



Faculdade de Ciências e Tecnologia da Universidade de Coimbra

Departamento de Matemática

Ano Letivo 2022/2023

# Modelação e Previsão de uma série temporal

Trabalho realizado por (Grupo 5)

Ana Carina Carvalho

Inês Queirós

Margarida Biscaia

Maria Lopes

Séries Temporais

# Conteúdo

<b>1 PARTE I. Geração de uma série temporal</b>	<b>2</b>
1.1 Exercício 1 . . . . .	2
1.2 Exercício 2 . . . . .	3
1.3 Exercício 3 . . . . .	5
1.4 Análise de Sensibilidade . . . . .	6
1.4.1 Alteração do parâmetro $\theta$ . . . . .	6
1.4.2 Alteração dos parâmetros $\alpha$ e $\beta$ . . . . .	6
1.4.3 Alteração da lei marginal de $Z$ . . . . .	7
<b>2 PARTE II. Modelação estocástica de uma série temporal</b>	<b>8</b>
2.1 Exercício 1 . . . . .	9
2.1.1 Alínea a . . . . .	9
2.1.2 Alínea b . . . . .	9
2.2 Exercício 2 . . . . .	10
2.3 Exercício 3 . . . . .	13
2.4 Exercício 4 . . . . .	16
<b>3 PARTE III. Modelação e previsão da cotação diária de uma ação do Índice PSI 20</b>	<b>18</b>
<b>Anexo</b>	<b>24</b>
A Resultados Importantes . . . . .	24
B Códigos Eviews . . . . .	24
B.1 Parte I . . . . .	24
B.2 Parte II . . . . .	26
B.3 Parte III . . . . .	27

# 1 PARTE I. Geração de uma série temporal

Prentende-se gerar um série temporal  $X = (X_t, t \in \{1, 2, \dots, 1000\})$ , verificando uma representação ARMA( $p, q$ ) com erros TGARCH( $r, s$ ) de potência  $\delta$ , isto é, tal que

$$X_t = \varphi_1 X_{t-1} + \dots + \varphi_p X_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q},$$

onde  $\varphi_i, \theta_j, (i = 1, \dots, p, j = 1, \dots, q)$  são número reais e  $\varepsilon = (\varepsilon_t, t \in \mathbb{Z})$  é um processo definido por

$$\begin{cases} \varepsilon_t = \sigma_t Z_t \\ \sigma_t^\delta = \alpha_0 + \sum_{i=1}^r [\alpha_i (\varepsilon_{t-i}^+)^{\delta} + \beta_i (\varepsilon_{t-i}^-)^{\delta}] + \sum_{j=1}^s \gamma_j \sigma_{t-j}^{\delta} \end{cases}$$

com  $Z = (Z_t, t \in \mathbb{Z})$  uma sucessão de variáveis aleatórias independentes, identicamente distribuídas, centradas e reduzidas tal que  $Z_t$  é independente de  $\varepsilon_{t-1}$  para todo o  $t \in \mathbb{Z}$ ,  $\varepsilon_t^+ = \varepsilon_t \mathbb{I}_{\{\varepsilon_t \geq 0\}}$ ,  $\varepsilon_t^- = -\varepsilon_t \mathbb{I}_{\{\varepsilon_t < 0\}}$ , e onde  $\delta > 0$ ,  $\alpha_0 > 0$ ,  $\alpha_i \geq 0$ ,  $\beta_i \geq 0$  ( $i = 1, \dots, r, j = 1, \dots, s$ ) são constantes reais.

**Nota:** Ao longo do trabalho iremos considerar  $\delta = 1$ ,  $p = 0$ ,  $q = 1$ ,  $r = 1$ ,  $s = 0$  e  $\alpha_0 = 2$ . Para níveis de significância usuais consideramos 0.05 e 0.01.

## 1.1 Exercício 1

Tendo em conta os valores indicados na nota anterior, temos

$$X_t = \varepsilon_t + \theta \varepsilon_{t-1},$$

onde  $\theta$  é um número real e  $\varepsilon = (\varepsilon_t, t \in \mathbb{Z})$  é um processo definido por

$$\begin{cases} \varepsilon_t = \sigma_t Z_t \\ \sigma_t = 2 + \alpha(\varepsilon_{t-1}^+) + \beta(\varepsilon_{t-1}^-) \end{cases}$$

com  $Z = (Z_t, t \in \mathbb{Z})$  uma sucessão de variáveis aleatórias independentes, identicamente distribuídas, centradas e reduzidas tal que  $Z_t$  é independente de  $\varepsilon_{t-1}$  para todo o  $t \in \mathbb{Z}$ ,  $\varepsilon_t^+ = \varepsilon_t \mathbb{I}_{\{\varepsilon_t \geq 0\}}$ ,  $\varepsilon_t^- = -\varepsilon_t \mathbb{I}_{\{\varepsilon_t < 0\}}$ .

O processo em causa admite então uma representação média móvel de ordem 1, MA(1), sendo definido por

$$X_t = \Theta(L) \varepsilon_t,$$

com  $\Theta(L) = 1 + \theta L$ . Assim, se considerarmos um valor para a raiz de  $\Theta$  em módulo estritamente superior a 1 (isto é, um valor para  $\theta$  em módulo estritamente inferior a 1), temos como objeto de trabalho uma solução estacionária canónica. Definimos então  $\theta = 0.5$ .

Consideremos  $Z = (Z_t, t \in \mathbb{Z})$  uma sucessão de variáveis gaussianas independentes, centradas e reduzidas tal que  $Z_t$  é independente de  $\varepsilon_{t-1}$ , para todo o  $t \in \mathbb{Z}$ .

Como  $\delta = 1$  e  $\gamma = 0$ , então  $\varepsilon$  segue um modelo TARCH(1).

Sendo a lei das v.a.r  $Z_t$  normal, então  $\varepsilon$  é condicionalmente gaussiano.

Assim, pelo **Teorema A.1**, como  $\varepsilon$  segue um modelo TARCH(1) condicionalmente gaussiano então  $\varepsilon$  é assintoticamente estacionário de 2<sup>a</sup> ordem se

$$\alpha^2 + \beta^2 < 2$$

Deste modo, escolhemos  $\alpha = 0.75$  e  $\beta = 0.5$ .

Tendo em conta esta escolha de parâmetros obtemos o processo estacionário

$$X_t = \varepsilon_t + 0.5\varepsilon_{t-1} ,$$

onde

$$\begin{cases} \varepsilon_t = \sigma_t Z_t \\ \sigma_t = 2 + 0.75(\varepsilon_{t-1}^+) + 0.5(\varepsilon_{t-1}^-) \end{cases}$$

## 1.2 Exercício 2

Após gerar 1000 observações de  $Z$ , obtivemos a seguinte trajetória do ruído (figura 1).

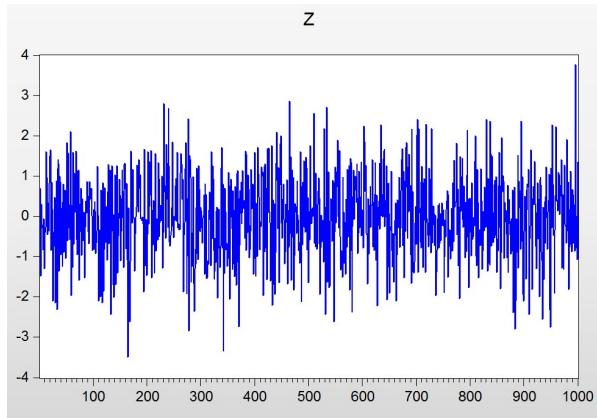


Figura 1: Trajetória de  $Z$

De modo a averiguar a independência do ruído  $Z$  gerado, realizámos o teste BDS. Este teste tem como hipóteses

$$H_0 : Z_t \text{ são i.i.d} \text{ vs } H_1 : \text{Existe dependência linear ou não linear em } Z$$

Obtivemos o resultado da figura 2.

BDS Test for Z					
Date: 11/28/22 Time: 19:17					
Sample: 1 1000					
Included observations: 1000					
Dimension	BDS Statistic	Std. Error	z-Statistic	Prob.	
2	-0.000603	0.001944	-0.310161	0.7564	
3	-0.001534	0.003084	-0.497355	0.6189	
4	-0.000955	0.003664	-0.260728	0.7943	
5	0.000264	0.003809	0.069429	0.9446	
6	-0.000295	0.003664	-0.080641	0.9357	
 Raw epsilon					
Pairs within epsilon					
Triples within epsilon					
Dimension	C(m,n)	c(m,n)	C(1,n-(m-1))	c(1,n-(m-1))	c(1,n-(m-1))^k
2	247203.0	0.495893	351256.0	0.704624	0.496496
3	173324.0	0.348388	350579.0	0.704677	0.349921
4	121739.0	0.245191	349722.0	0.704366	0.246147
5	85989.00	0.173536	348978.0	0.704280	0.173272
6	60917.00	0.123185	348963.0	0.705667	0.123481

Figura 2: Teste BDS

Pela observação dos  $p$ -valores associados ao teste, aceitamos a hipótese nula aos níveis de significância usuais, isto é, aceitamos a independência do ruído  $Z$  gerado.

De modo a validar a lei fixada, utilizámos o teste Jarque Bera e o Q-Q plot.

O teste Jarque Bera é geralmente utilizado para testar a normalidade, pelo que tem como hipóteses

$$H_0 : Z \text{ segue uma lei normal} \quad \text{vs} \quad H_1 : Z \text{ não segue uma lei normal}$$

Obtivemos os seguintes resultados (figura 3).

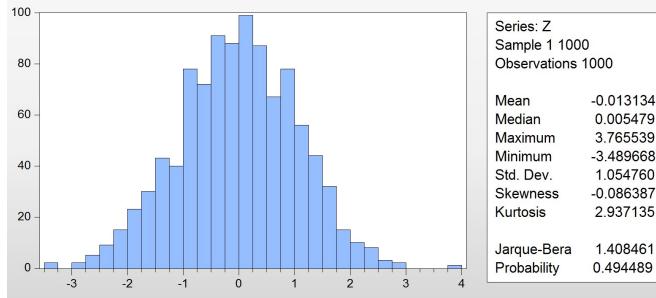


Figura 3: Teste Jarque Bera

Como o  $p$ -valor obtido é 0.494489, aceitamos a hipótese nula aos níveis de significância usuais, isto é, aceitamos que  $Z$  segue uma lei normal.

O Q-Q plot é um gráfico que compara duas distribuições de probabilidade, através dos quantis de ambas. Neste caso, vamos comparar a lei de  $Z$  com a lei normal, centrada e reduzida. Obtivemos o resultado da figura 4.

Por observação do gráfico, podemos concluir que a lei de  $Z$  se aproxima quase perfeitamente de uma lei normal, centrada e reduzida, pelo que validamos a lei fixada.

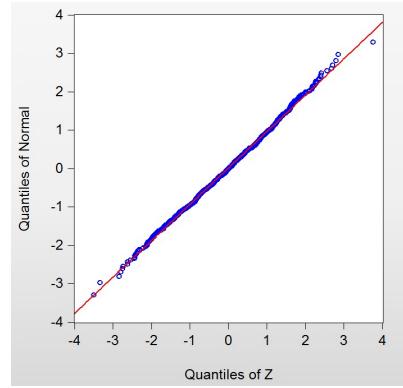


Figura 4: Q-Q plot

### 1.3 Exercício 3

De modo a obter a série dos erros ( $\varepsilon_t, t \in \{1, 2, \dots, 1000\}$ ), e sendo  $\varepsilon$  centrado (pois  $\alpha^2 + \beta^2 < 2$ ), considerámos como valor inicial  $\varepsilon_1 = 0$ , pelo que  $\sigma_1 = 2$ . Para a obter a série simulada ( $X_t, t \in \{1, 2, \dots, 1000\}$ ), considerámos como valor inicial  $X_1 = 0$ , uma vez que o processo  $X$  tem média nula (pois  $\varepsilon$  é centrado).

Obtivemos os seguintes resultados (figura 5 e 6).

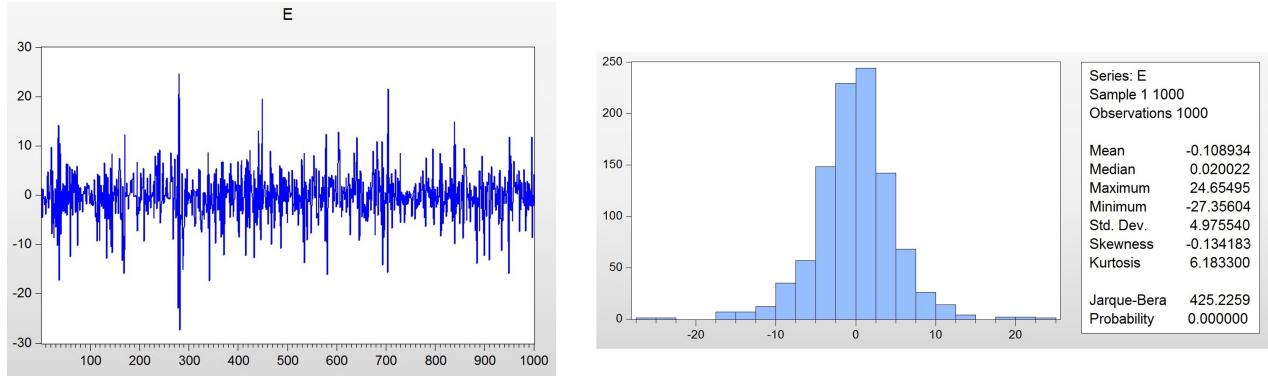


Figura 5: Trajetória e Análise Descritiva de  $\varepsilon$

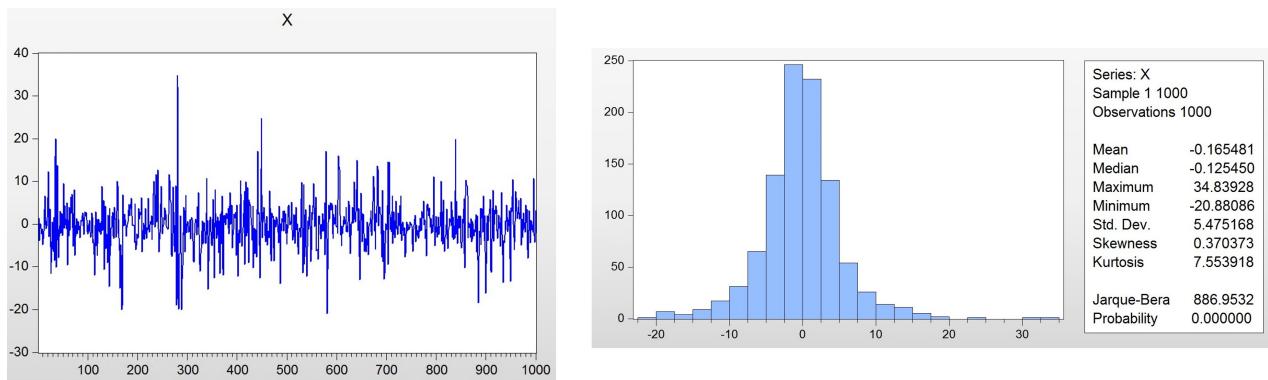


Figura 6: Trajetória e Análise Descritiva de  $X$

## 1.4 Análise de Sensibilidade

### 1.4.1 Alteração do parâmetro $\theta$

Consideremos a alteração do valor do parâmetro  $\theta$  para 10. Agora, a raíz de  $\Theta$  é em módulo inferior a 1, pelo que  $\Theta(L)$  não é invertível. Designaremos por  $X1$  a nova série obtida. É de esperar que a série  $X$  continue estacionária uma vez que temos um processo média móvel. No entanto, o resíduo não será a inovação. Obtivemos o seguinte resultado (figura 7).

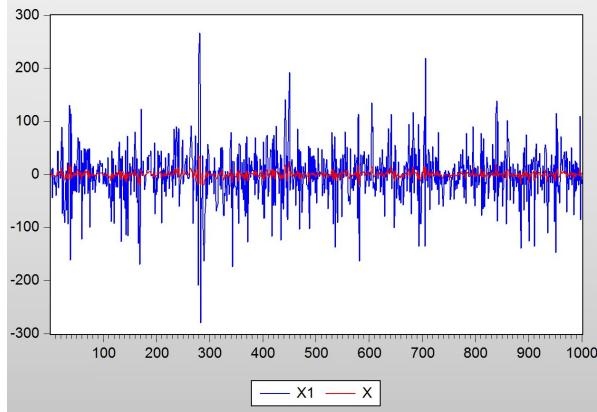


Figura 7: Trajetória de  $X$  em comparação com a trajetória de  $X1$

Como podemos observar, não houve alterações significativas na trajetória de  $X1$  em relação a  $X$ , apenas um aumento de volatilidade, como era de esperar.

### 1.4.2 Alteração dos parâmetros $\alpha$ e $\beta$

Consideremos a alteração dos valores dos parâmetros  $\alpha$  e  $\beta$  para 0.99 e 0.99, respectivamente. Designaremos por  $X2$  a nova série obtida, e por  $\sigma1$  o novo  $\sigma$ . Obtivemos os seguintes resultados (figura 8).

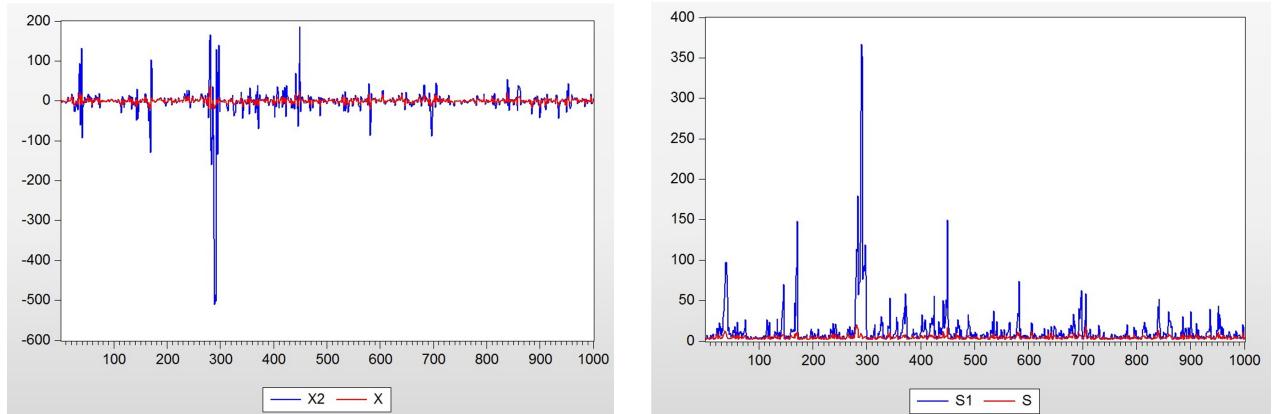


Figura 8: Trajetória de  $X$  em comparação com a trajetória de  $X2$  e trajetória de  $\sigma$  em comparação com a trajetória de  $\sigma1$

Por observação da figura podemos notar uma maior volatilidade na trajetória de  $\sigma_1$ , que poderá estar relacionada com a alteração dos parâmetros para valores próximos da zona de não estacionaridade de  $X$ .

### 1.4.3 Alteração da lei marginal de $Z$

Fixemos agora a lei de probabilidade marginal de  $Z$  como uma lei uniforme de intervalo  $[-\sqrt{3}, \sqrt{3}]$ . Designaremos por  $Z1$  o novo ruído obtido. Obtivemos a seguinte trajetória (figura 9).

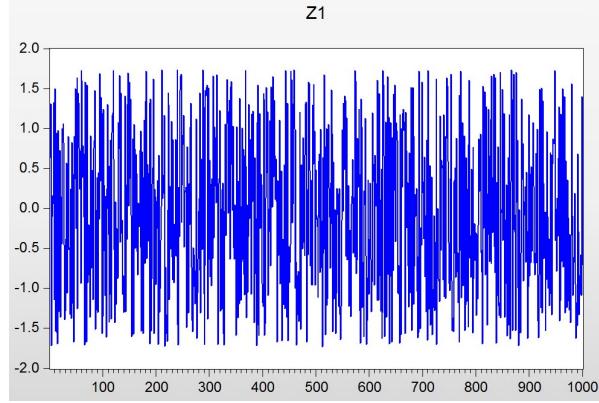


Figura 9: Trajetória de  $Z1$

Para validar a lei fixada, realizámos novamente o teste BDS e o Q-Q plot. Obtivemos os seguintes resultados (figura 10).

BDS Test for Z1					
Date: 11/28/22 Time: 19:15					
Sample: 1 1000					
Included observations: 1000					
Dimension	BDS Statistic	Std. Error	z-Statistic	Prob.	
2	-0.000152	0.001286	-0.118457	0.9057	
3	0.000448	0.002027	0.220918	0.8252	
4	0.000724	0.002392	0.302679	0.7621	
5	0.000970	0.002471	0.392615	0.6946	
6	0.000679	0.002361	0.287552	0.7737	
 Raw epsilon					
1.555725					
Pairs within epsilon					
701708.0					
Triples within epsilon					
5.13E+08					
Dimension	C(m,n)	c(m,n)	C(1,n-(m-1))	c(1,n-(m-1))	c(1,n-(m-1))^k
2	245084.0	0.491642	349589.0	0.701280	0.491794
3	172067.0	0.345861	349067.0	0.701638	0.345413
4	120727.0	0.243153	348394.0	0.701691	0.242429
5	84652.00	0.170838	347596.0	0.701491	0.169868
6	59286.00	0.119887	346921.0	0.701538	0.119208

Figura 10: Teste BDS

Pelos  $p$ -valores associados observados, aceitamos a independência do ruído  $Z$  gerado aos níveis de significância usuais.

Por observação da figura 11, podemos concluir que a lei de  $Z$  se aproxima quase perfeitamente de uma lei uniforme de intervalo  $[-\sqrt{3}, \sqrt{3}]$ .

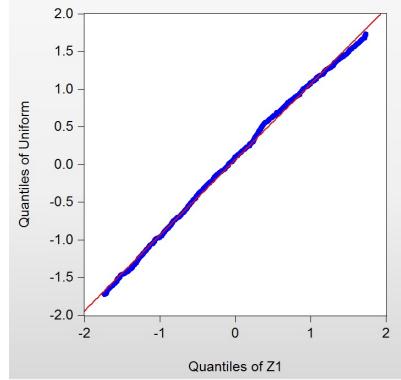


Figura 11: Q-Q plot de  $Z1$

Uma vez que estamos perante uma lei simétrica, centrada e reduzida, os parâmetros  $\alpha$  e  $\beta$  devem satisfazer a desigualdade  $\alpha^2 + \beta^2 < 2$ . Como tal, escolhemos os mesmos valores utilizados no exercício 2 ( $\alpha = 0.75$  e  $\beta = 0.5$ ). Designaremos por  $X3$  a nova série obtida. Obtivemos assim os seguintes resultados (figura 12).

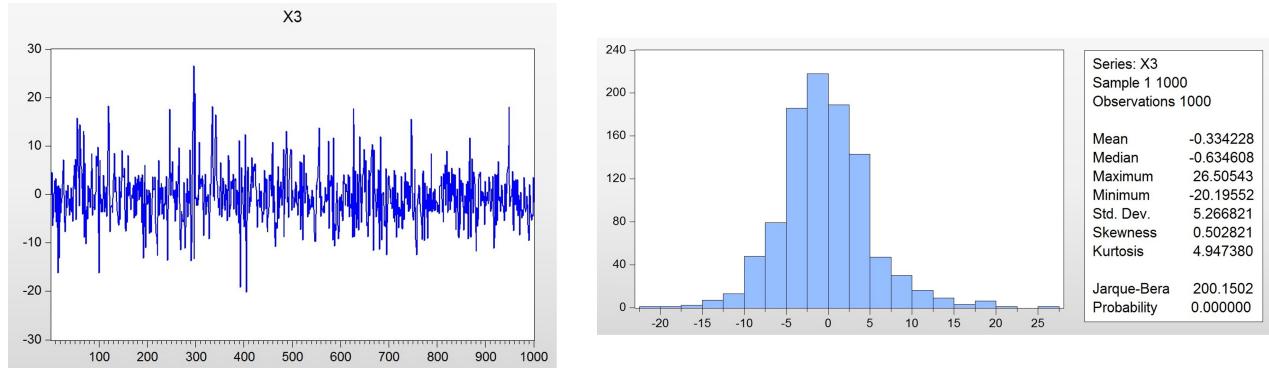


Figura 12: Trajetória e análise descritiva de  $X3$

## 2 PARTE II. Modelação estocástica de uma série temporal

Pretende-se agora ajustar um modelo estocástico a uma série temporal observada.

Escolha-se como série observada a sequência de valores  $(X_t, t \in \{51, 52, \dots, 995\})$ , obtida na primeira parte. Designaremos a série observada por  $Y_t$ .

## 2.1 Exercício 1

### 2.1.1 Alínea a

Queremos calcular as autocovariâncias, autocorrelações e autocorrelações parciais empíricas da série observada, através do *software*. Obtivemos o seguinte correlograma e análise descritiva (figura 13).

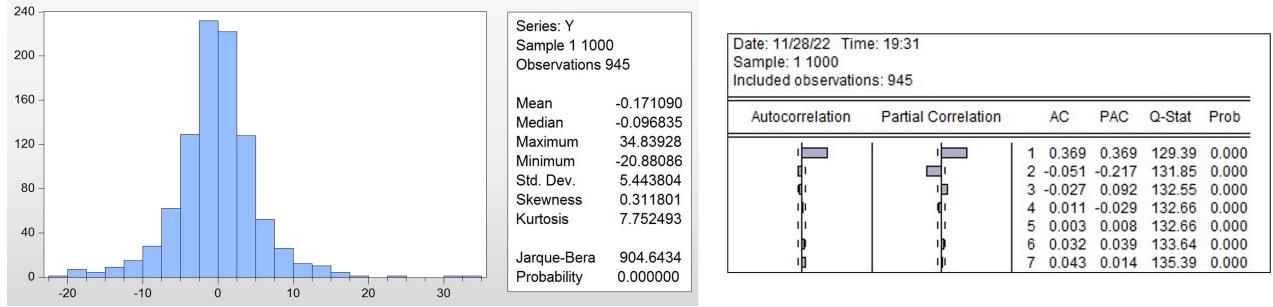


Figura 13: Análise descritiva e Correlograma de  $Y$

Pela análise do correlograma, observamos que as autocorrelações se anulam a partir da ordem 2 (inclusive).

Tem-se para a autocorrelação e autocorrelação parcial de ordem 1,  $\hat{\rho}(1) = \hat{r}(1) = 0.369$ .

Pela análise descritiva, tem-se para a variância empírica

$$\hat{\gamma}(0) = 5.443804^2 = 29.63500199$$

Assim, como  $\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)}$  vem que

$$\hat{\gamma}(h) = \hat{\gamma}(0)\hat{\rho}(h) = 29.63500199\hat{\rho}(h) ,$$

pelo que

$$\hat{\gamma}(1) = 29.63500199 \times \hat{\rho}(1) = 29.63500199 \times 0.369 = 10.93531573$$

### 2.1.2 Alínea b

Temos um MA(1) da forma  $X_t = \varepsilon_t + 0.5\varepsilon_{t-1}$ . Sabemos que  $E(X_t) = 0$ , pelo que

$$\begin{aligned} Cov(X_t, X_{t+h}) &= E(X_t X_{t+h}) = E[(\varepsilon_t + 0.5\varepsilon_{t-1})(\varepsilon_{t+h} + 0.5\varepsilon_{t+h-1})] = \\ &= E(\varepsilon_t \varepsilon_{t+h} + 0.5\varepsilon_t \varepsilon_{t+h-1} + 0.5\varepsilon_{t-1} \varepsilon_{t+h} + 0.5^2 \varepsilon_{t-1} \varepsilon_{t+h-1}) = \\ &= \begin{cases} (1 + 0.5^2)V(\varepsilon_t), & h = 0 \\ 0.5V(\varepsilon_t), & h = \pm 1 \\ 0, & h \in \mathbb{Z} \setminus \{0, \pm 1\} \end{cases} \equiv \gamma(h) \end{aligned}$$

Ora, pelo **Corolário A.2**,

$$V(\varepsilon_t) = \frac{2\alpha_0^2}{2 - (\alpha^2 + \beta^2)} \times \frac{\sqrt{2\pi} + \alpha + \beta}{\sqrt{2\pi} - (\alpha + \beta)} = \\ = \frac{2 \times 2^2}{2 - (0.75^2 + 0.5^2)} \times \frac{\sqrt{2\pi} + 0.75 + 0.5}{\sqrt{2\pi} - (0.75 + 0.5)} = 20.13945734 .$$

Logo,

$$\gamma(h) = \begin{cases} 25.17432168, & h = 0 \\ 10.06972867, & h = \pm 1 \\ 0, & h \in \mathbb{Z} \setminus \{0, \pm 1\} \end{cases}$$

Assim,

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)} = \frac{\gamma(h)}{25.17432168} = \begin{cases} 1, & h = 0 \\ 0.399(9), & h = \pm 1 \\ 0, & h \in \mathbb{Z} \setminus \{0, \pm 1\} \end{cases}$$

Além disso, tem-se para as autocorrelações parciais

$$r(1) = a_1(1) = \rho(1) = 0.399(9)$$

$$r(k) = 0, \quad k > 1$$

**NOTA:** não se pode concluir esta ultima afirmação pois não temos um AR

Temos então

	Empírico	Teórico
$\gamma(0)$	29.63500199	25.17432168
$\gamma(1)$	10.93531573	10.06972867
$\rho(1)$	0.369	0.399(9)
$r(1)$	0.369	0.399(9)

Tabela 1: Comparação entre valores empíricos e teóricos

Como os valores obtidos são muito próximos, validamos parcialmente o *software*.

## 2.2 Exercício 2

Na identificação de um modelo ARMA para a série observada, recorremos aos seguintes teoremas de caracterização:

- Uma condição necessária e suficiente para que um processo estacionário centrado admita uma representação AR é que a sucessão de autocorrelações parciais se anule a partir de uma certa ordem. A sucessão de autocorrelações parciais de um processo AR( $p$ ) anula-se a partir da ordem  $p + 1$ .
- Um processo estacionário regular admite uma representação MA de ordem finita se e só se a sucessão

das autocovariâncias se anula a partir de uma certa ordem. Se  $q + 1$  designa esta ordem mínima,  $q$  é a ordem da representação MA.

Para esta análise utilizámos o correlograma apresentado na figura 13. Verificámos pela análise do mesmo que as autocorrelações parciais se anulam praticamente a partir da ordem 2 (inclusive), o que indica que um modelo AR(1) poderá ser uma boa aproximação. Além disso, as autocorrelações também se anulam a partir da ordem 2 (inclusive), pelo que um modelo MA(1) também poderá ser uma boa aproximação.

É vantajoso que o número de parâmetros a estimar seja o menor possível, para que haja uma margem de erro menor. A nossa série original é MA(1). Se não tivéssemos informação sobre o modelo da série original e quiséssemos deduzir as ordens de um modelo ARMA, teria de ser implementado o método do Canto. Neste caso, utilizamos o método de Akaike e o método Schwarz, para concluir qual o modelo que melhor se ajusta.

	Akaike	Schwarz
AR(1)	6.085197	6.100598
MA(1)	<b>6.030708</b>	<b>6.046108</b>

Tabela 2: Comparação entre os critérios Akaike e Schwarz

Analizando a tabela, observamos que os menores valores dos testes referidos (realçados a negrito) correspondem a um modelo MA(1).

De seguida, apresenta-se o resultado da estimação de um modelo MA(2) com constante (figura 14).

Dependent Variable: Y Method: ARMA Maximum Likelihood (OPG - BHHH) Date: 11/30/22 Time: 12:05 Sample: 1 945 Included observations: 945 Convergence achieved after 37 iterations Coefficient covariance computed using outer product of gradients				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.171355	0.236082	-0.725825	0.4681
MA(1)	0.470448	0.017058	27.57967	0.0000
MA(2)	-0.022056	0.019806	-1.113646	0.2657
SIGMASQ	24.18529	0.681625	35.48183	0.0000
R-squared	0.183030	Mean dependent var	-0.171090	
Adjusted R-squared	0.180425	S.D. dependent var	5.443804	
S.E. of regression	4.928296	Akaike info criterion	6.032367	
Sum squared resid	22855.10	Schwarz criterion	6.052901	
Log likelihood	-2846.293	Hannan-Quinn criter.	6.040193	
F-statistic	70.27234	Durbin-Watson stat	1.995530	
Prob(F-statistic)	0.000000			
Inverted MA Roots	.04	.51		

Figura 14: Modelo MA(2) estimado

Como o  $p$ -valor associado ao coeficiente  $\theta_2$  (MA(2)), 0.2657, então aceitamos que este coeficiente não é significativamente diferente de 0 aos níveis de significância usuais. Assim, consideraremos a estimação para

um modelo MA(1) (figura 15).

Dependent Variable: Y Method: ARMA Maximum Likelihood (OPG - BHHH) Date: 11/30/22 Time: 12:11 Sample: 1 1945 Included observations: 945 Convergence achieved after 34 iterations Coefficient covariance computed using outer product of gradients				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.171069	0.239603	-0.713971	0.4754
MA(1)	0.482419	0.015557	31.00994	0.0000
SIGMASQ	24.19634	0.678538	35.65949	0.0000
R-squared	0.182657	Mean dependent var	-0.171090	
Adjusted R-squared	0.180922	S.D. dependent var	5.443804	
S.E. of regression	4.926804	Akaike info criterion	6.030708	
Sum squared resid	22865.54	Schwarz criterion	6.046108	
Log likelihood	-2846.509	Hannan-Quinn criter.	6.036577	
F-statistic	105.2574	Durbin-Watson stat	2.017555	
Prob(F-statistic)	0.000000			
Inverted MA Roots	-0.48			

Figura 15: Modelo MA(1) estimado

Como o  $p$ -valor associado à constante é 0.4754, logo , aceitamos que a constante não é significativamente diferente de 0 aos níveis de significância usuais, pelo que podemos considerar que a constante é nula. O  $p$ -valor associado ao coeficiente  $\theta_1$  (MA(1)) é nulo até à quarta casa decimal, pelo que aceitamos um modelo MA(1), da forma

$$Y_t = \varepsilon_t + 0.482419\varepsilon_{t-1}$$

Observando a Figura 16 verificamos que a raiz do polinómio MA(1) se encontra bem posicionada (isto é, é em módulo inferior a 1), pelo que temos um modelo estacionário em que o resíduo é a inovação.

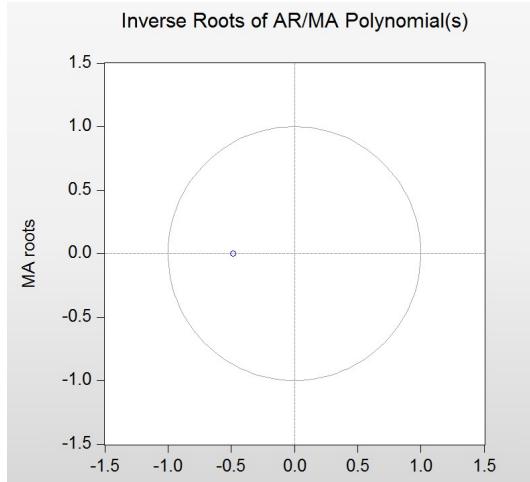


Figura 16: Raízes invertidas do polinómio MA(1)

### 2.3 Exercício 3

De forma a testar se a série dos resíduos obtidos é compatível com a hipótese do ruído branco, analisámos o respetivo correlograma e análise descritiva (figura 17).

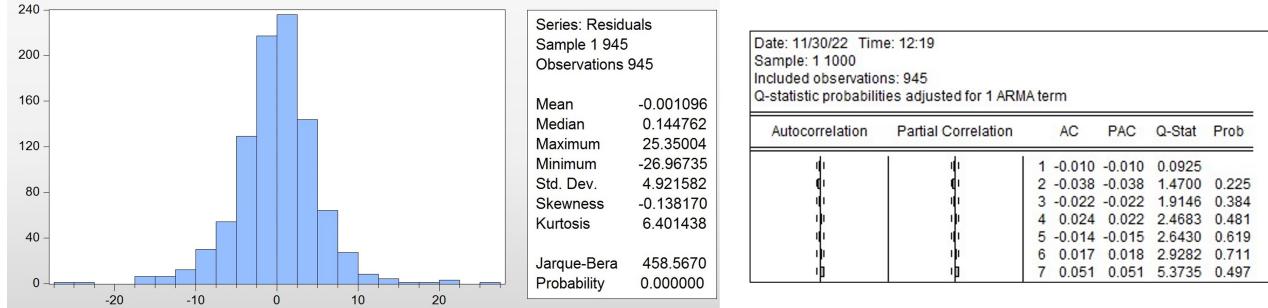


Figura 17: Análise descritiva e Correlograma dos resíduos

Sabemos que  $\varepsilon = (\varepsilon_t, t \in \mathbb{Z})$  é um ruído branco se  $\varepsilon_t$  são centrados, não correlacionados e com igual matriz de covariâncias.

Pelo correlograma, verificamos que as autocorrelações e as autocorrelações parciais são praticamente todas nulas, pelo que podemos considerar que as autocovariâncias também são nulas.

Para confirmar esta afirmação, recorremos ao histograma da série dos resíduos.

Pelo histograma, concluímos que estamos perante um processo centrado, com desvio-padrão 4.921582.

O coeficiente de curtose é superior ao da lei normal, pelo que estamos perante uma distribuição leptocúrtica, sendo a hipótese de normalidade rejeitada para níveis de significância superiores a  $10^{-6}$ .

Assim,  $\varepsilon_t$  é um processo centrado e não gaussiano. Podemos então concluir que a série dos resíduos é compatível com a de um ruído branco.

Testemos a heteroscedasticidade dos resíduos usando o teste ARCH LM, que testa a hipótese  $H_0$  : Ausência de heteroscedasticidade (figura 18).

Heteroskedasticity Test: ARCH				
F-statistic	320.6084	Prob. F(1,942)	0.0000	
Obs*R-squared	239.7056	Prob. Chi-Square(1)	0.0000	
 Test Equation:				
Dependent Variable: RESID^2				
Method: Least Squares				
Date: 11/30/22 Time: 19:11				
Sample (adjusted): 2 945				
Included observations: 944 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	11.97040	1.723778	6.944282	0.0000
RESID^2(-1)	0.503904	0.028142	17.90554	0.0000
R-squared	0.253925	Mean dependent var	24.17354	
Adjusted R-squared	0.253133	S.D. dependent var	56.29057	
S.E. of regression	48.64712	Akaike info criterion	10.60918	
Sum squared resid	2229283.	Schwarz criterion	10.61945	
Log likelihood	-5005.532	Hannan-Quinn criter.	10.61310	
F-statistic	320.6084	Durbin-Watson stat	2.171890	
Prob(F-statistic)	0.000000			

Figura 18: Teste de Heteroscedasticidade dos resíduos

Como o  $p$ -valor é nulo até à sexta casa decimal, rejeitamos a hipótese nula, ou seja, os resíduos são aceites como condicionalmente heteroscedásticos.

Procedemos ao ajustamento do modelo considerado.

Consideremos agora o correlograma do quadrado dos resíduos (figura 19).

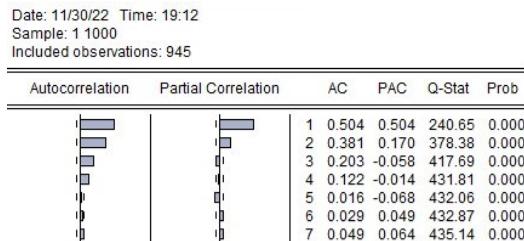


Figura 19: Correlograma do quadrado dos resíduos após modelação MA(1)

Observando a figura, fomos testar alguns modelos. Tendo em conta os critérios referidos anteriormente e o modelo da série original, optámos por fazer o ajustamento por um modelo MA(1) com processo de erro TARCH(1). As estimativas para este modelo encontram-se na figura 20.

Dependent Variable:	Y			
Method:	ML ARCH - Normal distribution (BFGS / Marquardt steps)			
Date:	12/13/22 Time: 15:43			
Sample (adjusted):	1 945			
Included observations:	945 after adjustments			
Failure to improve likelihood (non-zero gradients) after 73 iterations				
Coefficient covariance computed using outer product of gradients				
MA Backcast:	0			
Presample variance:	backcast (parameter = 0.7)			
@SQRT(GARCH) = C(3) + C(4)*(ABS(RESID(-1)) - C(5)*RESID(-1)) + C(6)				
*@SQRT(GARCH(-1))				
Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-0.073519	0.137815	-0.533465	0.5937
MA(1)	0.498934	0.021789	22.89893	0.0000
Variance Equation				
C(3)	1.818875	0.193705	9.389908	0.0000
C(4)	0.689578	0.051005	13.51993	0.0000
C(5)	-0.195311	0.056444	-3.460285	0.0005
C(6)	0.046309	0.049916	0.927750	0.3535
R-squared	0.182266	Mean dependent var	-0.171090	
Adjusted R-squared	0.181399	S.D. dependent var	5.443804	
S.E. of regression	4.925369	Akaike info criterion	5.613047	
Sum squared resid	22876.48	Schwarz criterion	5.643849	
Log likelihood	-2646.165	Hannan-Quinn criter.	5.624786	
Durbin-Watson stat	2.049561			
Inverted MA Roots	- .50			

Figura 20: Estimação do modelo MA(1)-TARCH(1)

Assim, aceitamos que todos os coeficientes são significativamente não nulos, exceto a constante  $C$  do modelo MA(1) e a constante  $C(6)$ .

Para obter os coeficientes  $\alpha$  e  $\beta$  do novo modelo tivemos em conta que

$$\alpha = C(4) \times (1 - C(5)) = 0.689578 \times (1 + 0.195311) = 0.824260168758$$

$$\beta = C(4) \times (1 + C(5)) = 0.689578 \times (1 - 0.195311) = 0.554895831242$$

Obtivemos o seguinte correlograma dos resíduos e a correspondente análise descritiva após o novo ajustamento (figura 21).

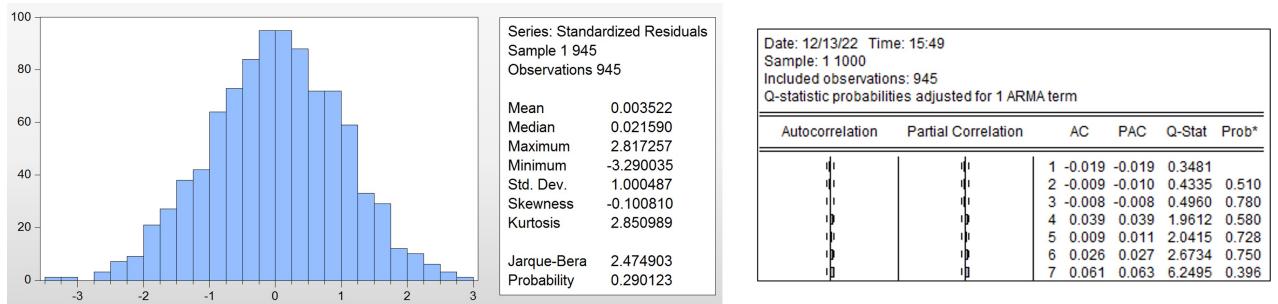


Figura 21: Correlograma dos resíduos e análise descritiva após modelação MA(1)-TARCH(1)

Analizando a figura notamos que o resíduo poderá ser compatível com um ruído branco gaussiano. Vejamos agora se no modelo ajustado se verifica heteroscedasticidade condicional dos resíduos, realizando novamente o teste ARCH-LM (figura 22).

Heteroskedasticity Test ARCH				
F-statistic	0.245399	Prob. F(1,942)	0.6204	
Obs*R-squared	0.245855	Prob. Chi-Square(1)	0.6200	
<hr/>				
Test Equation:				
Dependent Variable: WGT_RESID^2				
Method: Least Squares				
Date: 12/13/22 Time: 15:53				
Sample (adjusted): 2945				
Included observations: 944 after adjustments				
<hr/>				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.015270	0.055006	18.45741	0.0000
WGT_RESID^2(-1)	-0.016136	0.032573	-0.495377	0.6204
<hr/>				
R-squared	0.000260	Mean dependent var	0.999130	
Adjusted R-squared	-0.000801	S.D. dependent var	1.361111	
S.E. of regression	1.361655	Akaike info criterion	3.457396	
Sum squared resid	1746.567	Schwarz criterion	3.467672	
Log likelihood	-1629.891	Hannan-Quinn criter.	3.461312	
F-statistic	0.245399	Durbin-Watson stat	2.000289	
Prob(F-statistic)	0.620449			
<hr/>				

Figura 22: Teste ARCH aos resíduos do modelo ajustado

Sendo o  $p$ -valor do teste 0.6204, superior aos níveis de significância usuais, aceitamos o resíduo como homoscedástico.

Assim, podemos modelar a série observada da forma

$$Y_t = \varepsilon_t + 0.482419\varepsilon_{t-1}$$

onde  $\varepsilon = (\varepsilon_t, t \in \mathbb{Z})$  é um processo definido por

$$\begin{cases} \varepsilon_t = \sigma_t Z_t \\ \sigma_t = 1.818875 + 0.824260168758(\varepsilon_{t-1}^+) + 0.554895831242(\varepsilon_{t-1}^-) \end{cases}$$

## 2.4 Exercício 4

Após a estimação do modelo previmos os valores de  $(Y_t, t \in \{996, 997, 998, 999, 1000\})$ . A tabela seguinte mostra os valores observados para  $(X_t, t \in \{996, 997, 998, 999, 1000\})$ , e os valores previstos com base no modelo obtido na secção anterior.

$t$	Valores Observados	Valores Previstos
996	10.68695	10.31555
997	-2.660141	-3.421102
998	-4.260822	-4.533296
999	-2.140811	-1.948485
1000	3.082154	2.949376

Tabela 3: Comparação entre os valores previstos e observados

Observando a tabela podemos concluir que existe diferença entre os valores obtidos, mas que não é significativa.

Por fim apresentamos os gráficos das previsões da série ajustada (figura 23) e o gráfico do desvio padrão condicional (figura 24).

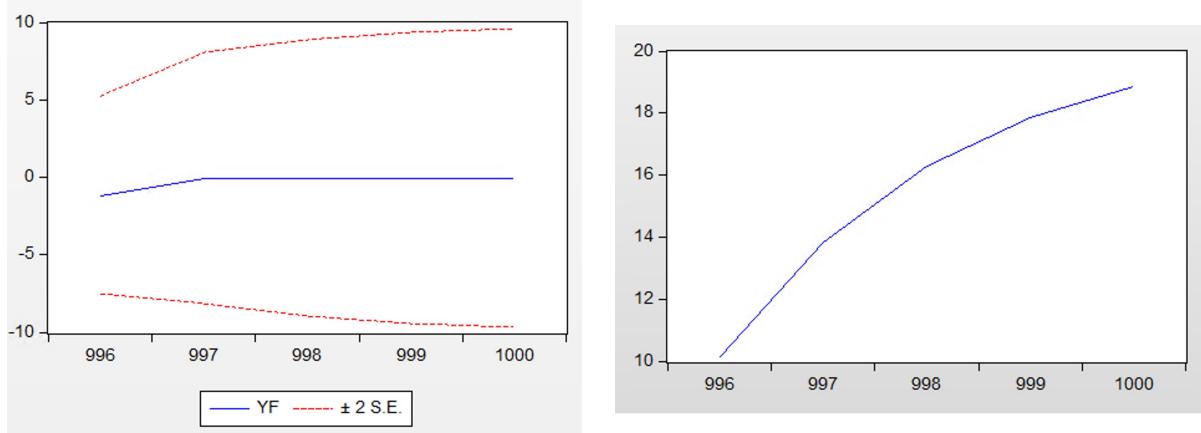


Figura 23: Previsão dos valores futuros e respetiva variância

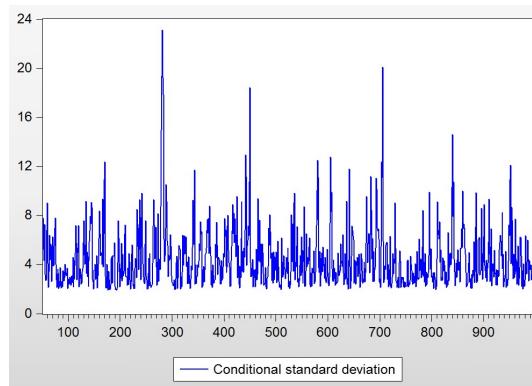


Figura 24: Desvio Padrão Condicional

Apresentemos também as trajetórias da série ajustada, da série observada e do resíduo (figura 25).

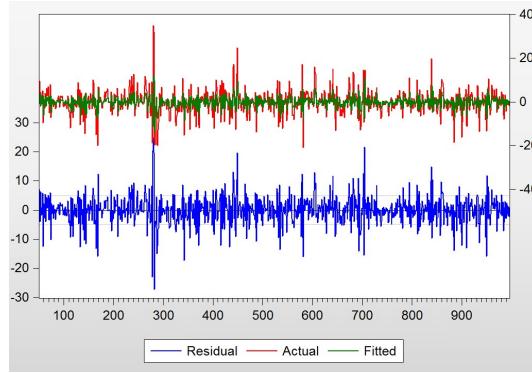


Figura 25: Gráfico da série observada, ajustada e resíduo

### 3 PARTE III. Modelação e previsão da cotação diária de uma ação do Índice PSI 20

Considere-se agora a série temporal relativa ao preço de fecho diário, entre 14/09/2020 e 12/09/2022, da ação do Índice PSI 20 correspondente à *Green Volt Energias*.

Analisemos a adequação de um modelo estocástico da classe ARMA com erros condicionalmente heteroscedásticos para reproduzir a dinâmica subjacente a esta série temporal.

Observação: A *Green Volt Energias* é uma empresa de energias renováveis criada em março de 2021, estando cotada na Euronext Lisboa desde julho de 2021. Deste modo, foram apenas considerados os dados a partir de 14/09/2021.

Comecemos por apresentar a trajetória e a análise descritiva relativas à série temporal em questão (figura 26).

Por observação da trajetória, deduzimos que há uma tendência crescente a partir de fevereiro de 2022. Através da análise descritiva, observamos que o valor da curtose é superior a 3, pelo que a distribuição é leptocúrtica. Além disso, o *p*-valor associado ao teste Jarque-Bera é nulo até à sexta casa decimal, pelo que a distribuição é não gaussiana.

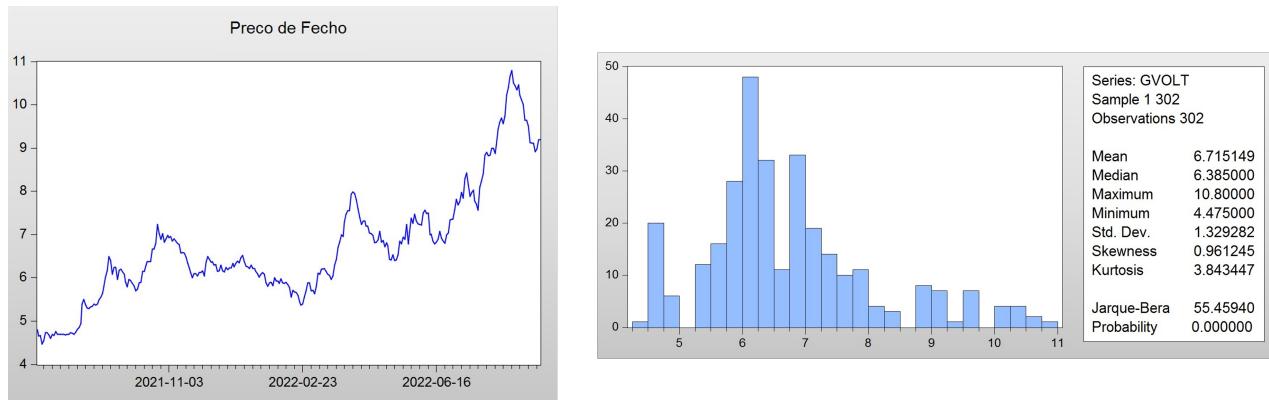


Figura 26: Cotações diárias e Análise descritiva da série GVolt

Observemos agora o correlograma da série GVolt (figura 27).

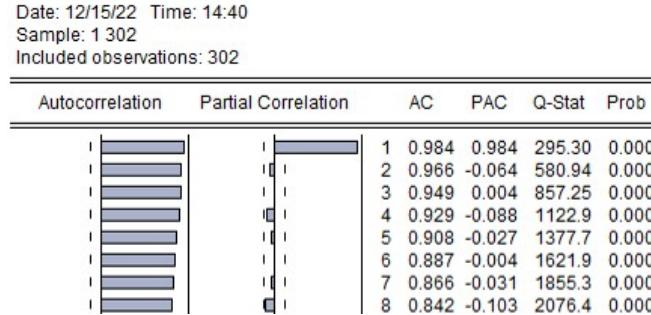


Figura 27: Correlograma da série GVolt

Como vemos pelo correlograma, a autocorrelação parcial de ordem 1 é muito próxima de 1. Além disso, as autocorrelações decrescem lentamente, o que é um indício de que a série é não estacionária. Assim, deve proceder-se à diferenciação da mesma. Este indício pode ser confirmado com o teste de estacionaridade Augmented Dickey-Fuller (figura 28).

Null Hypothesis: GVOLT has a unit root		
Exogenous: Constant		
Lag Length: 0 (Automatic - based on SIC, maxlag=15)		
	t-Statistic	Prob.*
<u>Augmented Dickey-Fuller test statistic</u>	-0.655646	0.8545
Test critical values:		
1% level	-3.451920	
5% level	-2.870931	
10% level	-2.571845	

Figura 28: Teste Augmented Dickey-Fuller

O *p*-valor do teste é 0.8545, superior aos níveis de significância usuais, pelo que aceitamos a não estacionaridade do processo.

Procedemos então à diferenciação da série GVolt obtendo uma nova série, que designaremos por DGvolt. Começámos por analisar a estacionaridade (figura 29).

Null Hypothesis: DGVOLT has a unit root		
Exogenous: Constant		
Lag Length: 0 (Automatic - based on SIC, maxlag=15)		
	t-Statistic	Prob.*
<u>Augmented Dickey-Fuller test statistic</u>	-15.87248	0.0000
Test critical values:		
1% level	-3.451993	
5% level	-2.870964	
10% level	-2.571862	

Figura 29: Teste Augmented Dickey-Fuller para a série diferenciada

O *p*-valor do teste é agora nulo até à quarta casa decimal pelo que aceitamos a estacionaridade da nova série. Analisemos agora o correlograma (figura 30).

Date: 12/18/22 Time: 18:17  
 Sample: 1 302  
 Included observations: 301

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
1	1	1	0.085	0.085	2.2184 0.136
2	2	2	0.045	0.038	2.8496 0.241
3	3	3	0.147	0.141	9.4520 0.024
4	4	4	-0.003	-0.029	9.4554 0.051
5	5	5	0.025	0.017	9.6447 0.086
6	6	6	0.047	0.025	10.327 0.112
7	7	7	0.068	0.068	11.780 0.108
8	8	8	0.015	-0.004	11.846 0.158

Figura 30: Correlograma da série diferenciada

Observando o correlograma, vemos que existe uma autocorrelação significativamente não nula de ordem 3, pelo que parece surgir a necessidade de usar um modelo auto-regressivo de ordem 3. Após testar alguns modelos e a heteroscedasticidade dos resíduos correspondentes, aceitámos a homoscedasticidade do resíduo. Tendo em conta os critérios utilizados anteriormente e o correlograma do quadrado dos resíduos, concluímos que o modelo que melhor se ajusta é o AR(3), sendo que o único coeficiente aceite como significativamente diferente de 0, aos níveis de significância usuais, é o  $\theta_3$  (AR(3)). Na figura 32 apresenta-se a estimativa do parâmetro do modelo escolhido.

Date: 12/18/22 Time: 18:31  
 Sample: 1 302  
 Included observations: 301  
 Q-statistic probabilities adjusted for 1 ARMA term

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
1	1	1	0.083	0.083	2.0887
2	2	2	0.031	0.024	2.3848 0.123
3	3	3	-0.004	-0.008	2.3890 0.303
4	4	4	-0.027	-0.027	2.6082 0.456
5	5	5	0.016	0.021	2.6876 0.611
6	6	6	0.028	0.026	2.9221 0.712
7	7	7	0.078	0.073	4.8157 0.568

Figura 31: Correlograma dos resíduos após estimação AR

Dependent Variable: DGVOLT				
Method: ARMA Maximum Likelihood (OPG - BHHH)				
Date: 12/18/22 Time: 18:31				
Sample: 2 302				
Included observations: 301				
Convergence achieved after 8 iterations				
Coefficient covariance computed using outer product of gradients				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.014548	0.010498	1.385715	0.1669
AR(3)	0.147743	0.048863	3.023593	0.0027
SIGMASQ	0.023286	0.001542	15.09683	0.0000
R-squared	0.021918	Mean dependent var	0.014618	
Adjusted R-squared	0.015353	S.D. dependent var	0.154553	
S.E. of regression	0.153362	Akaike info criterion	-0.901889	
Sum squared resid	7.008969	Schwarz criterion	-0.864941	
Log likelihood	138.7343	Hannan-Quinn criter.	-0.887104	
F-statistic	3.338903	Durbin-Watson stat	1.830829	
Prob(F-statistic)	0.036808			
Inverted AR Roots	.53	-.26+.46i	-.26-.46i	

Figura 32: Estimação dos parâmetros do modelo AR(3)

Pelo correlograma da figura 31, podemos inferir que a estrutura do resíduo é compatível com a de um ruído branco.

Para averiguar a estacionaridade do modelo, representámos graficamente os inversos das raízes do polinómio auto-regressivo (figura 33). Sendo todas as raízes de módulo inferior a 1, concluímos que estamos perante um modelo estacionário em que o ruído branco é a inovação.

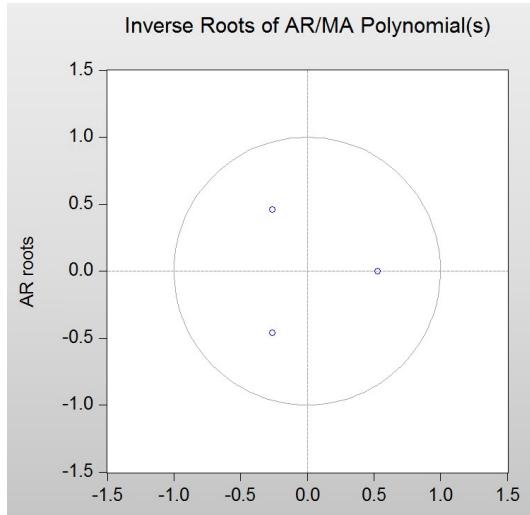


Figura 33: Inverso das raízes dos polinómios AR

De seguida apresentamos a análise descritiva do resíduo (figura 34). Pelo teste Jarque-Bera, sendo o *p*-valor nulo até à quarta casa decimal, aceitamos o resíduo como não gaussiano. Além disso, observamos que a média do resíduo é praticamente nula e que tem desvio padrão 0.152850. O coeficiente de assimetria é ligeiramente positivo e a curtose é superior a 3, pelo que estamos perante uma distribuição leptocúrtica.

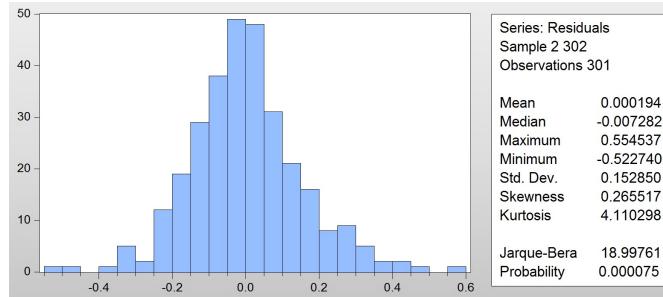


Figura 34: Análise descritiva dos resíduos após estimação AR

Utilizando o teste ARCH-LM (figura 35) testámos então a heteroscedasticidade dos resíduos, e sendo o *p*-valor associado superior aos níveis de significância usuais, aceitamos a homescedasticidade dos resíduos. Temos assim um ruído branco centrado não gaussiano de desvio padrão 0.152850.

Heteroskedasticity Test: ARCH				
F-statistic	0.434620	Prob. F(1,298)	0.5102	
Obs*R-squared	0.436900	Prob. Chi-Square(1)	0.5086	
<b>Test Equation:</b>				
Dependent Variable:	RESID^2			
Method:	Least Squares			
Date:	12/18/22	Time:	18:51	
Sample (adjusted):	3 302			
Included observations:	300 after adjustments			
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.022393	0.002739	8.175222	0.0000
RESID^2(-1)	0.038182	0.057916	0.659257	0.5102
R-squared	0.001456	Mean dependent var	0.023285	
Adjusted R-squared	-0.001894	S.D. dependent var	0.041213	
S.E. of regression	0.041252	Akaike info criterion	-3.531606	
Sum squared resid	0.507107	Schwarz criterion	-3.506914	
Log likelihood	531.7409	Hannan-Quinn criter.	-3.521724	
F-statistic	0.434620	Durbin-Watson stat	2.001628	
Prob(F-statistic)	0.510240			

Figura 35: Teste à heteroscedasticidade dos resíduos

O modelo retido é, então, sendo  $X_t$  a cotação de fecho da GVolt no dia  $t$ ,

$$Y_t = X_t - X_{t-1}$$

com

$$Y_t = 0.147743Y_{t-3} + \varepsilon_t$$

onde  $\varepsilon_t$  ruído branco centrado de desvio padrão 0.152850.

```

Estimation Command:
=====
LS(OPTMETHOD=OPG) DGVOLT C AR(3)

Estimation Equation:
=====
DGVOLT = C(1) + [AR(3)=C(2),UNCOND]

Substituted Coefficients:
=====
DGVOLT = 0.0145475127348 + [AR(3)=0.147742638603,UNCOND]

```

Figura 36: Modelo estimado

O software utilizado permite ainda comparar, visualmente, as trajetórias da série diferenciada e da correspondente série estimada bem como a trajetória do resíduo associado, como se ilustra na figura 37.

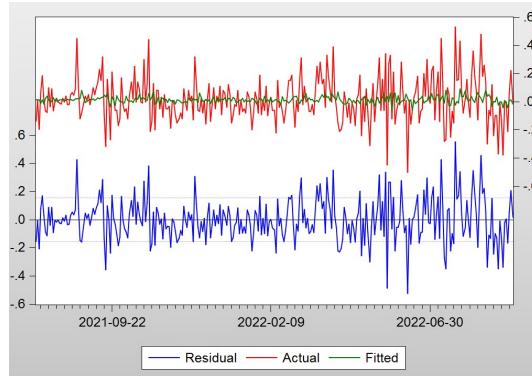


Figura 37: Gráficos das séries GVolt diferenciada, ajustada e residual

Façamos por fim a previsão dos valores da série DG Volt no período de 6 a 12 de setembro de 2022 (figura 38).

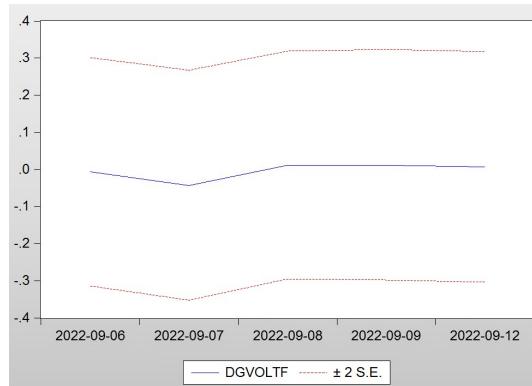


Figura 38: Previsão da série DG Volt

# Anexo

## A Resultados Importantes

### Teorema A.1

1. Se  $\varepsilon$  é um p.e.r. estacionário seguindo um modelo TARCH(1) condicionalmente Gaussiano então

$$\alpha^2 + \beta^2 < 2 .$$

2. Se  $\varepsilon$  segue um modelo TARCH(1) condicionalmente Gaussiano tal que

$$\alpha^2 + \beta^2 < 2 ,$$

então  $\varepsilon$  é assintoticamente estacionário de 2<sup>a</sup> ordem.

**Corolário A.2** Se  $\alpha^2 + \beta^2 < 2$ , então  $\varepsilon$  centrado e tem variância dada por

$$V(\varepsilon_t) = \frac{2\alpha_0^2}{2 - (\alpha^2 + \beta^2)} \times \frac{\sqrt{2\pi} + \alpha + \beta}{\sqrt{2\pi} - (\alpha + \beta)}$$

## B Códigos Eviews

### B.1 Parte I

#### B.1.1 Exercício 2

```
smpl @first 1000
genr z = nrnd
```

#### B.1.2 Exercício 3

GERAR ERROS:

```
smpl @first @last
series e=0
series ep=0
series en=0
series s=2
for !j=2 to 1000
s(!j)=2+0.75*ep(!j-1)+0.5*en(!j-1)
e(!j)=s(!j)*z(!j)
if e(!j)>=0 then
ep(!j)=e(!j)
en(!j)=0
```

```

else
ep(!j)=0
en(!j)=-e(!j)
endif
next
GERAR X:
smpl @first @last
series x=0
for !i=2 to 1000
x(!i)=e(!i)+0.5*e(!i-1)
next

```

### B.1.3 Análise de sensibilidade

ALTERAR TETA:

```

smpl @first @last
series x1=0
for !i=2 to 1000
x1(!i)=e(!i)+10*e(!i-1)
next

```

ALTERAR ALFA E BETA:

```

smpl @first @last
series e1=0
series ep1=0
series en1=0
series s1=2
for !j=2 to 1000
s1(!j)=2+0.99*ep1(!j-1)+0.99*en1(!j-1)
e1(!j)=s1(!j)*z(!j)
if e1(!j)>=0 then
ep1(!j)=e1(!j)
en1(!j)=0
else
ep1(!j)=0
en1(!j)=-e1(!j)
endif
next
smpl @first @last
series x2=0
for !i=2 to 1000

```

```

x2(!i)=e1(!i)+0.5*e1(!i-1)
next
ALTERAR LEI DE Z:
smpl @first 1000
genr z1 =2*sqr(3)*rnd-sqr(3)
smpl @first @last
series e2=0
series ep2=0
series en2=0
series s2=2
for !j=2 to 1000
s2(!j)=2+0.75*ep2(!j-1)+0.5*en2(!j-1)
e2(!j)=s2(!j)*z1(!j)
if e2(!j)>=0 then
ep2(!j)=e2(!j)
en2(!j)=0
else
ep2(!j)=0
en2(!j)=-e2(!j)
endif
next
smpl @first @last
series x3=0
for !i=2 to 1000
x3(!i)=e2(!i)+0.5*e2(!i-1)
next

```

## B.2 Parte II

### B.2.1 Exercício 1 (a)

```

smpl @first 1000
series y
for !i=51 to 945
y(!i)=x(!i)
next

```

### B.2.2 Exercício 4

CALCULAR OS VALORES PREVISTOS

```
smpl @first @last
```

```

series e3=0
series ep3=0
series en3=0
series s3=1.818875
for !j=2 to 1000
s3(!j)=1.818875+0.824260168758*ep3(!j-1)+0.554895831242*en3(!j-1)
e3(!j)=s3(!j)*z(!j)
if e3(!j)>=0 then
ep3(!j)=e(!j)
en3(!j)=0
else
ep3(!j)=0
en3(!j)=-e3(!j)
endif
next
smpl @first @last
series y3=0
for !i=2 to 1000
y3(!i)=e3(!i)+0.5*e3(!i-1)
next

```

### B.3 Parte III

```
series dgvolt=d(gvolt)
```