho}_{|-1|} = \hat{\rho}_{|1|}$. E $K(j)$ é uma função chamada de janela de truncamento tal que $K(-j) = K(j)$ e para a janela de Parzen temos a seguinte definição

$$K(j) = \begin{cases} 1 - 6 \left(\frac{j}{m}\right)^2 + 6 \left|\frac{j}{m}\right|^3 & \left|\frac{j}{m}\right| \leq 0.5 \\ 2 \left(1 - \left|\frac{j}{m}\right|\right)^3 & 0.5 \leq \left|\frac{j}{m}\right| \leq 1.0 \\ 0 & \left|\frac{j}{m}\right| > 1.0 \end{cases}$$

Por exemplo se a série for um ruído branco a única contribuição no somatório de (63) é para $j = 0$ e temos que $f(\omega) = \frac{\sigma^2}{2\pi}$ para $0 \leq \omega \leq \pi$, isto é a densidade espectral é constante.

O periodograma é definido por

$$p(\omega) = \frac{\hat{\gamma}_0}{2\pi} \sum_{j=-(T-1)}^{T-1} \hat{\rho}_{|j|} \cos(j\omega)$$

para $\omega = 0, 2\pi/T, 4\pi/T, \dots, ([\frac{T}{2}]2\pi)/T$ (64)

que é uma estimativa da densidade espectral

A densidade espectral para um ruído branco fracionário é dado por:

$$f(\omega) = \left[4 \sin^2 \left(\frac{\omega}{2} \right) \right]^{-d} f_\varepsilon(\omega) \quad (65)$$

onde ω é a frequência, e $f_\varepsilon(\omega)$ a densidade espectral de ε_t .

Observe que a diferença fracionária d pode ser estimada através da seguinte regressão:

$$\ln(f(\omega_j)) = \beta - d \ln \left[4 \sin^2 \left(\frac{\omega_j}{2} \right) \right] + e_j$$

para $j = 1, 2, \dots, T^{1/2}$.

Geweke and Porter-Hudak [1983] mostram que usando o periodograma como estimativa de $f(\omega_j)$, o estimador de M.Q.O. de d na regressão (7.3) tem distribuição normal, isto é,

$$\hat{d} \sim N\left(d, \frac{\pi^2}{T^{1/2} \cdot 6 \sum_{j=1}^{T^{1/2}} (U_j - \bar{U})^2}\right) \quad (66)$$

onde $U_j = \ln \left[4 \sin^2 \left(\frac{\omega_j}{2} \right) \right]$ e \bar{U} é a média amostral de U_j para $j = 1, 2, \dots, T^{1/2}$.

Sob a hipótese nula de não memória longa, isto é, $d = 0$ a estatística $t - Student$ é dada por:

$$t_{d=0} = \hat{d} \left(\frac{\pi^2}{6 \sum_{j=1}^{T^{1/2}} (U_j - \bar{U})^2} \right)^{-1/2}$$

Aplicando o teste GPH ao valor absoluto do retorno composto percentual do Índice BOVESPA usando o programa Ox-GARCH-TestGPH temos:

Tabela 40: Teste de GPH para Valor absoluto do retorno composto percentual do Índice BOVESPA

Log Periodogram Regression			
	Parameter	Stand Error	p-value
d parameter	0.24849	(0.0125097)	[0.0000]
N ⁰ of Observations: 5879;			
n ⁰ of periodogram points: 2939			

Como a estimativa de d foi de $\hat{d} = 0.24849$ que é significativo e está dentro do intervalo $[-0.5 : 0.5]$ esta série é fracamente estacionária e invertível.

7.4 Estimando um Modelo ARFIMA

Assumindo normalidade, isto é, $y_t \sim N(\mu, \Sigma)$ onde μ é a média não condicional da série e Σ é a matriz de variância e covariância para o *ARFIMA*

A função de verossimilhança do modelo *ARFIMA* é dada por

$$\log L(d, \phi, \theta, \sigma_a^2) = -\frac{T}{2} \ln(2\pi) - \frac{1}{2} \ln |\Sigma| - \frac{1}{2} \mathbf{y} \Sigma^{-1} \mathbf{y} \quad (67)$$

onde um algoritmo para calcular a função de autocovariância para um modelo *ARFIMA* pode ser obtida em Sowell [1992].

O programa Ox-GARCH-ARFIMA no OxMetrics estima um modelo ARFIMA(1,d,1), veja tabela 41. Como o coeficiente do

$AR(1)$ não é significativo, re-estimamos um ARFIMA(0,d,1) e os resultados estão na tabela 42.

Tabela 41: Estimação do Modelo ARFIMA(1,d,1)

Modified profile likelihood estimation of ARFIMA(1,d,1) model				
The estimation sample is : 2000-01-04 - 2022-07-15				
The dependent variable is; absDLIBOVM				
The dataset is: C:\Users\Pedro\Dropbox\Topicos_em_Financas_2022\Dados\DADOS_BR.xlsx				
	Coefficient	Std.Error	t-value	t-prob
d parameter	0.390890	0.02711	14.4	0.000
AR-1	−0.0489498	0.05520	−0.887	0.375
MA-1	−0.271438	0.07383	−3.68	0.000
Constant	1.25126	0.3937	3.18	0.001
loglikelihood	−8890.27278			
n ^o of observations	5879	n ^o of parameters:		5
AIC.T	14460.5456	AIC		3.0261
mean(absDLIBOVM)	1.22396	var(absDLIBOVM)		1.44577
σ	1.09754	σ ²		1.20459
Descriptive statistics for residuals:				
Normality test:	χ ² (2)	=	1676.9	[0.0000] **
ARCH 1-1 test:	F(1, 5873)	=	327.26	[0.0000] **
Portmanteau(76):	χ ² (73)	=	160.5	[0.0000] **

Tabela 42: Estimação do Modelo ARFIMA(0,d,1)

Modified profile likelihood estimation of ARFIMA(0,d,1) model				
The estimation sample is : 2000-01-04 - 2022-07-15				
The dependent variable is; absDLIBOVM				
The dataset is: C:\Users\Pedro\Dropbox\Topicos_em_Financas_2022\Dados\DADOS_BR.xlsx				
	Coefficient	Std.Error	t-value	t-prob
d parameter	0.404810	0.02374	17.1	0.000
MA-1	−0.331293	0.02845	−11.6	0.000
Constant	1.25290	0.4518	2.77	0.006
loglikelihood	−8890.68749			
n ^o of observations	5879	n ^o of parameters:		4
AIC.T	17789.395	AIC		3.0259
mean(absDLIBOVM)	1.22396	var(absDLIBOVM)		1.44577
σ	1.09763	σ ²		1.20478
Descriptive statistics for residuals:				
Normality test:	χ ² (2)	=	1671.9	[0.0000] **
ARCH 1-1 test:	F(1, 5873)	=	332.94	[0.0000] **
Portmanteau(76):	χ ² (74)	=	160.51	[0.0000] **

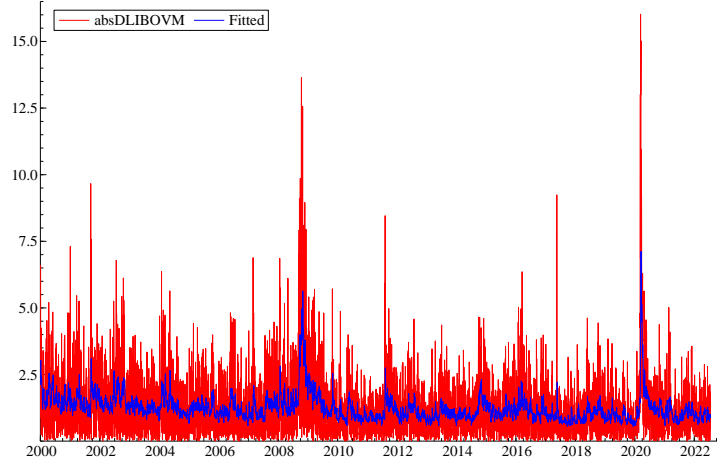
Tabela 43: Comparação entre os modelos ARFIMA(p,d,q)

Progress to date							
Model	T	p		Log-Lik	SC	HQ	AIC
ARFIMA(1,d,1)	5879	5	MPL	-8890.2728	3.0318	3.0281	3.0261
ARFIMA(0,d,1)	5879	4	MPL	-8890.6975	3.0305	3.0275	3.0259

Pelos critérios de informação o melhor modelo é $ARFIMA(0, d, 1)$ fracamente estacionário, uma vez que $\hat{d} = 0.3313$, que percente ao intervalo $[-0.5 : 0.5]$.

Abaixo é apresentado o observado, AbsDLIBOVM, e o ajustado pelo melhor *ARFIMA*. Observe que o modelo de memória longa acompanha os movimentos da série dos valores absolutos do retornos compostos percentuais do Índice BOVESPA.

Figura 35: AbsDLIBOVM X Ajustado melhor ARFIMA(0,d,1)



7.5 Modelo *FIGARCH* para Volatilidade

Como o correlograma dos retornos ao quadrado apresenta este padrão de memória longa, Baillie et al. [1996](BBM) introduziram o *GARCH* Fracionário (*FIGARCH*) onde a primeira diferença em (38) é substituída por $(1 - L)^d$ com d fracionário. A variância condicional do *FIGARCH*(s, d, m) será dado por:

$$\sigma_t^2 = \underbrace{\omega[1 - \beta_1(L)]^{-1}}_{\omega^*} + \underbrace{[1 - \phi(L)(1 - L)^d[1 - \beta_1(L)]^{-1}]a_t^2}_{\lambda(L)} \quad (68)$$

que pode ser re-escrito da seguinte forma

$$\sigma_t^2 = \omega^* + \lambda(L)a_t^2 \quad (69)$$

onde $0 \leq d \leq 1$.

É possível mostrar, veja Bollerslev and Mikkelsen [1996] no pé de página número 7, que $\omega > 0$, $\beta_1 - d \leq \phi_1 \leq \frac{2-d}{3}$ e $d(\phi_1 - \frac{1-d}{2}) \leq \beta_1 \times (\phi_1 - \beta_1 + d)$ são suficiente para garantir que a variância condicional do modelo *FIGARCH* seja positiva.

O programa Ox-GARCH-FIGARCH estima modelos *FIGARCH*(1, d , 1) com distribuições: normal, Student-t, GED e SkSt. Temos a table (44 com as comparações:

Tabela 44: Comparação entre os modelos FIGARCH(1,d,1)

Progress to date							
Model	T	p		Log-Lik	SC	HQ	AIC
N-FIGARCH(1,d,1)	5879	5	BFGS	-10813.968	3.6862	36825	3.6805
St-FIGARCH(1,d,1)	5879	6	BFGS	-10724.487	3.6573	3.6528	3.6504
GED-FIGARCH(1,d,1)	5879	6	BFGS	-10738.362	3.6620	3.6575	3.6552
SkSt-FIGARCH(1,d,1)	5879	7	BFGS	-10715.458	3.6557	3.6505	3.6477

O modelo escolhido é um SkSt-FIGARCH. Mas este modelo tem todos os critérios de informações superiores ao SkSt-GJR.

7.6 Modelo HYGARCH para Volatilidade

Davidson [2004] propõe o modelo HYGARCH que substitui em $\lambda(L)$ em (69) por $1 - [1 - \beta(L)]^{-1}\phi(L)\{1 + \alpha[(1 - L)^d]\}$.

Observe também que o modelo é não estacionário quando $d > 0$. Pois,

$$\begin{aligned}
(1 - L)^d &= \sum_{k=0}^{\infty} \frac{\Gamma(d+1)}{\Gamma(k+1)\Gamma(d-k+1)} L^k \\
&= 1 - dL - \frac{1}{2}d(1-d)L^2 - \frac{1}{6}d(1-d)(2-d)L^3 - \dots \\
&= 1 - \sum_{k=1}^{\infty} c_k(d)L^k
\end{aligned} \tag{70}$$

onde $c_1(d) = d$, $c_2(d) = \frac{1}{2}d(1-d)$, etc. Por construção temos que $\sum_{k=1}^{\infty} c_k(d) = 1$ para qualquer valor de d , e portanto, o modelo

FIGARCH pertence a mesma classe dos modelos *IGARCH*. Observe que no modelo *HYGARCH* os coeficientes de $c_k(d)$ são ponderados por α . Também temos que os modelos *FIGARCH* com $\alpha = 1$ (ou $\log(\alpha) = 0$) e estão encaixados nos modelos *HYGARCH* e são estacionários quando $\alpha < 1$ (ou $\log(\alpha) < 0$).

O programa Ox-GARCH-HYGARCH estimado modelos *HYGARCH*(1, d , 1) com distribuições: normal, Student-t, GED e SkSt. Temos a table (45 com as comparações:

Tabela 45: Comparação entre os modelos HYGARCH(1,d,1)

Progress to date							
Model	T	p		Log-Lik	SC	HQ	AIC
N-HYGARCH(1,d,1)	5879	5	MaxSA	-10810.269	3.6864	36820	3.6796
St-HYGARCH(1,d,1)	5879	6	MaxSA	-10722.127	3.6579	3.6527	3.6500
GED-HYGARCH(1,d,1)	5879	6	MaxSA	-10735.618	3.6625	3.6573	3.6546
SkSt-HYGARCH(1,d,1)	5879	7	MaxSA	-10715.458	3.6563	3.6504	3.6473

Observe que agora usando o critério *SC* o modelo *SkSt* – *FIGARCH* é superior ao modelo *SkSt* – *HYGARCH*. Mas se forem usados os critérios *AIC* e *HQ* o modelo *SkStHYGARCH* é superior. Comparando com o melhor modelo de memória curta, continuamos com o modelo *SkSt* – *GJR* como o melhor.

7.7 Modelo FIEGARCH para Volatilidade

Bollerslev and Mikkelsen [1996] também propõe a extensão do modelos EGARCH da equação 40 para FIEGARCH onde $[1 - \beta(L)] = \phi(L)(1 - L)^d$ com todas as raízes de $\phi(z) = 0$ estão for do círculo unitário.

A especificação do modelo *FIEGARCH*(1, d , 1) é dada por:

$$\log(\sigma_t^2) = \omega + \phi(L)^{-1}(1 - L)^{-d}[1 + \alpha(L)]g(\varepsilon_{t-1}) \quad (71)$$

onde $g(\varepsilon_t)$ está definido na equação 41.

O programa Ox-GARCH-FIEGARCH estimado modelos *FIEGARCH*(1, d , 1) com distribuições: normal, Student-t, GED e SkSt. Temos a table (46 com as comparações:

Tabela 46: Comparação entre os modelos FIEGARCH(1,d,1)

Progress to date							
Model	T	p		Log-Lik	SC	HQ	AIC
N-FIEGARCH(1,d,1)	5879	5	MaxSA	-10810.269	3.6864	36820	3.6796
St-FIEGARCH(1,d,1)	5879	6	MaxSA	-10722.127	3.6579	3.6527	3.6500
GED-FIEGARCH(1,d,1)	5879	6	MaxSA	-10735.618	3.6625	3.6573	3.6546
SkSt-FIEGARCH(1,d,1)	5879	7	MaxSA	-10715.458	3.6563	3.6504	3.6473

7.8 Modelo FIAPARCH para Volatilidade

Tse [1998] propõe o modelo *FIAPARCH* cuja especificação é dada por:

$$\sigma_t^\delta = \omega + \{1 - [1 - \beta(L)]^{-1} \phi(L)(1 - L)^{-d}\}(|\varepsilon_t| - \gamma\varepsilon_t)^\delta \quad (72)$$

O programa Ox-GARCH-FIAPARCH estima modelos *FIAPARCH*(1, d), para as quatro distribuições. Temos a tabela (47) com as comparações:

Tabela 47: Comparação entre os modelos FIAPARCH(1,d,1)

Progress to date							
Model	T	p		Log-Lik	SC	HQ	AIC
N-FIAPARCH(1,d,1)	5879	5	BFGS	-10741.253	3.6644	3.6593	3.6565
St-FIAPARCH(1,d,1)	5879	6	BFGS	-10666.843	3.6406	3.6347	3.6315
GED-FIAPARCH(1,d,1)	5879	6	BFGS	-10681.685	3.6457	3.6397	3.6366
SkSt-FIAPARCH(1,d,1)	5879	7	BFGS	-10654.593	3.6379	3.6312	3.6277

Observe que estes modelos são superiores ao melhores modelos até agora que eram o *APARCH*.

O programa Ox-GARCH-FIAPARCH estima o modelo *FIAPARCH*(1, d) para as quatro distribuições. Mas como a *SkSt - FIAPARCH* é o melhor modelo, sómente este será apresentado na tabela (48):

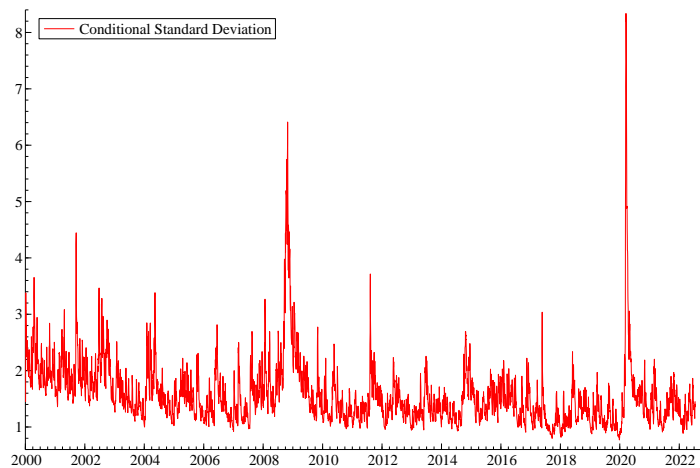
Tabela 48: Estimação do Modelo $SkSt - FIAPARCH(1, d, 1)$

The dependent variable is: DLIBOVM				
Mean Equation: $ARMA(0,0)$ Model				
No regressor in the conditional mean				
Variance Equation: $FIAPARCH(1,1)$ Model				
No regressor in the conditional variance				
Skewed Student distribution with 9.84732 degrees of freedom.				
and Asymetric doeficient ($\log \xi_i$) = -0.0924679 .				
Robust Standard Errors				
	Coefficient	Std.Error	t-value	t-prob
Const(M)	-0.008536	0.018196	-0.4691	0.6390
Const(V)	3.526094	0.50227	7.020	0.0000
d-Figarch	0.301206	0.026289	11.46	0.0000
ARCH(ϕ)	0.129260	0.026289	11.46	0.0000
GARCH(β_1)	0.370616	0.078551	4.718	0.0000
APARCH(γ_1)	0.700696	0.11661	6.009	0.0000
APARCH(δ)	1.263162	0.082222	15.36	0.0000
Asymmetry	-0.092468	0.018409	-5.023	0.0000
Tail	9.847318	1.2657	7.780	0.0000
loglikelihood -10654.593				
n ^o of observations	5879	n ^o of parameters:		9
AIC	3.627689	BIC		3.627685
mean(DLIBOVM)	-0.0000	var(DLIBOVM)		2.94385
Skewness(DLIBOVM)	-0.39646	Kurtosis(DLIBOVM)		10.36097

Os resíduos padronizados e os resíduos padronizados ao quadrado indicam que o modelo está bem especificado.

A seguir é apresentado o Desvio Padrão Condicional para este modelo

Figura 36: Desvio Padrão Condicional do modelo $SkSt - FIAPARCH(1, d, 1)$ para o retorno composto percentual do IBOVESPA



8 Modelos *Dynamic Conditional Score(DCS)* ou *Generalized Autoregressive Score(GAS)* para Volatilidade

8.1 Introdução

Este capítulo é baseado no manual GARCH do OxMetrics Laurent [2018] e também no livro Harvey [2014].

É sabido que as séries financeiras ocasionalmente apresentam grandes mudanças, também conhecidas como saltos(*jumps*).

O impacto dos saltos é, em geral, modelado assumindo uma distribuição de Poisson ou Bernoulli que, quando combinado com uma distribuição normal para a parte do movimento browniano, leva a misturas de Poisson ou Bernoulli de distribuições para retornos financeiros.

Alternativamente, alguns estudos assumem distribuições de cauda pesadas, como o SkSt (Student-t assimétrico) ou a distribuição e erro generalizado, para explicar a ocorrência de grandes mudanças nos retornos.

Vários autores mostraram que esses saltos afetam a volatilidade futura menos do que os modelos de volatilidade padrão preveem.

Muitos modelos de volatilidade, como o GARCH, são baseados na suposição de que cada observação de retorno tem o mesmo impacto relativo na volatilidade futura, independentemente da magnitude do retorno.

Esta suposição está em desacordo com um crescente corpo de evidências indicando que as maiores observações de retorno têm um efeito relativamente menor na volatilidade futura do que choques menores.

Um exemplo é o preço das ações da Bristol-Myers Squibb (BMY). Os retornos diários em % do BMY (no período 1999-2008) estão representados na Figura 37. A Bristol-Myers Squibb, uma das maiores empresas farmacêuticas dos EUA, retirou um pedido de

novo medicamento para um medicamento chamado Omapatrilat em 19 de abril de 2000.

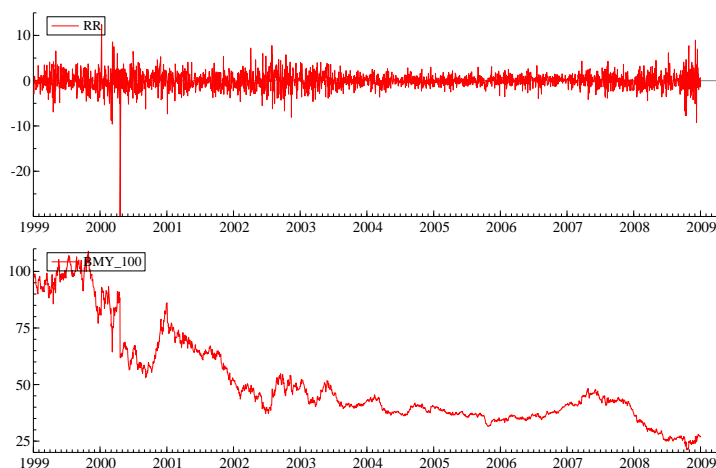
Isso foi geralmente visto como um grande golpe para a empresa, pois era para ser o próximo sucesso de vendas da empresa.

Esperava-se que o produto fosse o mais vendido entre todos os produtos farmacêuticos.

O mercado reagiu fortemente com uma perda de 30% em um único dia.

Este foi um evento único. O mercado se adaptou imediatamente às novas informações e os retornos da ação permaneceram calmos depois.

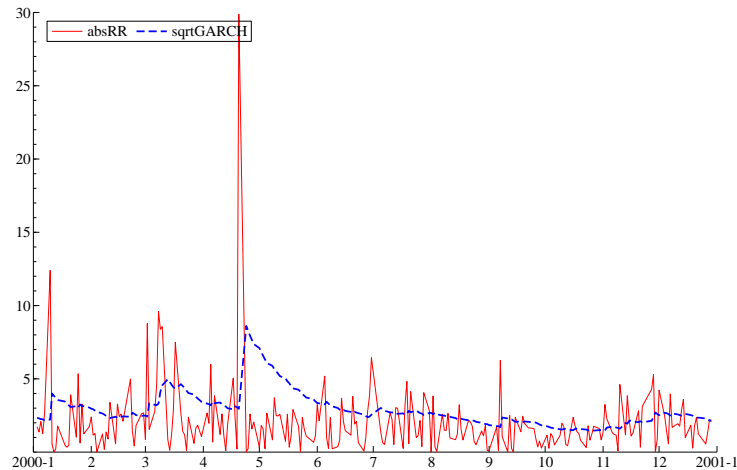
Figura 37: Preços e Retornos diários em % de BMY



A Figura 38 mostra os retornos absolutos (linha contínua) e o desvio padrão condicional do BMY para dois modelos no ano 2000.

A linha pontilhada corresponde a σ_t obtido para um $AR(1)$ – $GARCH(1, 1)$ com distribuição SkSt.

Figura 38: Retornos Absolutos e o desvio padr ao condicional de BMY



Os resultados (ver Tabela 49) sugerem que o termo $AR(1)$ não é significativo e que z_t tem caudas pesadas, mas não é assimétrica.

Tabela 49: Estimação do Modelo $AR(1) + GARCH(1, 1) + SkSt$

The dependent variable is: RR				
Mean Equation: $ARMA(1, 0)$ Model				
No regressor in the conditional mean				
Variance Equation: $GARCH(1, 1)$ Model				
No regressor in the conditional variance				
Skewed Student distribution with 7.17136 degrees of freedom.				
and Asymetric doeficient ($\log \xi_i$) = 0.0164003.				
Robust Standard Errors				
	Coefficient	Std.Error	t-value	t-prob
Const(M)	-0.014941	0.027028	-0.5528	0.5805
AR(1)	-0.008735	0.020674	-0.4225	0.6727
Const(V)	0.021686	0.011029	1.966	0.0494
ARCH(α_1)	0.074118	0.020461	3.622	0.0003
GARCH(β_1)	0.922183	0.021385	43.12	0.0000
Asymmetry	0.016400	0.027821	0.5895	0.5556
Tail	7.17356	1.1157	6.428	0.0000
loglikelihood	-4614.983	$\alpha_1 + \beta_1$		0.99630
n ^o of observations	2489	n ^o of parameters:		7
AIC	3.627689	BIC		3.627685
mean(Y)	-0.0349	var(Y)		3.49157
Skewness(Y)	-1.151021	Kurtosis(Y)		31.71335

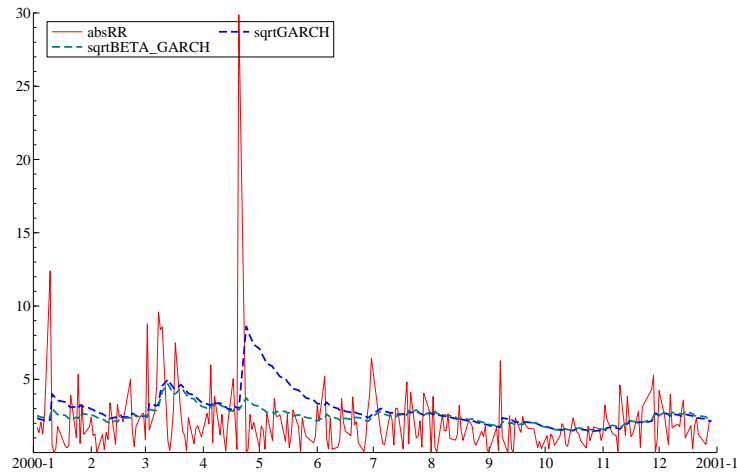
Curiosamente, a Figura 39 sugere que σ_t para o modelo $AR(1) - GARCH(1, 1) - SKST$ tem um aumento acentuado após o evento e permanece em um nível muito alto durante um mês e meio.

Para superar esse problema, Harvey and Chakravarty [2008] ou Creal et al. [2012] propuseram independentemente uma nova maneira de lidar com grandes retornos em um contexto GARCH. Harvey and Chakravarty [2008] chamaram estes modelos de *Dynamic*

Conditional Score - (DCS) e Creal et al. [2012] por *Generalized Autoregressive Score - (GAS)*.

Seus modelos contam com uma distribuição potencialmente não normal para as inovações (z_t) e uma equação tipo *GARCH* para a variância condicional derivado do *score* condicional da distribuição assumida em relação ao segundo momento. Isto pode ser visto pela linha pontilhada que representa o *SD – GAS – SkSt* na figura 39 cuja legenda é *sqrtBETA – GARCH*.

Figura 39: Retornos Absolutos, *SDSkSt – GARCH* e *SD – GAS – SkSt* para BMV



Antes de apresentar a especificação do Modelo GAS temos que mostrara relação do Filtro de Kalman com Extração de sinal de uma série de tempo.

8.2 Filtro de Kalman e Extração de Sinal

As recursões do Kalman Filter sem os componentes determinísticos, c_t e d_t , pode ser escrito como (veja Durbin and Koopman [2012] página 85):

$$\begin{aligned}
\nu_t &= y_t - Z_t a_t & F_t &= Z_t P_t Z_t' + G_t G_t' \\
a_{t|t} &= a_t + P_t Z_t' (F_t)^{-1} \nu_t & P_{t|t} &= P_t - P_t Z_t' (F_t)^{-1} Z_t P_t \\
a_{t+1} &= T_t a_t + K_t \nu_t & P_{t+1} &= T_t P_t (T_t - K_t Z_t)' + H_t H_t'
\end{aligned} \tag{73}$$

onde $K_t = T_t P_t Z_t' (F_t)^{-1}$ é o chamado **Ganho do Kalman Filter**.

Agora a_{t+1} e P_{t+1} são as previsões do vetor de estado α_{t+1} e sua variância no instante t .

Defina o **erro de estimativa do vetor de estado** por:

$$x_t = \alpha_t - a_t \tag{74}$$

com

$$Var(x_t) = P_t \tag{75}$$

As recursões do Filtro Kalman dadas em (73) podem ser reescritas usando o erro de estimativa do vetor de estado e sua relação com o erro de previsão um passo à frente, $\nu_t = y_t - E(y_t | Y_{t-1}) = y_t - Z_t a_t$

Como ν_t é a parte de y_t que não pode ser prevista pelo passado pode ser chamado de **inovações**.

O Filtro de Kalman em **forma de inovação** é dado por:

$$\begin{aligned}
\nu_t &= y_t - Z_t a_t \\
&= Z_t \alpha_t - Z_t a_t + G_t \varepsilon_t \\
&= Z_t x_t + G_t \varepsilon_t
\end{aligned} \tag{76}$$

e

$$\begin{aligned}
x_{t+1} &= \alpha_{t+1} - a_{t+1} \\
&= T_t \alpha_t + H_t \eta_t - T_t a_t - K_t \nu_t \\
&= T_t x_t + H_t \eta_t - K_t Z_t x_t - K_t G_t \varepsilon_t \\
&= L_t x_t + H_t \eta_t - K_t G_t \varepsilon_t
\end{aligned} \tag{77}$$

onde K_t é o ganho do Filtro Kalman e $L_t = T_t - K_t Z_t$.

A inovação análoga do modelo em representação em espaço de estado é dada por:

$$\nu_t = Z_t x_t + G_t \varepsilon_t \quad x_{t+1} = L_t x_t + H_t \eta_t - K_t G_t \varepsilon_t \tag{78}$$

com valor inicial: $x_1 = \alpha_1 - a_1$.

A recursão para P_{t+1} é dada por:

$$\begin{aligned}
P_{t+1} &= Var(x_{t+1}) = E[(\alpha_{t+1} - a_{t+1})x'_{t+1}] \text{ notag} \\
&= E(\alpha_{t+1}x'_{t+1}) \\
&= E(T_t \alpha_t + H_t \eta_t)(L_t x_t + H_t \eta_t - K_t G_t \varepsilon_t)' \\
&= T_t P_t L_t' + H_t H_t'
\end{aligned} \tag{79}$$

$$\tag{80}$$

com $Cov(x_t \eta_t) = 0$.

Os erros de previsão um passo à frente são independentes uns dos outros e a distribuição conjunta das observações usando a decomposição do erro de previsão pode ser escrita como:

$$p(\nu_1, \dots, \nu_T) = \prod_{t=1}^T p(\nu_t) \tag{81}$$

O modelo AR(1)+ruído ou o modelo de extração de sinal gaussiano mais ruído pode ser escrito como:

$$\begin{aligned}
y_t &= \mu_t + \varepsilon_t \\
\mu_{t+1} &= \phi \mu_t + \eta_t
\end{aligned} \tag{82}$$

com $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$ e $\eta_t \sim N(0, \sigma_\eta^2)$ e são mutuamente não correlacionados.

A **relação sinal-ruído** é dada por:

$$q = \frac{\sigma_\eta^2}{\sigma_\varepsilon^2} \quad (83)$$

e desempenha um papel fundamental na determinação de como as observações são ponderadas para previsão e extração de sinal.

A forma reduzida de 82 é dada por:

$$y_t = \phi y_{t-1} + \xi_t - \theta \xi_{t-1} \quad (84)$$

onde $\xi_t \sim N(0, \sigma^2)$ com restrições em θ .

Modelo (82) pode ser estimado pelo Filtro de Kalman e os parâmetros ϕ e q são estimados pelo Máxima Verossimilhança.

O Filtro de Kalman pode ser expresso como uma única equação que combina o estimador de μ_t com base na informação até o instante t com a t -ésima observação para produzir a ***innovation form (IF)*** do Filtro Kalman:

$$\begin{aligned} y_t &= \mu_{t|t-1} + v_t \\ \mu_{t+1|t} &= \phi \mu_{t|t-1} + K_t v_t \end{aligned} \quad (85)$$

e $K_t = \frac{\phi^2 P_{t+q}}{1 + \phi^2 P_{t+q}}$.

Agora a distribuição da t -ésima observação condicionada ao passado e os parâmetros variantes no tempo são atualizados por um filtro definido adequadamente.

Este modelo foi chamado de ***observation-driven*** por Cox [1981] e o modelo Gaussiano Não Observado foi chamado por Cox [1981] de ***parameter-driven model***.

O Filtro de Kalman dado em (85) é orientado (*driven*) pelo erro de previsão um passo à frente.

Se as distribuições não forem gaussianas v_t será substituído por uma variável u_t , que é proporcional ao ***score of the conditional distribution*** e a segunda equação de (85) torna-se:

$$\mu_{t+1|t} = \phi\mu_{t|t-1} + K_t u_t \quad (86)$$

onde K_t é tratado como um parâmetro desconhecido.

8.3 Especificação dos Modelos *Dynamical Conditional Score (DCS)* e *Generalized Autoregressive Score (GAS)*

Um modelo *DCS* é especificado em termos de uma distribuição condicional dada por:

$$\begin{aligned} f(y_t \mid \theta_{t|t-1}, Y_{t-1}) \quad \text{para } t = 1, \dots, T \\ \theta_{t+1|t} = g(\theta_{t|t-1}, \theta_{t-1|t-2}, \dots, Y_t) \end{aligned} \quad (87)$$

onde $Y_t = (y_1, \dots, y_t)$ e $\theta_{t|t-1}$ é um parâmetro que varia ao longo do tempo.

A segunda equação de (87) representa o Processo Gerador dos Dados ou uma forma de escrever o filtro que aproxima o modelo de Componentes não Observáveis não linear.

A especificação Normal-GARCH(1,1) é um ***observation-driven model***, uma vez que $\theta_{t|t-1} = \sigma_{t|t-1}^2$.

Este modelo pode ser escrito da seguinte forma:

$$y_t \mid Y_{t-1} \sim NI(0, \sigma_{t|t-1}^2) \quad (88)$$

$$\sigma_{t+1|t}^2 = \omega + \varphi\sigma_{t|t-1}^2 + \alpha v_t \quad \text{with } \omega > 0, \varphi \geq \alpha, \alpha \geq 0 \quad (89)$$

com

$$\varphi = \alpha + \beta \quad (90)$$

e

$$v_t = y_t^2 - \sigma_{t|t-1}^2$$

que é uma diferença martingal.

Quando a distribuição não é gaussiana na equação (89) usamos u_t em vez de v_t e obtemos o modelo $DCS - N - GARCH(1, 1)$.

O *score* para o modelo (88) é denotado por ∇_t e pode ser expresso da seguinte forma:

$$\begin{aligned} \nabla_t &= \partial \left(-\frac{1}{2}(\log(\sigma_{t|t-1}^2) + \frac{\varepsilon_t^2}{\sigma_{t|t-1}^2}) \right) / \partial \sigma_{t|t-1}^2 \\ &= -\frac{1}{2} \left(\frac{1}{\sigma_{t|t-1}^2} - \frac{y_t^2}{\sigma_{t|t-1}^4} \right) \\ &= \frac{1}{2} \left(\frac{y_t^2}{\sigma_{t|t-1}^4} - 1 \right) \sigma_{t|t-1}^2 \\ 2\nabla_t &= (z_t^2 - 1) \sigma_{t|t-1}^2 \end{aligned} \tag{91}$$

onde $z_t \sim N(0, 1)$.

Substituindo (91) em (89) temos:

$$\begin{aligned} \sigma_{t+1|t}^2 &= \omega + \varphi \sigma_{t|t-1}^2 + \alpha (z_t^2 - 1) \sigma_{t|t-1}^2 \\ &= \omega + \beta \sigma_{t|t-1}^2 + \underbrace{\alpha z_t^2 \sigma_{t|t-1}^2}_{y_{t-1}^2} \\ \sigma_{t+1|t}^2 &= \omega + \beta \sigma_{t|t-1}^2 + \alpha y_{t-1}^2 \end{aligned} \tag{92}$$

que é o D.G.P. para o $N - GARCH(1, 1)$.

Em geral, para um modelo $GAS(1, 1)$, a equação de atualização para os parâmetros que variam no tempo, ψ_t , é a conhecida função de atualização autorregressiva:

$$\theta_{t+1/t} = \omega + B_1\theta_{t/t-1} + A_1\kappa_t \quad (93)$$

Harvey and Chakravarty [2008] e Creal et al. [2012] propõem atualizar os parâmetros que variam no tempo através de:

$$\kappa_t = S_t \nabla_t \quad (94)$$

onde ∇_t é o *score* em relação ao parâmetro $\theta_{t/t-1}$, i.e.

$$\nabla_t = \partial(f(y_t \mid \theta_{t/t-1}, Y_{t-1})/\partial\theta_{t/t-1} \quad (95)$$

e S_t é uma matriz de escala dependente do tempo, para Creal et al. [2012] é recomendado usar $S_t = 1$ ou $S_t = [E_{t-1}(\nabla_t \nabla_t)]^{-1}$, enquanto que Harvey and Chakravarty [2008] usa $S_t = 2$.

O modelo $GAS(1, 1)$ de Harvey and Chakravarty [2008] combinado com distribuições: Normal, Student-t, GED ou SkSt é apresentado abaixo:

$$\sigma_{t+1|t}^2 = \omega + \varphi\sigma_{t|t-1}^2 + \alpha u_t \sigma_{t|t-1}^2 \quad (96)$$

onde

$$u_t = (z_t^2 - 1) \text{ se } z_t \sim N(0, 1) \quad (97)$$

$$u_t = \frac{(\nu + 1)z_t^2}{\nu - 2 + z_t^2} \text{ se } z_t \sim t(0, 1, \nu) \quad (98)$$

$$u_t = \frac{0.5\nu |z_t|^\nu}{\lambda_\nu^\nu} - 1 \text{ se } z_t \sim GED(0, 1, \nu) \quad (99)$$

$$u_t = \frac{(\nu + 1) z_t z_t^*}{(\nu - 2)g_t \xi^{I_t}} - 1 \text{ se } z_t \sim SkSt(0, 1, \xi, \nu) \quad (100)$$

onde para GED

$$\lambda_\nu = \sqrt{\frac{\Gamma(1/\nu) 2^{(-2/\nu)}}{\Gamma(3/\nu)}} \quad (101)$$

e para SkSt

$$z_t^* = sz_t + m \quad (102)$$

$$I_t = \text{sgn}(z_t^*) = I(z_t^* \geq 0) - I(z_t^* < 0) \quad (103)$$

$$g_t = 1 + \frac{z_t^{*2}}{(\nu - 2)\xi^{2I_t}} \quad (104)$$

$$m = \frac{\Gamma(\frac{\nu-1}{2})\sqrt{\nu-2}}{\sqrt{\pi}\Gamma(\frac{\nu}{2})} \left(\xi - \frac{1}{\xi} \right) \quad (105)$$

$$s = \sqrt{\left(\xi^2 + \frac{1}{\xi^2} - 1 \right) - m^2} \quad (106)$$

Harvey and Chakravarty [2008] chama o modelo *GAS* acima com distribuição Student-t de **Beta-t-GARCH(1,1)** porque $\frac{u_t+1}{\nu+1}$ tem uma distribuição Beta.

Harvey and Chakravarty [2008] também considera o GAS Exponencial (EGAS) que é dado por:

$$\log(\sigma_{t+1|t}^2) = \omega + \varphi \log(\sigma_{t|t-1}^2) + \alpha u_t \quad (107)$$

Também introduzem o efeito de alavancagem no modelo GAS usando o sinal das observações. O modelo (107) é reescrito da seguinte forma:

$$\log(\sigma_{t+1|t}^2) = \omega + \varphi \log(\sigma_{t|t-1}^2) + \alpha u_t + \gamma I_t \quad (108)$$

onde $I_t = \text{sgn}(-z_t)(u_t + 1)$ para distribuição simétrica e $I_t = \text{sgn}(-z_t^*)(u_t + 1)$ para *SkSt*. Este modelo é chamado de **Asymmetric Exponential GAS - AEGAS**.

8.4 Comparação do modelos GARCH com os GAS para o índice BOVESPA

Nesta sessão iremos comparar os modelos da classe GARCH com os da classe GAS.

A tabela abaixo apresenta os melhores modelos na classe GARCH.

Tabela 50: Comparação dos Modelos APARCH(1,1) para as quatro distribuições

Progress to date							
Model	T	p		log-likelihood	SC	HQ	AIC
N-APARCH(1,1)	5879	6	BFGS	-10767.014	3.6717	3.6673	3.6649
t-APARCH(1,1)	5879	6	BFGS	-10704.370	3.6519	3.6467	3.6439
GED-APARCH(1,1)	5879	6	BFGS	-10704.370	3.6519	3.6467	3.6439
SkSt-APARCH(1,1)	5879	7	BFGS	-10681.404	3.6456	3.6396	3.6365

E usando o programa Ox-GAS-EGAS-AEGAS, podemos comparar os modelos *GAS*.

Tabela 51: Comparação entre os modelos GAS(1,1)

Progress to date							
Model	T	p		Log-Lik	SC	HQ	AIC
N-GAS(1,1)	5879	4	MaxSA	-10821.229	3.6872	3.6843	3.6827
St-GAS(1,1)	5879	5	MaxSA	-10742.589	3.6619	3.6582	3.6563
GED-GAS(1,1)	5879	5	MaxSA	-10742.452	3.6643	3.66606	3.6586
SkSt-GAS(1,1)	5879	6	MaxSA	-10729.928	3.6591	3.6547	3.6523

Tabela 52: Comparação entre os modelos EGAS(1,1)

Progress to date							
Model	T	p		Log-Lik	SC	HQ	AIC
N-EGAS(1,1)	5879	4	BFGS	-108837.853	3.6929	3.6899	3.6883
St-EGAS(1,1)	5879	5	BFGS	-10739.427	3.6609	3.6572	3.6552
GED-EGAS(1,1)	5879	5	BFGS	-10749.624	3.6643	3.66606	3.6587
SkSt-EGAS(1,1)	5879	6	BFGS	-10726.755	3.6580	3.6536	3.6512

Tabela 53: Comparação entre os modelos AEGAS(1,1)

Progress to date							
Model	T	p		Log-Lik	SC	HQ	AIC
N-AEGAS(1,1)	5879	5	BFGS	-108812.076	3.6856	3.6819	3.6799
St-AEGAS(1,1)	5879	6	BFGS	-10698.182	3.6483	3.6439	3.6415
GED-AEGAS(1,1)	5879	6	BFGS	-10716.191	3.6544	3.6500	3.6476
SkSt-AEGAS(1,1)	5879	7	BFGS	-10685.584	3.6455	3.6403	3.6376

Na classe dos modelos *GAS*, o modelo *AEGA* é superior aos outros dois modelos, a saber, *GAS* e *EGAS*, para todas as quatro distribuições.

Ao se comprar o melhor modelo da classe *GAS*, que são os *AE GAS*, com os melhores da classe *GARCH*, que são os *APARCH*, só para o critério de SC, que o modelo *AE GAS* é superior, mas para o HQ e AIC, o melhor modelo é o *APARCH*.

Para todos os melhores modelos a distribuição que foi mais adequada sempre foi a *SkSt*.

9 Comparação do Modelo de Volatilidade Estocástica e do GARCH(1,1)

9.1 Introdução

A idéia básica do modelo de volatilidade estocástica é tratar a volatilidade como um componente não-observado, com seu logaritmo sendo diretamente modelado como um processo auto-regressivo. A estrutura dessa classe de modelos é uma discretização dos processos estocásticos em tempo contínuo, a partir dos quais são construídas generalizações do modelo de Black and Scholes [1973], como em Hull and White [1987] e Taylor [1994].

9.2 Especificação do Modelo

O modelo volatilidade estocástico (VE) univariado foi proposto por Taylor [1980] e tem a seguinte estrutura (na forma discreta):

$$r_t = \exp\left(\frac{h_t}{2}\right)a_t \text{ com } a_t \sim NI(0, 1) \quad (109)$$

com a variância seguindo um processo *AR*(1), isto é:

$$h_t = \gamma_0 + \gamma_1 h_{t-1} + \eta_t \text{ com } \eta_t \sim NI(0, \sigma_\eta^2) \quad (110)$$

O que implica que a distribuição condicional dos retornos é normal, isto é, $r_t|h_t \sim N(0, \exp(h_t))$.

Assim h_t , o log-volatilidade, é um componente não-observado. Se $|\gamma_1| < 1$, então é estritamente estacionário, com média $\frac{\gamma_0}{1-\gamma_1}$ e variância $\frac{\sigma_\eta^2}{1-\gamma_1^2}$. Como r_t é o produto de dois processos estritamente estacionários, segue-se que r_t também é estritamente estacionário. Além disso, temos que apresenta excesso de curtoses em relação à distribuição normal, capturando o fato estilizado presente em séries financeiras mencionado anteriormente.

Esse modelo também pode ser reescrito da seguinte forma:

$$\ln(r_t^2) = h_t + \ln(a_t^2) \quad (111)$$

com a evolução da volatilidade dada por (110).

Essa versão nos logs é uma representação em espaço de estado linear para o modelo, onde a primeira equação seria a equação de medida, e a segunda, a equação de transição.

A classe de modelos de volatilidade estocástica (VE) tem sua origem tanto em finanças matemáticas quanto em econometria de finanças. Por exemplo, Clark [1973] sugeriu modelar retornos de ativos como função de um processo aleatório da chegada de informação ao mercado. Essa abordagem, chamada de deformação temporal, implica um modelo para o retorno de ativos com volatilidade que varia ao longo do tempo. Tauchen and Pitts [1983] propõem uma mistura de distribuições com dependência temporal na chegada de informação. Hull and White [1987] expõem um modelo de precificação de opções européias, assumindo que o ativo subjacente segue um modelo de volatilidade estocástica a tempo contínuo. Uma outra abordagem surgiu do trabalho de Taylor [1986], que especifica um modelo de volatilidade estocástica a tempo discreto como uma alternativa ao modelo autorregressivo com heteroscedasticidade condicional (ARCH). Até há pouco tempo, a estimação do modelo de Taylor, ou de qualquer outro modelo da classe VE, era bastante difícil, mas uma técnica

moderna de econometria permitiu que a estimação dessa classe de modelos se tornasse factível. Desse modo, essa classe tornou-se uma alternativa aos modelos ARCH.

Os modelos VE são capazes de reproduzir alguns dos fatos estilizados presentes em séries financeiras. Considere-se o modelo VE dado por (111) e (110), devido à estrutura do modelo, pode-se mostrar que a distribuição de gerada por esse modelo tem caudas pesadas; além disso, devido à estrutura auto-regressiva do logaritmo da volatilidade, ela gera agrupamentos de volatilidade. Se os choques que influenciam os retornos e o logaritmo da volatilidade, são negativamente correlacionados, temos o efeito alavancagem. As evidências empíricas na estimação desse modelo são de que γ_1 é muito próximo de 1, implicando alta persistência ou mesmo memória longa.

A maior dificuldade com os modelo VE é que não se podem obter explicitamente expressões para as funções de verossimilhança, como no caso de outros modelos de volatilidade condicional, tais como os da família GARCH. Existe uma série de métodos de estimação para contornar esse problema, dentre os quais: o método generalizado de momentos (GMM), o método de quase máxima verossimilhança (QMV) e o método das cadeias de Markov de Monte Carlo (MCMC).

Como $\ln(a_t^2)$ tem distribuição logarítmo de uma qui-quadrada com um grau de liberdade e se aproximarmos por uma distribuição normal podemos estimar o modelo (111) e (110) pelo Filtro de Kalman.

9.2.1 Programa no EVIEWS para estimar um modelo VE

Nas linhas 188 até 219 do programa GARCH-EVIEWS estima um modelo de VE e faz a comparação com a volatilidade estimada por um GARCH(1,1)

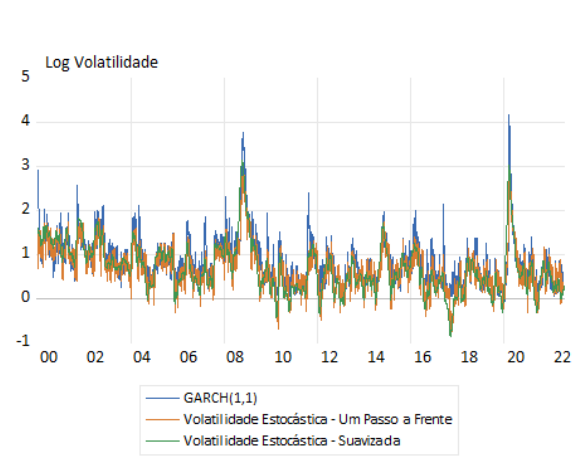
Tem-se os seguinte resultados para as estimativas dos parâmetros:

Tabela 54: Modelo de Volatilidade Estocástica para o Índice BOVESPA.

Sspace: VOLATILIDADEESTOCASTICA				
Method: Maximum likelihood (BFGS / Marquardt steps)				
Date: 08/25/22 Time: 10:20				
Sample: 1/03/2000 8/08/2022				
Included observations: 5896				
Valid observations: 5879				
Convergence achieved after 21 iterations				
Coefficient covariance computed using outer product of gradients				
	Coefficient	Std. Error	z-Statistic	Prob.
C(1)	0.014205	0.004168	3.407949	0.0007
C(2)	0.979232	0.004913	199.3231	0.0000
C(3)	-3.933308	0.241184	-16.30833	0.0000
	Final State	Root MSE	z-Statistic	Prob.
HT	0.318331	0.593054	0.536765	0.5914
Log likelihood	-13264.43	Akaike info criterion		4.513499
Parameters	3	Schwarz criterion		4.516907
Diffuse priors	0	Hannan-Quinn criter.		4.514684

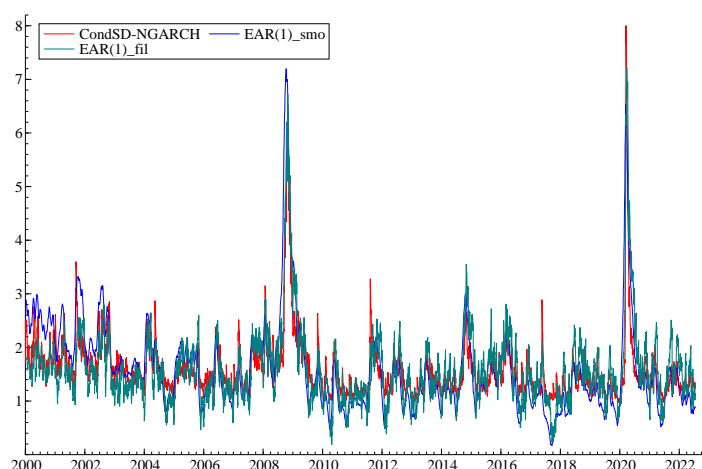
E o gráfico abaixo compara a volatilidade estimada pelo modelo VE, filtrada e suavizada, com o volatilidade estimada por um GARCH(1,1)

Figura 40: Volatilidade Estocástica e N-GARCH(1,1) para IBOVESPA



O programa GARCH-Ox-Estimation-VE estima um modelo de VE e faz a comparação com a volatilidade estimada por um GARCH(1,1)

Figura 41: Volatilidade Estocástica e N-GARCH(1,1) para IBOVESPA estimado no Ox-Metrics



Referências

- R. T. Baillie, T. Bollerslev, and H. O. Mikkelsen. Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 74(1):3–30, sep 1996. doi: 10.1016/s0304-4076(95)01749-6.
- F. Black. The pricing of commodity contracts. *Journal of Financial Economics*, 3(1-2):167–179, jan 1976. doi: 10.1016/0304-405x(76)90024-6.
- F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654, may 1973. doi: 10.1086/260062.
- T. Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307–327, apr 1986. doi: 10.1016/0304-4076(86)90063-1.
- T. Bollerslev and H. O. Mikkelsen. Modeling and pricing long memory in stock market volatility. *Journal of Econometrics*, 73(1):151–184, jul 1996. doi: 10.1016/0304-4076(95)01736-4.
- P. K. Clark. A subordinated stochastic process model with finite variance for speculative prices. *Econometrica*, 41(1):135–155, jan 1973. doi: 10.2307/1913889.
- D. Cox. Statistical analysis of time series: Some recent developments [with discussion and reply]. *Scandinavian Journal of Statistics*, 8(2):93–115, 1981. ISSN 03036898, 14679469. URL <http://www.jstor.org/stable/4615819>.
- D. Creal, S. J. Koopman, and A. Lucas. GENERALIZED AUTOREGRESSIVE SCORE MODELS WITH APPLICATIONS. *Journal of Applied Econometrics*, 28(5):777–795, jan 2012. doi: 10.1002/jae.1279.
- J. Davidson. Moment and memory properties of linear conditional heteroscedasticity models, and a new model. 22(1):16–29, 2004. doi: 10.1198/073500103288619359.
- Z. Ding, C. W. Granger, and R. F. Engle. A long memory property of stock market returns and a new model. 1(1):83–106, 1993. doi: 10.1016/0927-5398(93)90006-d.

- J. Durbin and S. J. Koopman. *Time Series Analysis by State Space Methods*. Oxford University Press, may 2012. doi: 10.1093/acprof:oso/9780199641178.001.0001.
- R. Engle and A. Patton. What good is a volatility model? *Quantitative Finance*, 1(2):237–245, 2001. doi: 10.1088/1469-7688/1/2/305.
- R. F. Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom. *Econometrica*, 50:987–1007, 1982.
- E. F. Fama. Mandelbrot and the stable paretian hypothesis. *The Journal of Business*, 36(4): 420–429, jan 1963. doi: 10.1086/294633.
- E. F. Fama. The behavior of stock-market prices. *The Journal of Business*, 38(1):34–105, 1965.
- J. Geweke. Comment. *Econometric Reviews*, 5(1):57–61, 1986. doi: 10.1080/07474938608800097.
- J. Geweke and S. Porter-Hudak. The estimation and application of long memory time series models. *Journal of Time Series Analysis*, 4(4):221–238, 1983. doi: 10.1111/j.1467-9892.1983.tb00371.x.
- L. R. Glosten, R. Jagannathan, and D. E. Runkle. On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance*, 48(5): 1779–1801, dec 1993. doi: 10.1111/j.1540-6261.1993.tb05128.x.
- A. Harvey and T. Chakravarty. Beta-t(e)garch. Cambridge working papers in economics, Faculty of Economics, University of Cambridge, 2008. URL <https://EconPapers.repec.org/RePEc:cam:camdae:0840>.
- A. C. Harvey. *Dynamic Models for Volatility and Heavy Tails*. Cambridge University Press, 2014. ISBN 1107034728. URL https://www.ebook.de/de/product/20340707/a_c_harvey_andrew_c_harvey_dynamic_models_for_volatility_and_heavy_tails.html.
- M. L. Higgins and A. K. Bera. A class of nonlinear arch models. *International Economic Review*, 33(1):137–158, 1992. doi: 10.2307/2526988.
- J. Hull and A. White. The pricing of options on assets with stochastic volatilities. *The Journal of Finance*, 42(2):281–300, jun 1987. doi: 10.1111/j.1540-6261.1987.tb02568.x.
- H. E. Hurst. Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, 116(1):770–799, 1951. doi: 10.1061/TACEAT.0006518. URL <https://ascelibrary.org/doi/abs/10.1061/TACEAT.0006518>.
- S. Laurent. *Estimating and forecasting ARCH models using G@rch 8*. Timberlake Consultants, London Union, NJ, 2018. ISBN 9780957170896.
- G. M. Ljung and G. E. P. Box. On a measure of lack of fit in time series models. *Biometrika*, 65(2):297–303, aug 1978. doi: 10.1093/biomet/65.2.297.
- A. W. Lo. Long-term memory in stock market prices. *Econometrica*, 59(5):1279. doi: 10.2307/2938368.
- B. Mandelbrot. The variation of certain speculative prices. *The Journal of Business*, 36(4):394, jan 1963. doi: 10.1086/294632.
- B. B. Mandelbrot. Limit theorems on the self-normalized range for weakly and strongly dependent processes. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 31(4): 271–285, 1975. doi: 10.1007/bf00532867.

- D. B. Nelson. Stationarity and persistence in the garch(1,1) model. *Econometric Theory*, 6(3): 318–334, 1990.
- D. B. Nelson. Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, 59(2):347, mar 1991. doi: 10.2307/2938260.
- F. Sowell. Maximum likelihood estimation of stationary univariate fractionally integrated time series models. *Journal of Econometrics*, 53(1-3):165–188, 1992. doi: 10.1016/0304-4076(92)90084-5.
- G. E. Tauchen and M. Pitts. The price variability-volume relationship on speculative markets. *Econometrica*, 51(2):485, 1983. doi: 10.2307/1912002.
- S. J. Taylor. Conjectured models for trends in financial prices, tests and forecasts. *Journal of the Royal Statistical Society. Series A (General)*, 143(3):338, 1980. doi: 10.2307/2982133.
- S. J. Taylor. *Modelling Financial Time Serie*. Wiley, Chichester, 1986.
- S. J. Taylor. Modelling stochastic volatility: a review and comparative study. *Mathematical Finance*, 4(2):183–204, apr 1994. doi: 10.1111/j.1467-9965.1994.tb00057.x.
- Y. K. Tse. The conditional heteroscedasticity of the yen-dollar exchange rate. *Journal of Applied Econometrics*, 13(1):49–55, 1998. doi: 10.1002/(sici)1099-1255(199801/02)13:1<49::aid-jae459>3.0.co;2-o.
- P. L. Valls Pereira, L. K. Hotta, L. A. R. d. Souza, and N. M. C. G. d. Almeida. Alternative models to extract asset volatility: a comparative study. *Brazilian Review of Econometrics*, 19(1):57–109, may 1999. doi: 10.12660/bre.v19n11999.2793.
- J.-M. Zakoian. Threshold heteroskedastic models. *Journal of Economic Dynamics and Control*, 18(5):931–955, 1994. doi: 10.1016/0165-1889(94)90039-6.