

Modelo ARIMA(p, d, q):

Um processo $\{X_t, t \in Z\}$ segue um modelo ARIMA(p,d,q) se $\Delta^d X_t$ seguir um modelo ARMA(p,q), ou seja, temos

$$\phi(B)\Delta^d X_t = \theta(B)\varepsilon_t. \quad (3.3)$$

No estabelecimento de um modelo ARIMA para uma série temporal há três estágios a considerar:

- (i) identificação;
- (ii) estimação;
- (iii) diagnóstico.

Identificação

é feita principalmente com base nas **autocorrelações** e **autocorrelações parciais** estimadas.

a f.a.c. ρ_j é estimada por

$$r_j = \frac{c_j}{c_0}, \quad j = 0, 1, \dots, T-1, \quad (3.4)$$

onde c_j é a estimativa da f.a.c.v γ_j ,

$$c_j = \frac{1}{T} \sum_{t=1}^{T-j} [(X_t - \bar{X})(X_{t+j} - \bar{X})], \quad j = 0, 1, \dots, T-1, \quad (3.5)$$

sendo $\bar{X} = \frac{1}{T} \sum_{t=1}^T X_t$ a média amostral. Como as autocorrelações verdadeiras são funções pares, colocamos $c_{-j} = c_j$ e $r_{-j} = r_j$.

Uma expressão aproximada para a variância de r_j , para um processo estacionário normal, é dada por

$$\text{Var}(r_j) \simeq \frac{1}{T} \sum_{v=-\infty}^{\infty} [\rho_v^2 + \rho_{v+j}\rho_{v-j} - 4\rho_j\rho_v\rho_{v-j} + 2\rho_v^2\rho_j^2]. \quad (3.6)$$

Para um processo em que as autocorrelações são nulas para $v > q$, todos os termos do lado direito de (3.6) anulam-se para $j > q$, exceto o primeiro, obtendo-se

$$\text{Var}(r_j) \simeq \frac{1}{T} \left[1 + 2 \sum_{v=1}^q \rho_v^2 \right], \quad j > q. \quad (3.7)$$

Como desconhecemos as autocorrelações ρ_v , substituímo-las por r_v , obtendo-se uma estimativa para (3.7),

$$\widehat{\sigma}^2(r_j) \simeq \frac{1}{T} \left[1 + 2 \sum_{v=1}^q r_v^2 \right], \quad j > q. \quad (3.8)$$

Para T suficientemente grande e sob a hipótese que $\rho_j = 0$, para $j > q$, a distribuição de r_j é aproximadamente normal, com média igual a zero e variância dada por (3.7) (Jenkins e Watts, 1968, p.187). Assim, pode-se construir um intervalo de confiança aproximado para as autocorrelações,

$$r_j \pm t_\gamma \cdot \widehat{\sigma}(r_j), \quad (3.9)$$

onde t_γ é o valor da estatística t de Student com $T - 1$ graus de liberdade, tal que $P(-t_\gamma < t < t_\gamma) = \gamma$. Na prática usa-se $t_\gamma = 2$, correspondendo, aproximadamente, a $\gamma = 0,95$, de modo que podemos considerar ρ_j significativamente diferente de zero se

$$|r_j| > 2\widehat{\sigma}(r_j), \quad j > q. \quad (3.10)$$

Os processos AR(p), MA(q) e ARMA(p, q) apresentam f.a.c. com características especiais:

- (i) um processo AR(p) tem f.a.c. que decai de acordo com exponenciais e/ou senoides amortecidas, infinita em extensão;
- (ii) um processo MA(q) tem f.a.c. finita, no sentido que ela apresenta um corte após o “lag” q ;
- (iii) um processo ARMA(p, q) tem f.a.c. infinita em extensão, que decai de acordo com exponenciais e/ou senoides amortecidas após o “lag” $q - p$.

f.a.c. é útil para identificar modelos **MA**, mas não para **AR** e **ARMA**.

Box et al. (1994) propõem a utilização de um outro instrumento para facilitar este procedimento de identificação: a *função de autocorrelação parcial* (f.a.c.p.).

Vamos denotar por ϕ_{kj} o j -ésimo coeficiente de um modelo AR(k), de tal modo que ϕ_{kk} seja o último coeficiente. Sabemos que

$$\rho_j = \phi_{k1}\rho_{j-1} + \phi_{k2}\rho_{j-2} + \cdots + \phi_{kk}\rho_{j-k}, \quad j = 1, \dots, k,$$

a partir das quais obtemos as equações de Yule-Walker

$$\begin{bmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-2} \\ \vdots & \vdots & \vdots & \vdots & \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & 1 \end{bmatrix} \begin{bmatrix} \phi_{k1} \\ \phi_{k2} \\ \vdots \\ \phi_{kk} \end{bmatrix} = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_k \end{bmatrix} \quad (3.11)$$

Resolvendo essas equações sucessivamente para $k = 1, 2, 3, \dots$ obtemos

$$\begin{aligned} \phi_{11} &= \rho_1, \\ \phi_{22} &= \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}, \\ \phi_{33} &= \frac{\begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}} \end{aligned}$$

e, em geral,

$$\phi_{kk} = \frac{|\mathbf{P}_k^*|}{|\mathbf{P}_k|},$$

onde \mathbf{P}_k é a matriz de autocorrelações e \mathbf{P}_k^* é a matriz \mathbf{P}_k com a última coluna substituída pelo vetor de autocorrelações.

A quantidade ϕ_{kk} , encarada como função de k , é chamada *função de autocorrelação parcial*.

Pode-se demonstrar que ϕ_{kk} é igual à correlação parcial entre as variáveis X_t e X_{t-k} ajustadas às variáveis intermediárias $X_{t-1}, \dots, X_{t-k+1}$. Ou seja, ϕ_{kk} mede a correlação remanescente entre X_t e X_{t-k} depois de eliminada a influência de $X_{t-1}, \dots, X_{t-k+1}$.

Exemplo 3.1. A correlação entre os valores ajustados $X_t - \phi_{11}X_{t-1}$ e $X_{t-2} - \phi_{11}X_{t-1}$ é dada por

$$\begin{aligned} & \text{Corr}(X_t - \phi_{11}X_{t-1}, X_{t-2} - \phi_{11}X_{t-1}) \\ &= \text{Corr}(X_t - \rho_1X_{t-1}, X_{t-2} - \rho_1X_{t-1}) \\ &= \frac{\text{Cov}(X_t - \rho_1X_{t-1}, X_{t-2} - \rho_1X_{t-1})}{(\text{Var}(X_t - \rho_1X_{t-1})\text{Var}(X_{t-2} - \rho_1X_{t-1}))^{1/2}} \\ &= \frac{\gamma_2 - 2\rho_1\gamma_1 + \rho_1^2\gamma_0}{\gamma_0 - 2\rho_1\gamma_1 + \rho_1^2\gamma_0} \\ &= \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2} = \phi_{22}. \end{aligned}$$

Será necessário estimar a f.a.c.p. de um processo AR, MA ou ARMA. Uma maneira consiste em estimar, sucessivamente, modelos autorregressivos de ordens $p = 1, 2, 3, \dots$ por mínimos quadrados e tomar as estimativas do último coeficiente de cada ordem.

Outra maneira consiste em substituir nas equações de Yule-Walker as f.a.c. ρ_j por suas estimativas, ou seja, considerar

$$r_j = \hat{\phi}_{k1}r_{j-1} + \dots + \hat{\phi}_{kk}r_{j-k}, \quad j = 1, \dots, k,$$

e resolver (3.11) com ρ_j substituída por r_j e ϕ_{kj} substituída por $\hat{\phi}_{kj}$.

Quenouille (1949) mostra que, sob a suposição que o processo seja AR(p), as f.a.c.p. estimadas de ordem $p + 1, p + 2, \dots$ são, aproximadamente, independentemente distribuídas, com

$$\text{Var}(\hat{\phi}_{kk}) \approx \frac{1}{T}, \quad k \geq p + 1.$$

Se o número de observações, T , for suficientemente grande, $\hat{\phi}_{kk}$ tem distribuição aproximada normal, o que permite a construção de intervalos de confiança para ϕ_{kk} .

- (i) um processo AR(p) tem f.a.c.p. $\phi_{kk} \neq 0$, para $k \leq p$ e $\phi_{kk} = 0$, para $k > p$;
- (ii) um processo MA(q) tem f.a.c.p. que se comporta de maneira similar à f.a.c. de um processo AR(p): é dominada por exponenciais e/ou senoides amortecidas;
- (iii) um processo ARMA(p, q) tem f.a.c.p. que se comporta como a f.a.c.p. de um processo MA puro.

f.a.c.p. é útil para identificar modelos **AR puros**, não sendo tão útil para identificar modelos **MA e ARMA**.

sob a hipótese que o processo seja $AR(p)$,

$$\text{Var}(\hat{\phi}_{jj}) \simeq \frac{1}{T}, \quad j > p, \quad (3.12)$$

de modo que

$$\hat{\sigma}(\hat{\phi}_{jj}) \simeq \frac{1}{\sqrt{T}}, \quad j > p. \quad (3.13)$$

Além disso, para T grande e sob a hipótese que o processo seja $AR(p)$, $\hat{\phi}_{jj}$ terá distribuição aproximadamente normal, com média zero e variância (3.12), de modo que consideraremos ϕ_{jj} significativamente diferente de zero se

Exemplo 3.2. Na Tabela 3.1, temos as estimativas das autocorrelações e de seus respectivos desvios padrões, bem como as autocorrelações parciais estimadas, para as séries simuladas das figuras 2.6, 2.8 e 2.9 , a saber,

$$\begin{aligned} \text{AR}(1) : \quad & X_t = 0,8X_{t-1} + \varepsilon_t; \\ \text{MA}(1) : \quad & X_t = \varepsilon_t - 0,8\varepsilon_{t-1}; \\ \text{ARMA}(1,1) : \quad & X_t = 0,8X_{t-1} + \varepsilon_t - 0,3\varepsilon_{t-1}. \end{aligned}$$

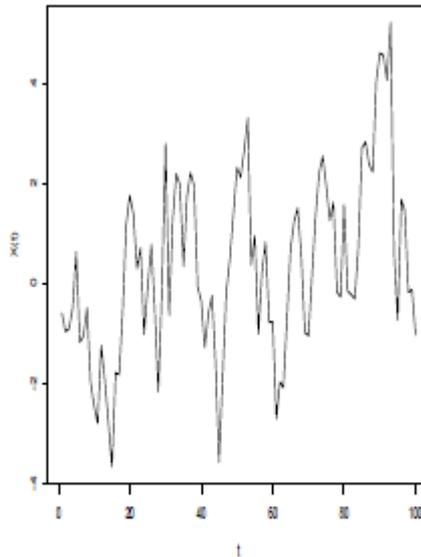
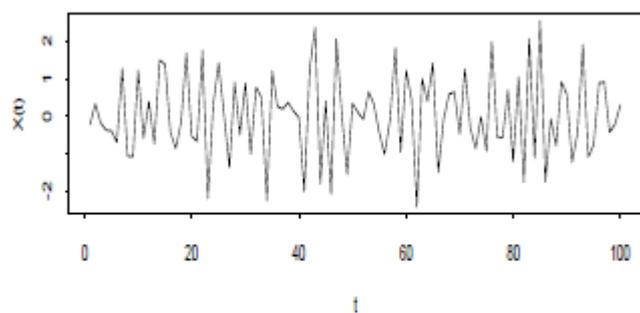


Figura 2.6: Processo $AR(1)$ simulado, $\phi = 0,8$



Series : ma1

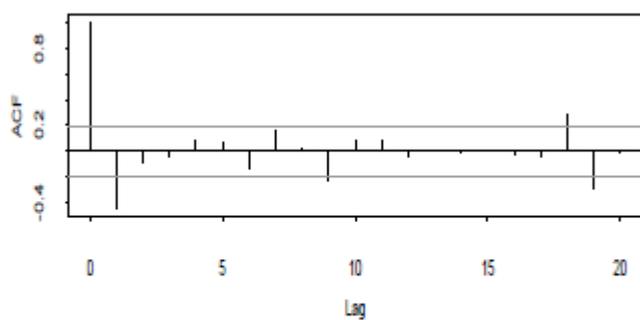
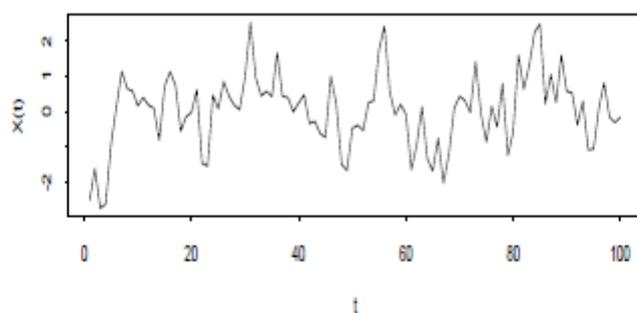


Figura 2.8: Processo MA(1) simulado, $\theta = 0,8$ e f.a.c.



Series : arma

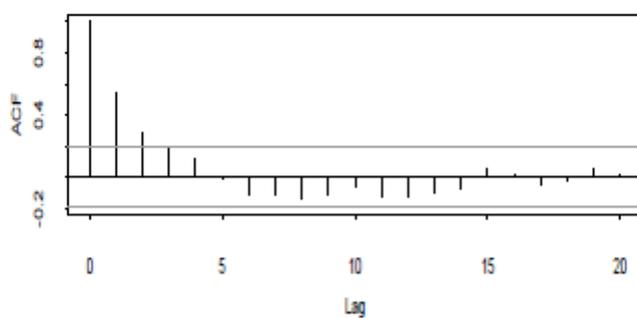


Figura 2.9: Processo ARMA(1,1) simulado, $\phi = 0,8, \theta = 0,3$ e f.a.c.

Tabela 3.1: Autocorrelações amostrais e respectivos desvios padrões e autocorrelações parciais amostrais para as séries simuladas AR(1), MA(1) e ARMA(1,1).

lag	AR(1)			MA(1)			ARMA(1)		
	r_j	$\hat{\sigma}(r_j)$	$\hat{\phi}_{jj}$	r_j	$\hat{\sigma}(r_j)$	$\hat{\phi}_{jj}$	r_j	$\hat{\sigma}(r_j)$	$\hat{\phi}_{jj}$
1	0,765*	0,10	0,765*	-0,352*	0,10	-0,352*	0,567*	0,10	0,567*
2	0,568*	0,15	-0,041	-0,127	0,11	-0,286*	0,370*	0,13	0,071
3	0,464*	0,17	0,103	0,036	0,11	-0,152	0,225	0,14	-0,015
4	0,344	0,18	-0,088	-0,063	0,11	-0,182	0,269	0,14	0,193
5	0,197	0,19	-0,124	-0,004	0,11	-0,153	0,161	0,15	-0,096
6	0,071	0,19	-0,082	-0,102	0,11	-0,006	0,056	0,15	-0,090
7	0,041	0,19	0,109	0,031	0,11	0,073	0,084	0,15	0,136
8	0,005	0,19	-0,039	-0,173	0,11	-0,125	0,123	0,15	0,034
9	-0,052	0,19	-0,042	0,051	0,12	-0,071	0,125	0,15	0,010
10	-0,084	0,19	-0,024	-0,026	0,12	-0,116	0,084	0,15	0,022
\bar{X}		0,532			-0,042			1,077	
S		2,462			1,284			1,312	

Na tabela também estão indicadas as médias e os desvios padrões amostrais de cada série, bem como destacados com um asterisco (*) os valores que caem fora do intervalo de dois desvios padrões.

Na Figura 3.1, temos os gráficos de r_j , $\hat{\phi}_{jj}$ para cada um dos modelos. Os intervalos de confiança para $\hat{\phi}_{jj}$ estão indicados em linha pontilhada.

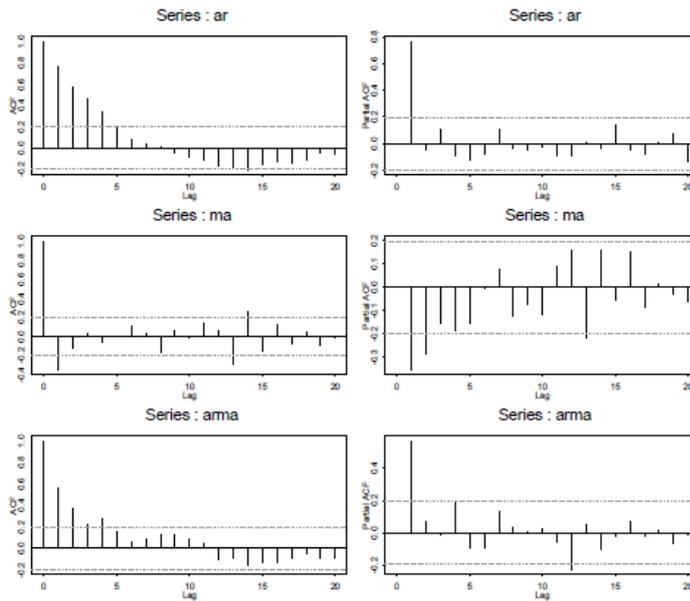


Figura 3.1: Autocorrelações e autocorrelações parciais amostrais para as séries simuladas AR(1), MA(1) e ARMA(1,1).

Note que os valores estimados podem diferir bastante dos valores verdadeiros. Por exemplo, no caso do modelo AR(1), $\phi = 0,8$ e $\rho_1 = 0,8$, $\rho_2 = 0,64$, etc. Para o modelo MA(1), $\theta = 0,8$ e $\rho_1 = -\theta/(1 + \theta^2) = -0,488$, sendo as demais nulas. Para o ARMA(1,1), $\rho_1 = 0,622$, $\rho_2 = 0,498$, etc. O valor de T aqui é 100. À medida que aumentarmos o tamanho da série, os valores estimados estarão mais próximos dos valores verdadeiros.

Procedimento de Identificação

O objetivo da identificação é determinar os valores de p , d e q do modelo $ARIMA(p, d, q)$. O procedimento de identificação consiste de três partes:

- (a) verificar se existe necessidade de uma transformação na série original, com o objetivo de estabilizar sua variância;
- (b) tomar diferenças da série, obtida no item (a), tantas vezes quantas necessárias para se obter uma série estacionária, de modo que o processo $\Delta^d X_t$ seja reduzido a um $ARMA(p, q)$. O número de diferenças, d , necessárias para que o processo se torne estacionário, é alcançado quando a f.a.c. amostral de $W_t = \Delta^d X_t$ decresce rapidamente para zero. Nesse estágio, a utilização de um teste para verificar a existência de raízes unitárias no polinômio autorregressivo, pode ser de grande utilidade. O teste de Dickey e Fuller (1979) será estudado no Capítulo 4;
- (c) identificar o processo $ARMA(p, q)$ resultante por meio da análise das autocorrelações e autocorrelações parciais estimadas, cujos comportamentos devem imitar os comportamentos das respectivas quantidades teóricas. Esses comportamentos foram abordados no Capítulo 2 e um resumo das propriedades destes modelos encontra-se na Tabela 3.2, extraída de Box et al. (1994).

Tabela 3.2: Comportamento das f.a.c. e f.a.c.p. de um processo ARIMA(p, d, q)

Ordem	(1, $d, 0$)	(0, $d, 1$)
comportamento de ρ_k	decai expon.	somente $\rho_1 \neq 0$
comportamento de ϕ_{kk}	somente $\phi_{11} \neq 0$	decaim. exp. dom.
estimativas iniciais	$\phi = \rho_1$	$\rho_1 = -\frac{\theta}{1+\theta^2}$
região de admissibilidade	$-1 < \phi < 1$	$-1 < \theta < 1$
Ordem	(2, $d, 0$)	(0, $d, 2$)
comportamento de ρ_k	mistura de expon. ou senoides amortecidas	só $\rho_1 \neq 0$ e $\rho_2 \neq 0$
comportamento de ϕ_{kk}	só $\phi_{11} \neq 0$ e $\phi_{22} \neq 0$	dom. por mist. exp. ou senoides amort.
estimativas iniciais	$\begin{cases} \phi_1 = \frac{\rho_1(1-\rho_2)}{1-\rho_1^2}, \\ \phi_2 = \frac{\rho_2-\rho_1^2}{1-\rho_1^2} \end{cases}$	$\begin{cases} \rho_1 = \frac{-\theta_1(1-\theta_2)}{1+\theta_1^2+\theta_2^2}, \\ \rho_2 = \frac{-\theta_2}{1+\theta_1^2+\theta_2^2} \end{cases}$
região de admissibilidade	$\begin{cases} -1 < \phi_2 < 1 \\ \phi_2 - \phi_1 < 1 \\ \phi_2 + \phi_1 < 1 \end{cases}$	$\begin{cases} -1 < \theta_2 < 1 \\ \theta_2 - \theta_1 < 1 \\ \theta_2 + \theta_1 < 1 \end{cases}$
Ordem	(1, $d, 1$)	
comportamento de ρ_k	decai exponencialmente após o lag 1	
comportamento de ϕ_{kk}	dominada por decaimento expon. após o lag 1	
estimativas iniciais	$\rho_1 = \frac{(1-\phi\theta)(\phi-\theta)}{1+\theta^2-2\phi\theta}$, $\rho_2 = \rho_1\phi$	
região de admissibilidade	$-1 < \phi < 1$, $-1 < \theta < 1$	

Tabela 3.3: Variâncias aproximadas para \bar{W} , onde $W_t = \Delta^d X_t$, $n = T - d$

AR(1)	MA(1)	ARMA(1,1)
$\frac{c_0(1+r_1)}{n(1-r_1)}$	$\frac{c_0(1+2r_1)}{n}$	$\frac{c_0}{n} \left[1 + \frac{2r_1^2}{r_1-r_2} \right]$
AR(2)		MA(2)
$\frac{c_0(1+r_1)(1-2r_1^2+r_2)}{n(1-r_1)(1-r_2)}$		$\frac{c_0(1+2r_1+2r_2)}{n}$

Exemplo 3.3. Suponha que temos os seguintes dados:

j	1	2	3	4	5	6	7	8
r_j	0,81	0,69	0,58	0,44	0,30	0,26	0,19	0,15
$\hat{\phi}_{jj}$	0,81	0,11	-0,03	-0,12	-0,13	0,17	-0,01	0,02

$$T = 50, \quad \bar{X} = 0,5327, \quad S^2 = 6,0579.$$

Temos que

$$\hat{\sigma}(\hat{\phi}_{jj}) \simeq \frac{1}{\sqrt{T}} = \frac{1}{\sqrt{50}} \simeq 0,14, \quad \text{logo } 2\hat{\sigma}(\hat{\phi}_{jj}) = 0,28,$$

o que mostra que só $\phi_{11} \neq 0$; os valores de $\{r_j\}$ sugerem um decaimento e, portanto, $\{\hat{\phi}_{jj}\}$ sugere um processo AR(1).

Para um processo AR(1), usando a Tabela 3.3, temos que

$$\text{Var}(\bar{X}) = \frac{c_0(1+r_1)}{n(1-r_1)} = \frac{6,0579(1+0,81)}{50(1-0,81)} = 1,1542$$

e $\hat{\sigma}(\bar{X}) = 1,0743$; como $\bar{X} = 0,5327$, a média pode ser considerada igual a zero e o modelo sugerido é

$$X_t = \phi X_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim \text{RB}(0, \sigma_\varepsilon^2).$$

Exemplo 3.4. Vamos agora identificar um ou mais modelos preliminares para a série ICV (arquivo m-icv70.79.dat), no período de janeiro de 1970 a junho de 1979, utilizando $T = 114$ observações.

Inicialmente, vamos considerar a transformação logarítmica

$$Y_t = \ln(ICV), \quad t = 1, \dots, 114.$$

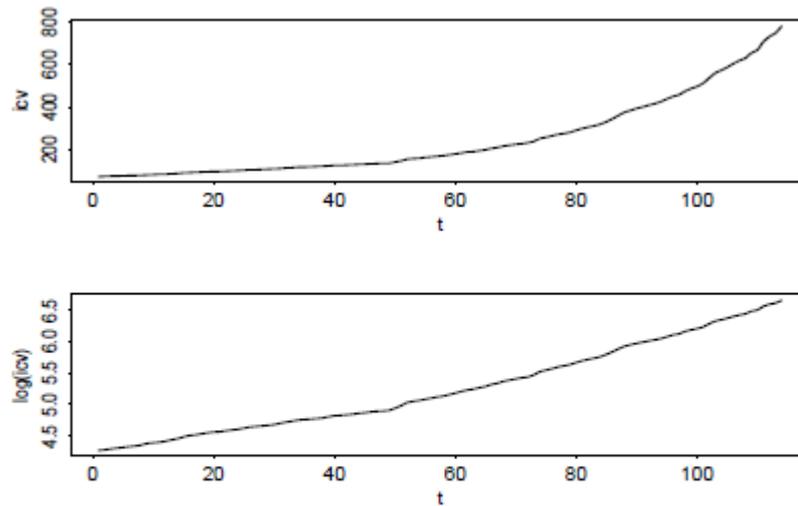


Figura 3.2: Gráfico das séries ICV e $\ln(ICV)$

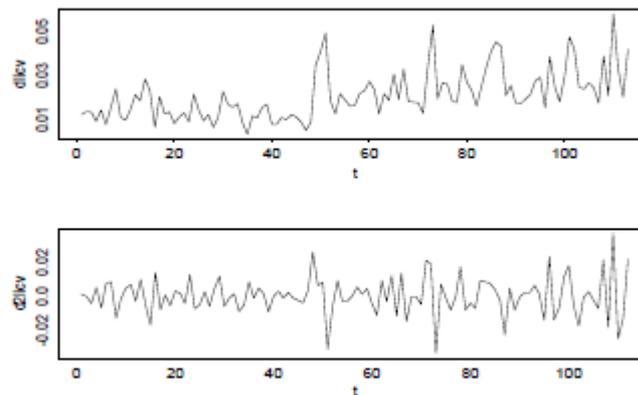


Figura 3.4: Gráfico das séries $(1 - B) \ln(ICV)$ e $(1 - B)^2 \ln(ICV)$

Tabela 3.4: autocorrelações e autocorrelações parciais estimadas das séries (a) $\ln(\text{ICV})$, (b) $(1 - B)\ln(\text{ICV})$, (c) $(1 - B)^2\ln(\text{ICV})$.

lag	(a)		(b)		(c)	
	f.a.c.	f.a.c.p.	f.a.c.	f.a.c.p.	f.a.c.	f.a.c.p.
1	0,97	0,97	0,49	0,49	-0,32	-0,32
2	0,94	-0,01	0,33	0,12	-0,12	-0,24
3	0,92	-0,02	0,26	0,08	0,00	-0,15
4	0,89	-0,01	0,16	-0,03	-0,10	-0,23
5	0,86	0,00	0,21	0,14	-0,04	-0,24
6	0,84	-0,02	0,27	0,15	0,08	-0,14
7	0,81	-0,01	0,29	0,11	0,00	-0,13
8	0,78	-0,02	0,29	0,08	0,05	-0,06
9	0,76	-0,01	0,25	0,02	-0,02	-0,07
10	0,73	-0,02	0,22	0,04	-0,02	-0,06
11	0,70	-0,02	0,21	0,05	-0,09	-0,17
12	0,68	-0,02	0,28	0,14	0,03	-0,14
13	0,65	0,00	0,30	0,08	0,09	-0,03
14	0,62	0,00	0,26	0,01	0,01	-0,01
15	0,60	-0,01	0,23	0,00	-0,10	-0,14
16	0,57	-0,01	0,27	0,12	0,14	0,06
17	0,55	-0,02	0,16	-0,09	-0,12	-0,05
18	0,53	-0,01	0,19	0,05	0,04	0,05
19	0,50	-0,02	0,17	-0,06	0,06	0,11
20	0,48	-0,01	0,08	-0,11	-0,19	-0,14

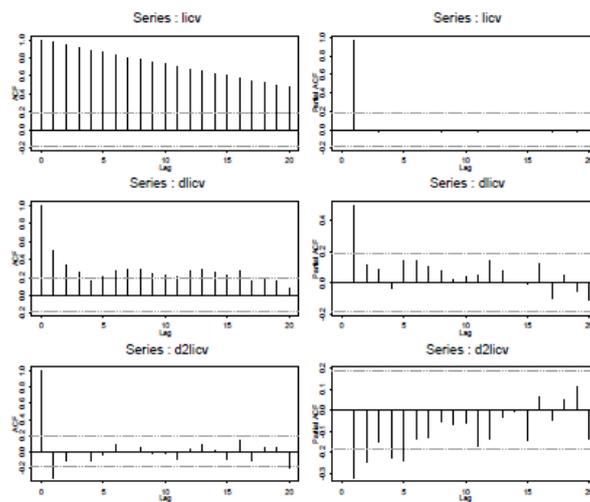


Figura 3.3: (a) Autocorrelações e (b) autocorrelações parciais das séries $\ln(\text{ICV})$, $(1 - B)\ln(\text{ICV})$ e $(1 - B)^2\ln(\text{ICV})$

Observando a f.a.c. amostral de $(1 - B)^2 Y_t$, verificamos que $r_1 = -0,32$, que além de ser negativo, assume um valor próximo de $-0,5$. Calculando as variâncias das diferenças das séries, temos que $\text{Var}((1 - B)Y_t) = 0,00013$ e $\text{Var}((1 - B)^2 Y_t) = 0,000162$ indicando um aumento de cerca de 25% da variância da série com duas diferenças em relação àquela com uma diferença. Assim, escolhemos $d = 1$.

Analisando o comportamento das f.a.c. e f.a.c.p da série $(1 - B)\ln(\text{ICV})$, Tabela 3.4 e Figura 3.3, podemos notar que somente $\phi_{11} \neq 0$, pois $\hat{\phi}_{11} = 0,49$ e $\hat{\sigma}(\hat{\phi}_{11}) = 1/\sqrt{114} = 0,094$. Assim, um modelo preliminar para $\ln(\text{ICV})$ é um ARIMA(1,1,0).

Para verificar se uma constante deve, ou não, ser adicionada ao modelo, observamos que

$$\begin{aligned}\bar{W} &= 0,02116, \\ \text{Var}(\bar{W}) &= \frac{c_0(1 + r_1)}{n(1 - r_1)} = \frac{0,00014(1 + 0,49)}{113(1 - 0,49)} = 3,62 \times 10^{-6},\end{aligned}$$

ou seja, $\hat{\sigma}(\bar{W}) = 0,0019$. Assim, podemos concluir que a média da série $(1 - B)Y_t$ é diferente de zero e que uma constante deve ser incluída no modelo.

Logo, o modelo preliminar para a série $\ln(\text{ICV})$ é um ARIMA(1,1,0) com uma constante, isto é,

$$(1 - B)(1 - \phi_1 B)Y_t = \theta_0 + \varepsilon_t, \quad (3.16)$$

onde $Y_t = \ln(\text{ICV})$ e $\varepsilon_t \sim \text{RB}(0, \sigma^2)$.

Exemplo 3.5. Considere a série de retornos mensais do Ibovespa, de junho de 1994 a agosto de 2001 (arquivo m-ibv94.01.dat), com $T = 86$ observações. Se denotarmos a série de índices por X_t , então os retornos são dados por

$$r_t = \log \left(\frac{X_t}{X_{t-1}} \right) = \log X_t - \log X_{t-1} = \Delta(\log X_t).$$

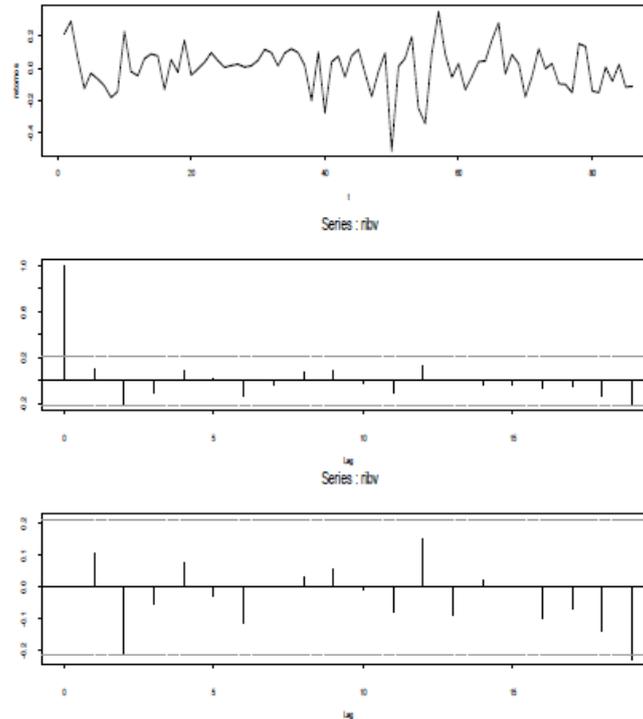


Figura 3.5: Retornos Ibovespa (a) gráfico da série; (b) f.a.c. amostral; (c) f.a.c.p. amostral

O gráfico da série, as f.a.c. e f.a.c.p. amostrais estão na Figura 3.5, mostrando que essa série pode ser considerada um ruído branco. Como a média amostral é 0,003872 e o desvio padrão amostral é 0,137428, podemos concluir que um modelo adequado para descrever esses retornos é

$$r_t = \Delta(\log X_t) = \varepsilon_t, \quad \text{com } \varepsilon_t \sim \text{RB}(0, \sigma^2),$$

e $\hat{\sigma}^2 = 0,01889$, ou ainda,

$$\log X_t = \log X_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim \text{RB}(0, \sigma^2),$$

ou seja, os logaritmos dos índices do Ibovespa seguem um passeio casual com média zero (sem *drift*).

Como veremos mais adiante, esta é, em geral, uma característica de retornos de ativos financeiros: são não correlacionados. Mas podem não ser independentes, o que se pode verificar considerando-se os quadrados dos retornos.

Formas alternativas de identificação

A. Critério de Informação AIC

Akaike (1973, 1974) sugere escolher o modelo cujas ordens k e l minimizam o critério

$$AIC(k, l) = \ln \hat{\sigma}_{k,l}^2 + \frac{2(k+l)}{T} \quad (3.18)$$

onde $\hat{\sigma}_{k,l}^2$ é o estimador de máxima verossimilhança de σ^2 para um modelo ARMA(k, l) (veja a seção 3.3).

O que se faz, então, é estipular valores limites superiores K e L para k e l e calcular (3.18) para todas as possíveis combinações (k, l) com $0 \leq k \leq K$ e $0 \leq l \leq L$. Em geral, K e L são funções de T , por exemplo, $K = L = \ln T$.

Para o caso de modelos AR(p), o critério AIC reduz-se a

$$AIC(k) = \ln \hat{\sigma}_k^2 + \frac{2k}{T}, \quad k \leq K. \quad (3.19)$$

Existem várias correções para melhorar o comportamento do AIC, no sentido de diminuir a probabilidade de selecionar uma ordem maior do que a verdadeira.

Hurvich e Tsai (1989) propõem uma correção para o AIC, que no caso de um processo AR é dada por

$$AIC_c(k) = AIC(k) + \frac{2(k+1)(k+2)}{T-k+2}, \quad k \leq K$$

e utilizando simulações, mostram que essa correção é útil quando T é pequeno ou quando K é uma fração “moderadamente grande” de T .

B. Critério de Informação Bayesiano BIC

Akaike (1977), Rissanen (1978) e Schwarz (1978) sugerem minimizar o Critério de Informação Bayesiano, dado por

$$BIC = -2 \ln(FV) + (\text{número de parâmetros}) \ln T,$$

na qual FV é a função de verossimilhança.

No caso de um processo ARMA, essa quantidade é dada por

$$BIC(k, l) = \ln \hat{\sigma}_{k,l}^2 + (k+l) \frac{\ln T}{T}. \quad (3.20)$$

Hannan (1980) mostra que, sob determinadas condições, as estimativas k e l que minimizam a expressão (3.20) são fortemente consistentes.

Exemplo 3.6. Na Figura 3.6, temos o gráfico da série de retornos r_t diários da Petrobras, de 18/08/1998 a 29/09/2010, com $T = 2998$ observações, suas f.a.c. e f.a.c.p. amostrais; essas indicam que a série é autocorrelacionada e que um modelo autorregressivo é apropriado. Provavelmente, um modelo AR(3) é suficiente, mas vamos fixar uma ordem máxima $K = 15$. Na Tabela 3.5, temos alguns valores de AIC, BIC e do logaritmo da verossimilhança para cada modelo estimado. Como os critérios escolhem ordens diferentes, vamos nos fixar no AIC, e ajustar um AR(15). Estimando este modelo, somente os coeficientes $\phi_1, \phi_2, \phi_3, \phi_{10}$ e ϕ_{15} são significativos (veja a seção 3.3), ou seja, o modelo final é

$$r_t = \phi_1 r_{t-1} + \phi_2 r_{t-2} + \phi_3 r_{t-3} + \phi_{10} r_{t-10} + \phi_{15} r_{t-15} + \varepsilon_t.$$

Lembremos que o retorno é definido como a primeira diferença do logaritmo da série original de preços das ações.

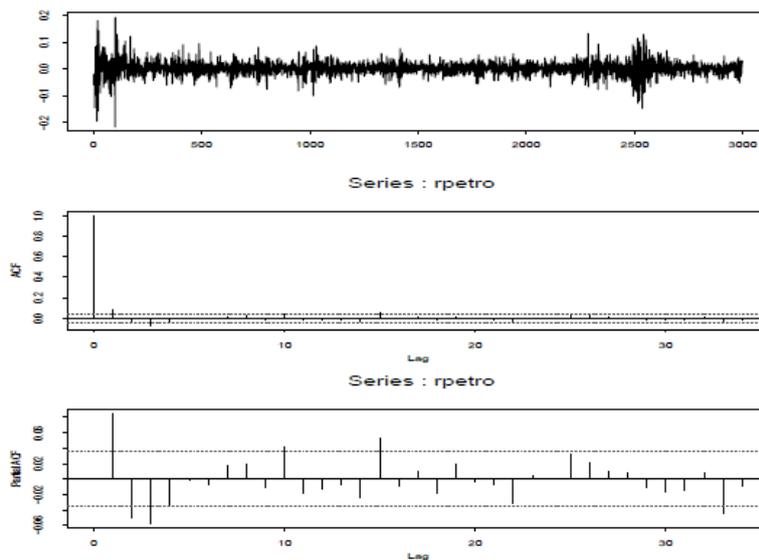


Figura 3.6: Retornos Petrobras (a) gráfico da série (b) f.a.c. amostral (c) f.a.c.p. amostral

Diagnóstico

Após estimar o modelo temos que verificar se ele representa, ou não, adequadamente, os dados. Qualquer insuficiência revelada pode sugerir um modelo alternativo como adequado.

A verificação pode ser feita analisando os resíduos. Suponha que o modelo proposto seja

$$\phi(B)W_t = \theta(B)\varepsilon_t,$$

com

$$W_t = \Delta^d X_t.$$

Se esse modelo for verdadeiro, então os “erros verdadeiros” ε_t constituirão um ruído branco.

3.4.1 Teste de autocorrelação residual

Estimados ϕ e θ , as quantidades

$$\hat{\varepsilon}_t = \hat{\theta}^{-1}(B)\hat{\phi}(B)W_t$$

são chamadas *resíduos estimados* ou simplesmente *resíduos*. Se o modelo for adequado, os $\hat{\varepsilon}_t$ deverão estar próximos dos ε_t e, portanto, deverão ser aproximadamente não correlacionados. Se indicarmos por \hat{r}_k as autocorrelações dos resíduos $\hat{\varepsilon}_t$, então deveríamos ter $\hat{r}_k \simeq 0$. Em particular, deveríamos ter, aproximadamente,

$$\hat{r}_k \sim \mathcal{N}\left(0, \frac{1}{n}\right),$$

onde $n = T - d$ é o número efetivo de observações e sempre sob a suposição que o modelo ajustado seja apropriado. As autocorrelações \hat{r}_k são calculadas por

$$\hat{r}_k = \frac{\sum_{t=k+1}^n \hat{\varepsilon}_t \hat{\varepsilon}_{t-k}}{\sum_{t=1}^n \hat{\varepsilon}_t^2}.$$

Contudo, o desvio padrão de \hat{r}_k pode ser consideravelmente menor que $1/\sqrt{n}$, especialmente para pequenos valores de k , como mostrou Durbin (1970). Ele provou que para um AR(1), $\text{Var}(\hat{r}_k) \simeq \phi^2/n$, que pode ser bem menor que $1/n$. Box et al. (1994) provaram que, para um modelo AR(1), tem-se

$$\begin{aligned}\text{Var}(\hat{r}_k) &\simeq \frac{1}{n}[1 - \phi^{2(k-1)}(1 - \phi^2)] \\ \text{Cov}(\hat{r}_i, \hat{r}_j) &\simeq \frac{1}{n}\{\delta_{ij} - \phi^{i+j-2}(1 - \phi^2)\},\end{aligned}$$

onde δ_{ij} é o delta de Kronecker. Daqui, temos que, para k grande ou moderado, a variância de \hat{r}_k é, aproximadamente, $1/n$, e as autocorrelações estimadas dos resíduos são não correlacionadas.

De qualquer modo, a comparação de \hat{r}_k com os limites $\pm 2/\sqrt{n}$ fornece uma indicação geral de possível quebra de comportamento de ruído branco em ε_t , com a condição de que seja lembrado que, para pequenos valores de k , estes limites subestimarão a significância de qualquer discrepância.

3.4.2 Teste de Box-Pierce-Ljung

Box e Pierce (1970) sugeriram um teste para as autocorrelações dos resíduos estimados, que, apesar de não detectar quebras específicas no comportamento de ruído branco, pode indicar se esses valores são muito altos. Uma modificação deste teste foi proposta por Ljung e Box (1978), que é apresentada a seguir.

Se o modelo for apropriado, a estatística

$$Q(K) = n(n+2) \sum_{k=1}^K \frac{\hat{r}_k^2}{(n-k)} \quad (3.26)$$

terá uma distribuição χ^2 com $K - p - q$ graus de liberdade. A hipótese de ruído branco para os resíduos é rejeitada para valores grandes de $Q(\cdot)$. Em geral, basta utilizar as 15 ou 20 primeiras \hat{r}_k .

Exemplo 3.11. Vamos, agora, verificar se o modelo ARIMA(1,1,0) com θ_0 , proposto para a série $\ln(ICV)$, é adequado (veja os exemplos 3.4 e 3.9). A Figura

3.7 apresenta as f.a.c. e f.a.c.p dos resíduos com os correspondentes intervalos de confiança. Os valores das autocorrelações e autocorrelações parciais não indicam nenhuma quebra de comportamento de ruído branco nos resíduos.

Assim, o modelo ajustado à série $\ln(ICV)$ é dada por

$$(1 - 0,5119B)(1 - B)\ln(ICV) = 0,01036 + \varepsilon_t, \quad (3.27)$$

com $\hat{\sigma}^2 = 0,0000923$.

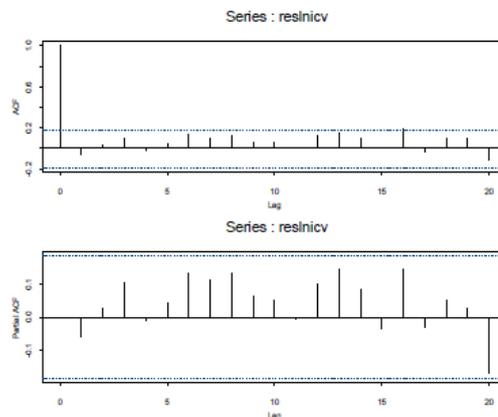


Figura 3.7: F.a.c. e f.a.c.p. dos resíduos do modelo ARIMA(1,1,0) ajustado à série $\ln(ICV)$.

Exemplo 3.12. Retomemos o exemplo 3.10, no qual obtivemos as estimativas dos parâmetros do modelo AR(15) ajustado à série de retornos da Petrobras. Uma estimativa da variância residual é $\hat{\sigma}^2 = 0,00069$. Na Figura 3.8, temos os resíduos, f.a.c. e f.a.c.p. amostrais dos resíduos, mostrando que podemos aceitar a hipótese de que estes sejam um ruído branco. No Capítulo 5, veremos que essa série apresenta, na realidade, uma variância condicional não constante, de modo que será necessário ajustar um modelo da classe ARCH-GARCH para levar esse fato em conta. O gráfico da série, Figura 3.6 (a) mostra grupos de retornos com variabilidade (volatilidade) grande e outros com variabilidade menor.

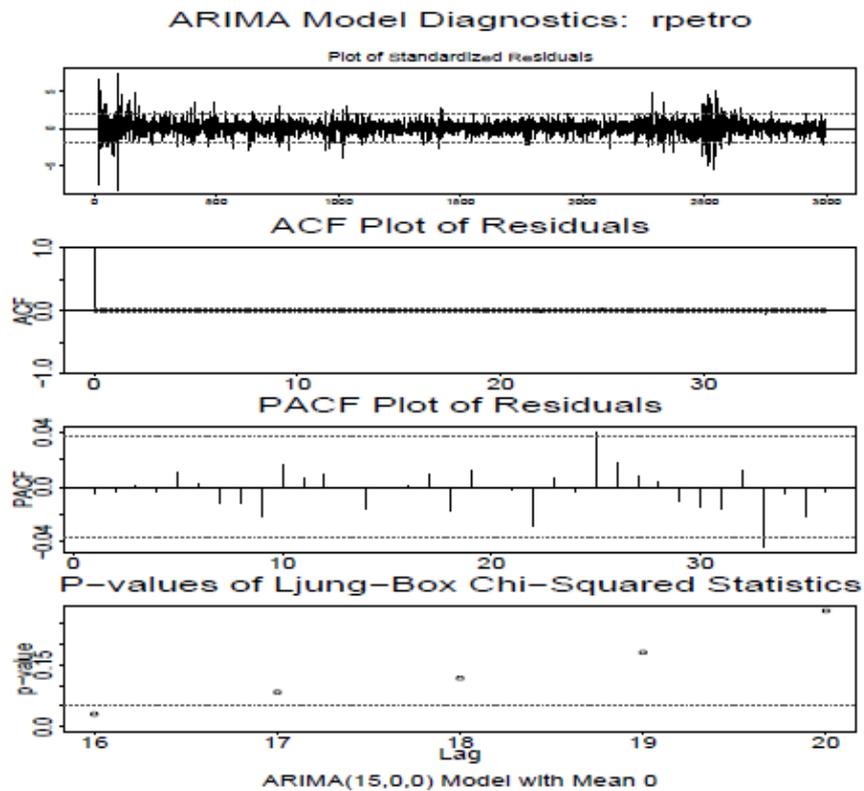
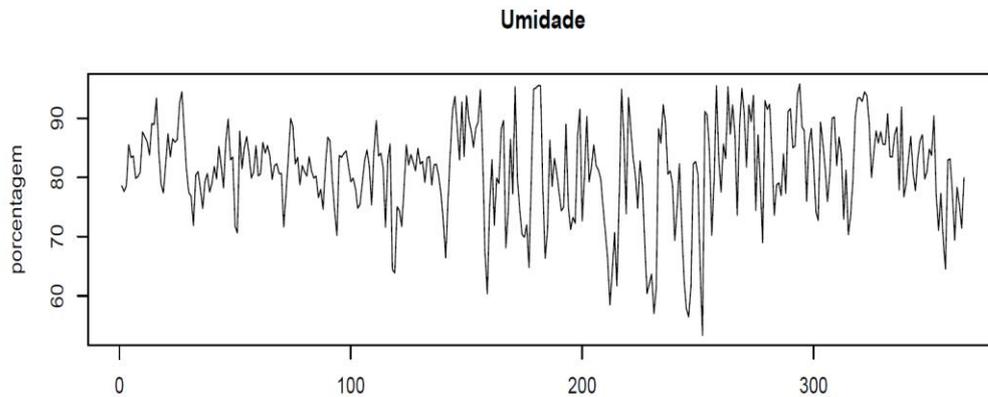


Figura 3.8: Diagnóstico do modelo do exemplo 3.10.

Exemplo de Identificação

Exemplo 6.4. Considere a série Atmosfera–Umidade (simplesmente, Umidade), de 1^o de janeiro a 24 de dezembro de 1997, com $N = 358$ observações.



A série foi dividida em grupos de 7 observações consecutivas (uma semana completa), calculando-se para cada grupo a média e a amplitude (ou desvio padrão), cuja representação gráfica aparece na Figura 6.6. Na Tabela 6.5, apresentamos os valores da estatística d_λ , dada por (1.8), utilizando as transformações usuais. A análise da Figura 6.6 sugere que a amplitude e, também, o desvio padrão não dependem da média da série original, indicando a não necessidade de transformação para estabilizar a variância. Os resultados da Tabela 6.6 indicam que o mínimo de d_λ ocorre para $\lambda = 1$, significando que os dados originais são simétricos e, provavelmente, seguem uma distribuição normal. Assim, trabalhamos com a série original, isto é, $Z_t = (\text{Umidade})_t$.

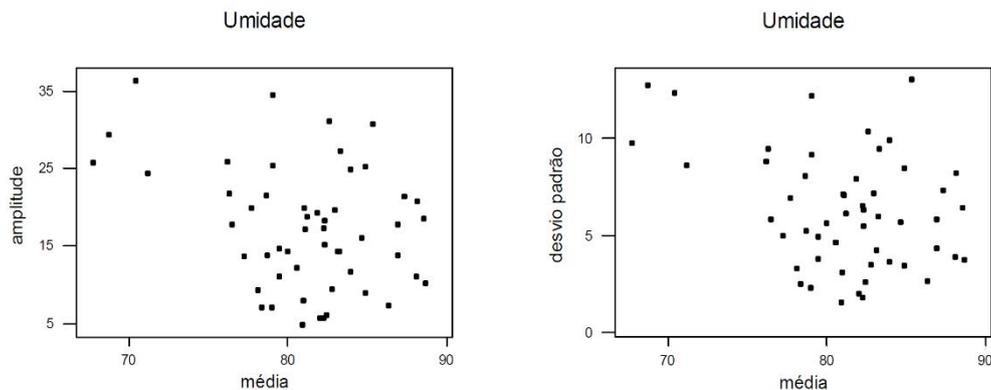


Figura 6.6: Gráficos da série Umidade: (a) amplitude \times média e (b) desvio padrão \times média

Tabela 6.5: Valores da estatística d_λ , dada por (1.8).

λ	d_λ	Transformação
-1,0	0,187	$1/Z_t$
-0,5	0,168	$1/(Z_t)^{1/2}$
0	0,147	$\ln(Z_t)$
0,5	0,125	$(Z_t)^{1/2}$
1,0	0,103	Z_t

$$Z_t^{(\lambda)} = \begin{cases} \frac{e^{\lambda Z_t} - 1}{\lambda}, & \text{se } \lambda \neq 0, \\ Z_t, & \text{se } \lambda = 0, \end{cases} \quad (1.8)$$

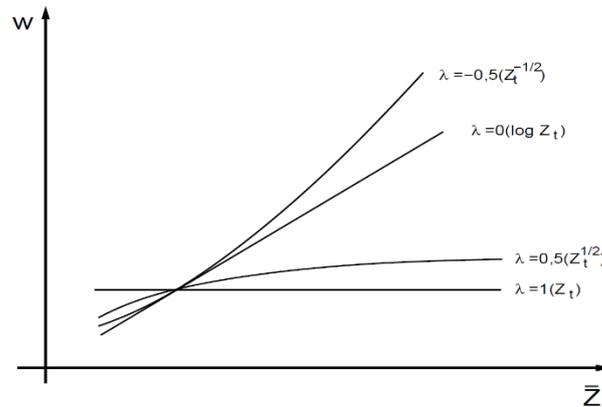


Figura 1.6: Gráficos amplitude \times média, ilustrando alguns valores possíveis de λ .

Hinkley (1977) sugere que se calcule a média, mediana e um estimador de escala (desvio padrão ou algum estimador robusto) e, então, para a transformação $Z_t^{(\lambda)}$ e λ tomando valores $\dots, -2, -1, -1/2, -1/4, 0, 1/4, 1/2, 1, 2, \dots$, escolha-se o valor de λ que minimize

$$d_\lambda = \frac{|\text{média} - \text{mediana}|}{\text{medida de escala}}, \quad (1.10)$$

que pode ser vista como uma medida de assimetria; numa distribuição simétrica, $d_\lambda = 0$.

A Tabela 6.6 e a Figura 6.7 apresentam os valores e as correspondentes representações gráficas das funções de autocorrelação e autocorrelação parcial da série Z_t e de suas diferenças de ordem 1 e 2.

Tabela 6.6: Autocorrelações e autocorrelações parciais estimadas das séries (a) Umidade, (b) $(1 - B)Umidade$, (c) $(1 - B)^2Umidade$.

lag	(a)		(b)		(c)	
	acf	pacf	acf	pacf	acf	pacf
1	0,53	0,53	-0,16	-0,16	-0,45	-0,45
2	0,21	-0,10	-0,27	-0,30	-0,17	-0,45
3	0,14	0,10	0,00	-0,12	0,14	-0,27
4	0,08	-0,04	-0,07	-0,20	-0,08	-0,33
5	0,08	0,07	0,07	-0,03	0,10	-0,17
6	0,01	-0,09	-0,01	-0,10	-0,01	-0,12
7	-0,05	-0,03	-0,09	-0,13	-0,07	-0,12
8	-0,02	0,02	0,00	-0,11	0,04	-0,10
9	0,00	0,01	0,00	-0,11	-0,02	-0,14
10	0,03	0,03	0,05	-0,05	0,05	-0,07
11	0,01	-0,03	0,00	-0,08	0,00	-0,03
12	-0,01	0,01	-0,06	-0,11	-0,06	-0,05
13	0,03	0,05	0,01	-0,09	0,09	0,08
14	0,07	0,03	-0,12	-0,26	-0,17	-0,17
15	0,21	0,23	0,15	-0,03	0,15	-0,06
16	0,22	0,00	0,08	-0,07	0,02	-0,04
17	0,16	0,04	-0,06	-0,05	-0,09	-0,01
18	0,15	0,03	0,01	-0,06	0,02	-0,05
19	0,14	0,03	0,04	0,02	0,04	0,04
20	0,08	-0,04	-0,01	-0,02	-0,04	0,00

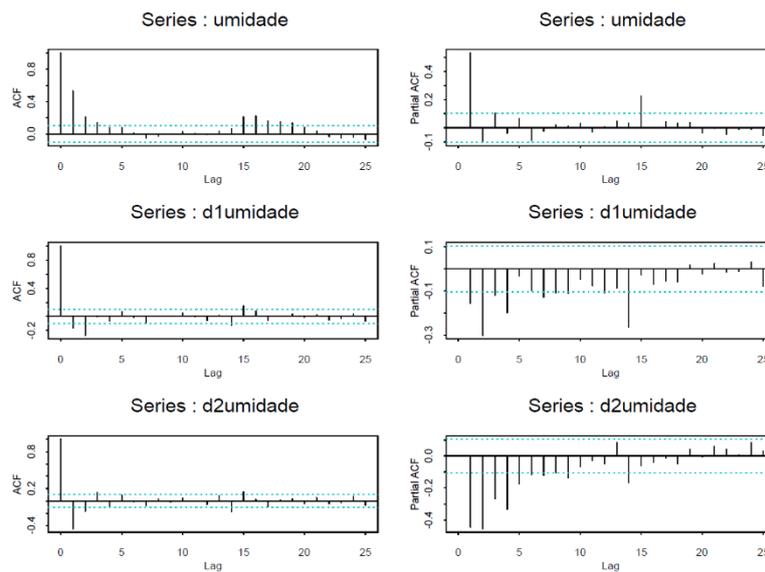


Figura 6.7: Autocorrelações e autocorrelações parciais estimadas das séries Umidade, $(1 - B)Umidade$ e $(1 - B)^2Umidade$.

O comportamento da fac de Z_t não indica, aparentemente, a necessidade de diferenças na série original, com o objetivo de torná-la estacionária.

A aplicação do teste de Dickey-Fuller (Apêndice B) foi feita com $p = 1$ (sugerido pelo comportamento da fac de Z_t), fornecendo o modelo

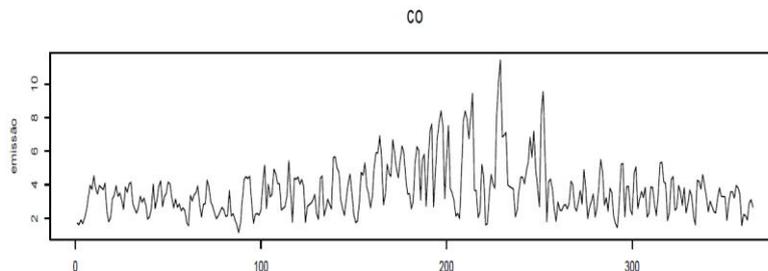
$$\Delta Z_t = \phi_0^* + \phi_1^* Z_{t-1} + a_t$$

com $\hat{\phi}_1^* = -0,4680$, $\hat{\sigma}(\hat{\phi}_1^*) = 0,045$ e, conseqüentemente, $\hat{\tau}_u = -10,4156$. O valor crítico de $\hat{\tau}_u$, fornecido pelo R, com o nível de 1%, é $-3,44$ indicando que a hipótese de raiz unitária é rejeitada com o nível de 1%.

Pelo que foi mencionado anteriormente e de acordo com o comportamento das funções de autocorrelação e autocorrelação parcial da série Z_t , um modelo preliminar para a série de Umidade é um processo AR(1) com uma constante, isto é,

$$(1 - \phi_1 B)Z_t = a_t + \theta_0. \quad (6.15)$$

Exemplo 6.5. Vejamos, agora, a série Poluição-CO (simplesmente, CO), de 1^o de janeiro a 24 de dezembro de 1997, com $N = 358$ observações.



Esta série de dados diários também foi dividida em grupos de uma semana (sete dias) para os quais calculamos a média e a amplitude (ou desvio padrão); os resultados são apresentados na Figura 6.9. Na Tabela 6.7, temos os valores da estatística d_λ , dada por (1.10). A análise da Figura 6.9 sugere um aumento, das medidas de variabilidade, proporcional à média, indicando a necessidade de uma transformação logarítmica para estabilizar a variabilidade da série. Esta transformação também é indicada para tornar os dados mais simétricos e, possivelmente, com distribuição mais próxima de uma normal; tal indicação pode ser visualizada nos resultados da Tabela 6.7, que apresenta um valor mínimo para d_λ , quando $\lambda = 0$.

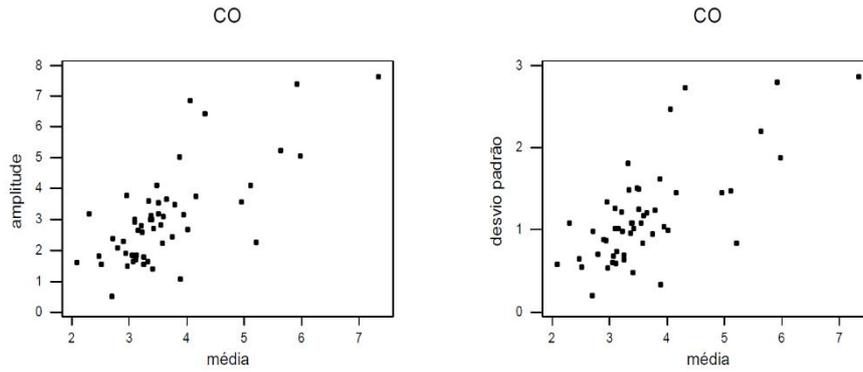


Figura 6.9: Gráficos da série CO: (a) amplitude \times média, (b) $dp \times$ amplitude.

Tabela 6.7: Valores da estatística d_λ , dada por (1.8).

λ	d_λ	Transformação
-1,0	0,1856	$1/Z_t$
-0,5	0,096	$1/(Z_t)^{1/2}$
0	0,002	$\ln(Z_t)$
0,5	0,098	$(Z_t)^{1/2}$
1,0	0,180	Z_t

$$Z_t = \ln(\text{CO})_t. \quad (6.16)$$

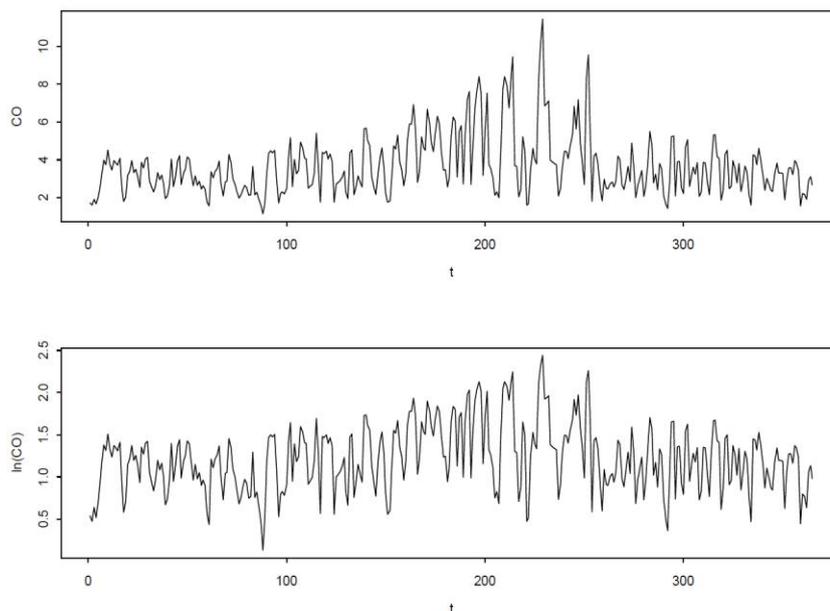


Figura 6.10: Gráficos das séries CO e $\ln(\text{CO})$.

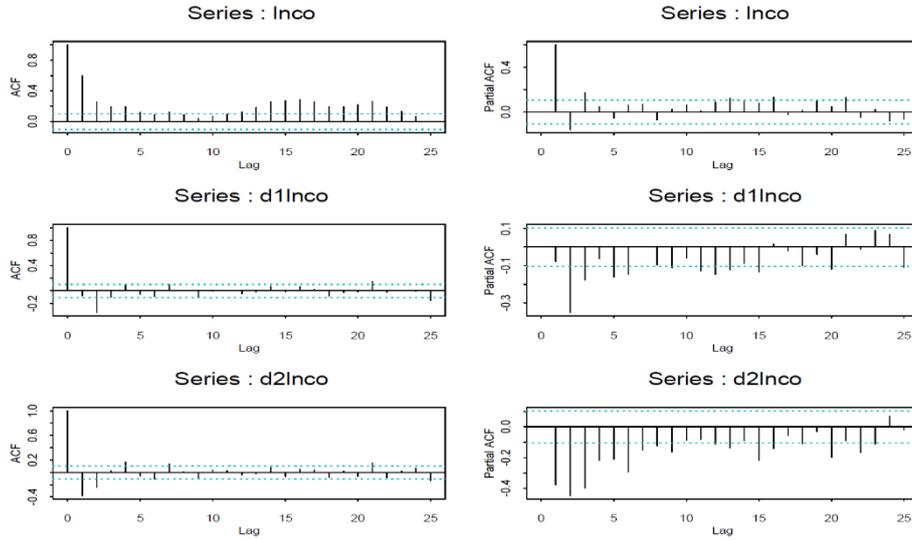


Figura 6.11: Autocorrelações e autocorrelações parciais estimadas das séries $\ln(\text{CO})$, $(1 - B)\ln(\text{CO})$ e $(1 - B)^2\ln(\text{CO})$.

Para confirmar a suspeita da não necessidade de tomar diferenças, fizemos o teste da raiz unitária, utilizando $p = 3$, pois temos as três primeiras autocorrelações parciais significativamente diferentes de zero, como podemos observar na Figura 6.11.

O modelo ajustado, de acordo com (B.27), foi

$$\Delta Z_t = \phi_0^* + \phi_1^* Z_{t-1} + \phi_2^* \Delta Z_{t-1} + \phi_3^* \Delta Z_{t-2} + a_t,$$

com

$$\begin{aligned} \hat{\phi}_0^* &= 0,4815(0,06628); \\ \hat{\phi}_1^* &= -0,3897(0,05228); \\ \hat{\phi}_2^* &= 0,1019(0,05404); \\ \hat{\phi}_3^* &= -0,1734(0,05223), \end{aligned}$$

implicando em $\hat{\tau}_u = (-0,38973)/(0,05228) = -7,454$.

O valor crítico, para a estatística $\hat{\tau}_u$, com o nível de 1%, é $-3,44$, o que nos leva a rejeitar a hipótese de existência de raiz unitária.

Assim, um modelo preliminar para a série CO é um processo AR(3), dado por

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \phi_3 Z_{t-3} + \theta_0 + a_t, \quad (6.17)$$

onde $Z_t = \ln(\text{CO})_t$.

Exemplo de diagnóstico

3.4.1 Teste de autocorrelação residual

Estimados ϕ e θ , as quantidades

$$\hat{\varepsilon}_t = \hat{\theta}^{-1}(B)\hat{\phi}(B)W_t$$

são chamadas *resíduos estimados* ou simplesmente *resíduos*. Se o modelo for adequado, os $\hat{\varepsilon}_t$ deverão estar próximos dos ε_t e, portanto, deverão ser aproximadamente não correlacionados. Se indicarmos por \hat{r}_k as autocorrelações dos resíduos $\hat{\varepsilon}_t$, então deveríamos ter $\hat{r}_k \simeq 0$. Em particular, deveríamos ter, aproximadamente,

$$\hat{r}_k \sim \mathcal{N}\left(0, \frac{1}{n}\right),$$

onde $n = T - d$ é o número efetivo de observações e sempre sob a suposição que o modelo ajustado seja apropriado. As autocorrelações \hat{r}_k são calculadas por

$$\hat{r}_k = \frac{\sum_{t=k+1}^n \hat{\varepsilon}_t \hat{\varepsilon}_{t-k}}{\sum_{t=1}^n \hat{\varepsilon}_t^2}.$$

Contudo, o desvio padrão de \hat{r}_k pode ser consideravelmente menor que $1/\sqrt{n}$, especialmente para pequenos valores de k , como mostrou Durbin (1970). Ele provou que para um AR(1), $\text{Var}(\hat{r}_k) \simeq \phi^2/n$, que pode ser bem menor que $1/n$. Box et al. (1994) provaram que, para um modelo AR(1), tem-se

$$\begin{aligned}\text{Var}(\hat{r}_k) &\simeq \frac{1}{n}[1 - \phi^{2(k-1)}(1 - \phi^2)] \\ \text{Cov}(\hat{r}_i, \hat{r}_j) &\simeq \frac{1}{n}\{\delta_{ij} - \phi^{i+j-2}(1 - \phi^2)\},\end{aligned}$$

onde δ_{ij} é o delta de Kronecker. Daqui, temos que, para k grande ou moderado, a variância de \hat{r}_k é, aproximadamente, $1/n$, e as autocorrelações estimadas dos resíduos são não correlacionadas.

De qualquer modo, a comparação de \hat{r}_k com os limites $\pm 2/\sqrt{n}$ fornece uma indicação geral de possível quebra de comportamento de ruído branco em ε_t , com a condição de que seja lembrado que, para pequenos valores de k , estes limites subestimarão a significância de qualquer discrepância.

3.4.2 Teste de Box-Pierce-Ljung

Box e Pierce (1970) sugeriram um teste para as autocorrelações dos resíduos estimados, que, apesar de não detectar quebras específicas no comportamento de ruído branco, pode indicar se esses valores são muito altos. Uma modificação deste teste foi proposta por Ljung e Box (1978), que é apresentada a seguir.

Se o modelo for apropriado, a estatística

$$Q(K) = n(n+2) \sum_{k=1}^K \frac{\hat{r}_k^2}{(n-k)} \quad (3.26)$$

terá uma distribuição χ^2 com $K - p - q$ graus de liberdade. A hipótese de ruído branco para os resíduos é rejeitada para valores grandes de $Q(\cdot)$. Em geral, basta utilizar as 15 ou 20 primeiras \hat{r}_k .

8.2.3 Teste da autocorrelação cruzada

A verificação das autocorrelações \hat{r}_k dá informação sobre novos termos de médias móveis a serem incluídos no modelo. Por exemplo, se $|\hat{r}_8| > 2/\sqrt{n}$, um termo $\theta_8 a_{t-8}$ deverá ser incluído no modelo.

Um outro teste que pode auxiliar no procedimento de identificação é aquele baseado na correlação cruzada entre valores passados da série e o valor presente do ruído.

De fato, se o modelo for adequado, então a_t e Z_{t-k} serão não-correlacionadas, para $k \geq 1$, logo $\text{Cov}\{a_t, Z_{t-k}\} = \gamma_{az}(k) = 0, k \geq 1$. Isto sugere investigar a função de correlação cruzada (fcc)

$$s_k = \frac{\sum a_t(Z_{t-k} - \bar{Z})}{[\sum a_t^2 \sum (Z_t - \bar{Z})^2]^{1/2}}, \quad k = 1, 2, 3, \dots \quad (8.6)$$

Se para um dado k_0 , s_{k_0} tem valor “grande”, isto sugere que o modelo é inadequado. Em particular, se o modelo tentativo é um $\text{AR}(p)$, um novo termo auto-regressivo no lag k_0 deve ser incluído no modelo. Como não conhecemos os verdadeiros a_t , consideramos os resíduos estimados \hat{a}_t e substituímos s_k por

$$\hat{s}_k = \frac{\sum_{t=k+1}^n \hat{a}_t(Z_{t-k} - \bar{Z})}{[\sum_{t=1}^n \hat{a}_t^2 \sum_{t=1}^n (Z_t - \bar{Z})^2]^{1/2}}, \quad k = 1, 2, \dots \quad (8.7)$$

Pode-se demonstrar que, se Z_t for estacionário, com fac ρ_k , então

$$E(s_k) \rightarrow 0, \quad (8.8)$$

$$\text{Var}(s_k) \rightarrow \frac{1}{n-k} \simeq \frac{1}{n}, \quad k > 0, \quad (8.9)$$

$$\text{Cov}\{s_k, s_{k+l}\} \rightarrow \frac{1}{n-k} \rho_l \simeq \frac{1}{n} \rho_l, \quad k > 0. \quad (8.10)$$

As relações (8.8) e (8.9) mostram que $\gamma_{az}(k)$ é significativamente diferente de zero se $|s_k| > 2/\sqrt{n}$. Mas (8.8) e (8.9) não são válidas quando usamos os resíduos estimados \hat{a}_t .

Contudo, Hokstad (1983) mostra que $1/n$ é um limite superior para $\text{Var}(\hat{s}_k)$, quando $Z_t \sim \text{AR}(p)$. Portanto, o critério de julgar s_k significante quando $|\hat{s}_k| > 2/\sqrt{n}$ é razoável, exceto para k pequeno. Observe que para k pequeno o mesmo problema ocorre para o teste da fac dos resíduos.

Assim, podemos utilizar esses resultados para construir um modelo ARMA:

- (i) começando com um $\text{AR}(p)$ de baixa ordem, podemos incluir novos termos autorregressivos, analisando a fcc \hat{s}_k ;
- (ii) quando s_k não se apresentar mais significante, a fac \hat{r}_k pode indicar termos de médias móveis a serem incluídos;
- (iii) se termos de médias móveis são incluídos num estágio anterior de identificação, a interpretação de valores grandes para $|s_k|$ não é tão óbvia.

Exemplo 8.1. Vamos apresentar um exemplo analisado por Hokstad (1983). Considere a série de manchas solares de Wolf, com $N = 264$ observações (1705–1968). As autocorrelações e autocorrelações parciais sugerem um modelo AR(2), que estimado resulta

$$Z_t = 16,7 + 1,39Z_{t-1} - 0,70Z_{t-2} + a_t, \quad \hat{\sigma}_a^2 = 253. \quad (8.11)$$

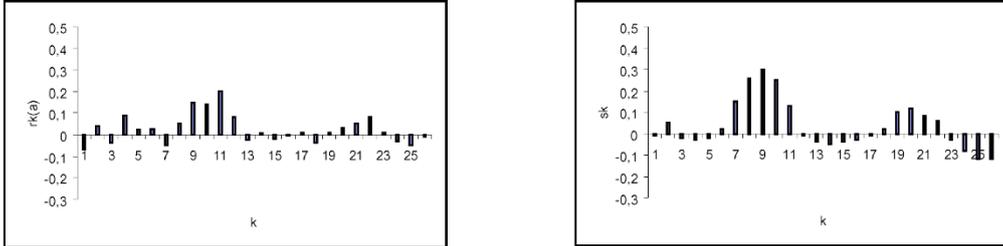


Figura 8.1: Função de autocorrelação estimada de a_t e função de correlação cruzada estimada de a_t e Z_{t-k} , modelo (8.11).

As autocorrelações \hat{r}_k sugerem que termos de médias móveis de ordens 9, 10 e 11 podem eventualmente ser incluídos no modelo (8.11). O novo modelo estimado fica

$$Z_t = 13,9 + 1,27Z_{t-1} - 0,57Z_{t-2} + a_t + 0,23a_{t-9} + 0,22a_{t-10} + 0,20a_{t-11}, \quad (8.12)$$

com $\hat{\sigma}_a^2 = 224,8$.

Por outro lado, as \hat{s}_k sugerem que um modelo autorregressivo de ordem nove deve ser incluído, resultando

$$Z_t = 6,5 + 1,23Z_{t-1} - 0,54Z_{t-2} + 0,18Z_{t-9} + a_t, \quad \hat{\sigma}_a^2 = 221,4. \quad (8.13)$$

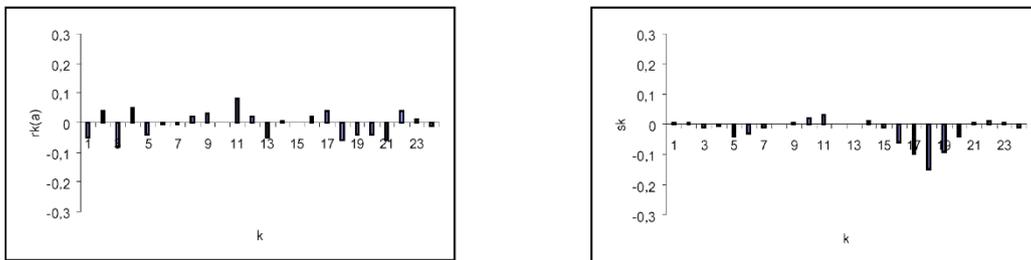


Figura 8.2: Função de autocorrelação estimada de a_t e função de correlação cruzada estimada de a_t e Z_{t-k} , modelo (8.13).

As correlações \hat{r}_k e \hat{s}_k resultantes do modelo (8.13), dadas na Figura 8.2, sugerem que um termo autorregressivo de ordem 18 deve ser incluído em (8.13), obtendo-se

$$Z_t = 9,8 + 1,19Z_{t-1} - 0,53Z_{t-2} + 0,24Z_{t-9} - 0,10Z_{t-18} + a_t, \quad \hat{\sigma}_a^2 = 220,8 \quad (8.14)$$

que, por sua vez, apresenta correlações residuais \hat{r}_k e \hat{s}_k adequadas. Como (8.13) e (8.14) apresentam variâncias residuais muito próximas, o modelo (8.13) deve ser preferido. Neste exemplo, $\hat{s}_9 = 0,32$ (modelo 8.12) $\hat{s}_{18} = 0,135$ (modelo 8.13) e $2/\sqrt{n} = 0,12$, de modo que ambas são significativas.

8.3 Uso dos resíduos para modificar o modelo

Suponha que os resíduos b_t do modelo ajustado

$$\phi_0(B)\Delta^{d_0}Z_t = \theta_0(B)b_t \quad (8.17)$$

não sejam aleatórios. Usando o método de identificação da seção anterior, podemos descrever os resíduos através do modelo

$$\bar{\phi}(B)\Delta^{\bar{d}}b_t = \bar{\theta}(B)a_t. \quad (8.18)$$

Substituindo (8.18) em (8.17) temos um novo modelo

$$\phi_0(B)\bar{\phi}(B)\Delta^{d_0}\Delta^{\bar{d}}Z_t = \theta_0(B)\bar{\theta}(B)a_t, \quad (8.19)$$

cujos resíduos são aleatórios, e que deverá ser ajustado aos dados. O ciclo de identificação, estimação e verificação deve ser continuado, até que um modelo satisfatório seja encontrado.

Exemplo 6.3. Vamos agora identificar um ou mais modelos preliminares para a série M-ICV (simplesmente ICV), no período de janeiro de 1970 a junho de 1979, utilizando $N = 114$ observações.

Inicialmente, a série foi dividida em grupos com 8 observações consecutivas, calculando-se para cada grupo a média e a amplitude (ou desvio padrão), como sugerido na seção 1.6. Na Figura 6.2 temos as representações gráficas desse procedimento. O comportamento de ambos os gráficos sugere que uma transformação logarítmica é necessária para estabilizar a variância da série ICV (ver Figura 1.6).

Entretanto, quando calculamos os valores de d_λ , expressão (1.8), observamos que o menor valor absoluto é obtido para $\lambda = -0,5$ (ver Tabela 6.3), indicando a necessidade da transformação $1/\sqrt{z_t}$ para tornar os dados mais simétricos ou mais próximos de uma distribuição normal. Assim, neste caso, não existe uma transformação que estabilize a variância e, ao mesmo tempo, torne os dados mais simétricos. Diante deste fato, optamos pela transformação que estabiliza a variância.

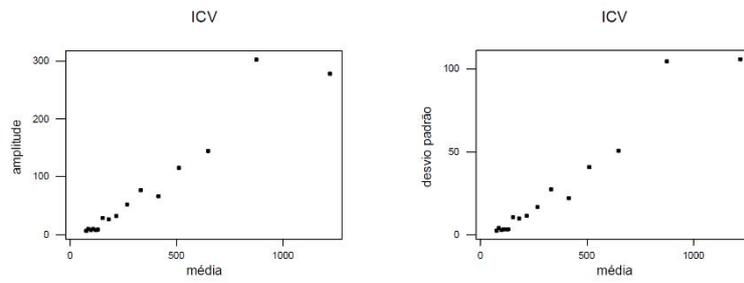


Figura 6.2: Gráficos da série ICV: (a) amplitude \times média e (b) desvio padrão \times média

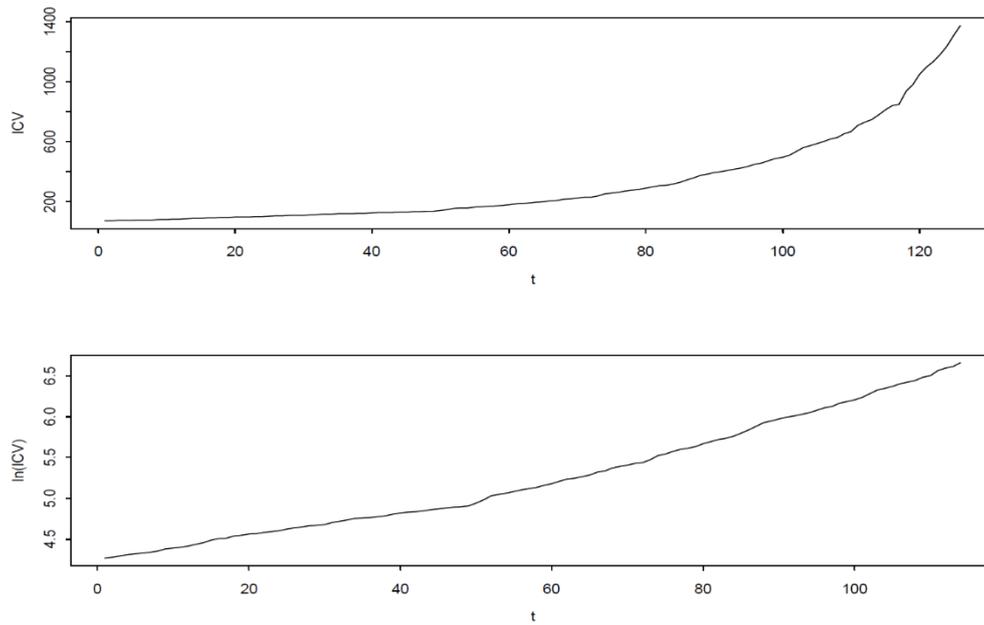


Figura 6.3: Gráficos das séries ICV e $\ln(ICV)$.

Tabela 6.3: Valores da estatística d_λ , dada por (1.8)

λ	d_λ	Transformação
-1,0	0,125	$1/Z_t$
-0,5	0,032	$1/(Z_t)^{1/2}$
0	0,192	$\ln(Z_t)$
0,5	0,328	$(Z_t)^{1/2}$
1,0	0,423	Z_t

Daqui em diante, trabalharemos com a transformação

$$Y_t = \ln(ICV), \quad t = 1, \dots, 114. \quad (6.13)$$

A Figura 6.3 apresenta a série original e a série transformada.

A Tabela 6.4 e a Figura 6.4 apresentam os valores e as correspondentes representações gráficas das funções de autocorrelação e autocorrelação parcial da série Y_t e de suas diferenças de ordens 1 e 2.

Analisando o comportamento das funções vemos que a fac de Y_t não decresce rapidamente para zero, indicando a não-estacionariedade da série. Para selecionarmos um valor apropriado para d , podemos verificar, graficamente, quantas diferenças são necessárias para que a fac convirja rapidamente para zero. Este fato parece ocorrer para $d = 1$ ou, talvez, $d = 2$.

A Figura 6.5 apresenta a primeira e a segunda diferenças do $\ln(ICV)$.

Tabela 6.4: Autocorrelações e autocorrelações parciais estimadas das séries (a) $\ln(ICV)$, (b) $(1 - B)\ln(ICV)$, (c) $(1 - B)^2\ln(ICV)$.

lag	(a)		(b)		(c)	
	acf	pacf	acf	pacf	acf	pacf
1	0,97	0,97	0,49	0,49	-0,32	-0,32
2	0,94	-0,01	0,33	0,12	-0,12	-0,24
3	0,92	-0,02	0,26	0,08	0,00	-0,15
4	0,89	-0,01	0,16	-0,03	-0,10	-0,23
5	0,86	0,00	0,21	0,14	-0,04	-0,24
6	0,84	-0,02	0,27	0,15	0,08	-0,14
7	0,81	-0,01	0,29	0,11	0,00	-0,13
8	0,78	-0,02	0,29	0,08	0,05	-0,06
9	0,76	-0,01	0,25	0,02	-0,02	-0,07
10	0,73	-0,02	0,22	0,04	-0,02	-0,06
11	0,70	-0,02	0,21	0,05	-0,09	-0,17
12	0,68	-0,02	0,28	0,14	0,03	-0,14
13	0,65	0,00	0,30	0,08	0,09	-0,03
14	0,62	0,00	0,26	0,01	0,01	-0,01
15	0,60	-0,01	0,23	0,00	-0,10	-0,14
16	0,57	-0,01	0,27	0,12	0,14	0,06
17	0,55	-0,02	0,16	-0,09	-0,12	-0,05
18	0,53	-0,01	0,19	0,05	0,04	0,05

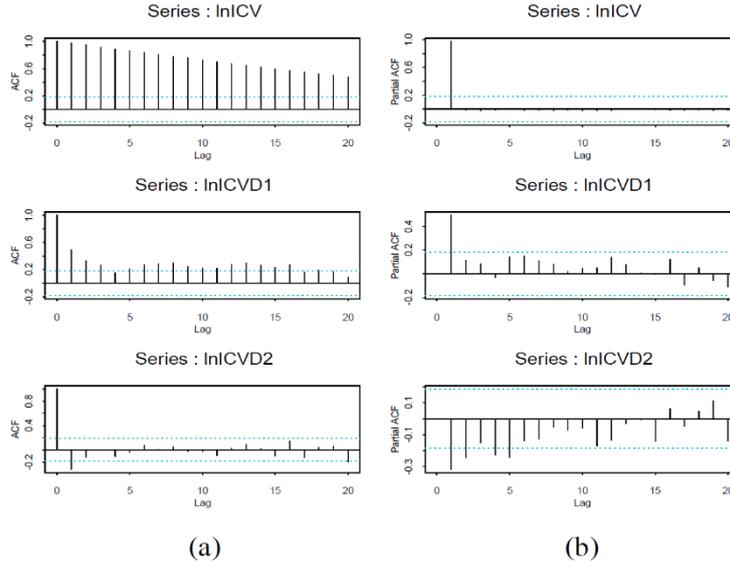


Figura 6.4: Autocorrelações (a) e autocorrelações parciais (b) das séries $\ln(\text{ICV})$, $(1 - B)\ln(\text{ICV})$ e $(1 - B)^2\ln(\text{ICV})$.

A escolha entre esses dois valores, $d = 1$ ou $d = 2$, pode ser feita analisando-se as observações feitas por McLeod (1983), mencionadas no procedimento de identificação e/ou aplicando o teste da raiz unitária de Dickey-Fuller (Apêndice B).

Observando a fac amostral de $(1 - B)^2 Y_t$, verificamos que $r_1 = -0,32$, que além de ser negativo, assume um valor próximo de $-0,5$. Calculando as variâncias das diferenças das séries, temos que $\text{Var}((1 - B)Y_t) = 0,00014$ e $\text{Var}((1 - B)^2 Y_t) = 0,000162$ indicando um aumento de variância da série com duas diferenças em relação àquela com uma diferença. Assim, de acordo com McLeod (1983), escolhemos $d = 1$.

Para a aplicação do teste de Dickey-Fuller utilizamos a função `ur.df` da biblioteca `urca` do repositório R. Escolhemos $p = 1$, uma vez que a facp amostral da série $\ln(\text{ICV})$ exibe um único valor alto, $\hat{\phi}_{11} = 0,97$, indicando um valor de ϕ_{11}

diferente de zero (ver Tabela 6.4). De acordo com (B.19) foi ajustado o modelo

$$\Delta Y_t = \phi_0^* + \phi_1^* Y_{t-1} + a_t$$

com $\hat{\phi}_1^* = 0,009207$ e $\hat{\sigma}(\hat{\phi}_1^*) = 0,001267$, implicando em $\hat{\tau}_\mu = 7,2668$, com valores críticos da estatística $\hat{\tau}_\mu$, dados por $-3,46$, $-2,88$ e $-2,57$, com os níveis de significância $0,01$, $0,05$ e $0,10$, respectivamente.

Assim, $\hat{\tau}_\mu = 7,2668 > -3,46$, o que implica em não rejeitar a hipótese de uma raiz unitária no operador autorregressivo, ou seja, $d \geq 1$, ao nível de 1%.

Para testar a existência de duas raízes unitárias, não podemos usar os testes DF ou ADF do Apêndice B. Há um procedimento para testar se a série contém no máximo duas raízes unitárias. Usamos este procedimento aqui e concluímos que $d = 1$.

Analisando o comportamento das fac e facp da série $(1 - B) \ln(\text{ICV})$, Tabela

6.4 e Figura 6.4, podemos notar que somente $\phi_{11} \neq 0$, pois $\hat{\phi}_{11} = 0,49$ e $dp(\hat{\phi}_{11}) = \frac{1}{\sqrt{114}} = 0,094$. Assim, um modelo preliminar para $\ln(\text{ICV})$ é um ARIMA(1,1,0).

Para verificar se uma constante deve, ou não, ser adicionada ao modelo, observamos que

$$\begin{aligned} \bar{W} &= 0,02116 \text{ e} \\ \text{Var}(\bar{W}) &= \frac{c_0(1 + r_1)}{n(1 - r_1)} = \frac{0,00014(1 + 0,49)}{113(1 - 0,49)} = 3,62 \times 10^{-6} \end{aligned}$$

ou seja, $DP(\bar{W}) = 0,0019$.

Assim, podemos concluir que a média da série $(1 - B)Y_t$ é diferente de zero e que uma constante deve ser incluída no modelo.

Logo, o modelo preliminar para a série $\ln(\text{ICV})$ é uma ARIMA(1,1,0) com uma constante, isto é,

$$(1 - B)(1 - \phi_1 B)Y_t = \theta_0 + a_t, \quad (6.14)$$

onde $Y_t = \ln(\text{ICV})$.

Exemplo 8.2. Vamos agora testar se o modelo ARIMA(1,1,0) com θ_0 (equação (6.14)), proposto para a série $\ln(\text{ICV})$, é adequado.

O Quadro 8.1 apresenta os valores dos parâmetros estimados e os respectivos valores da estatística t , além dos p -valores. Temos, também, o valor do critério AIC corrigido. Para os parâmetros estimados, os p -valores indicam que ambos são não nulos. A estimativa denominada “constant” é, na realidade, a estimativa da média ($\hat{\mu}$). Desse modo, a estimativa da constante do modelo é dada por

$$\hat{\theta}_0 = (1 - \hat{\phi})\hat{\mu} = 0,01045.$$

Para verificar o ajustamento do modelo, analisamos os resultados apresentados na Figura 8.5. Os valores das autocorrelações residuais não indicam nenhuma quebra de comportamento de ruído branco nos resíduos. Os p -valores associados à estatística de Ljung-Box, para $K \leq 12$, também indicam um bom ajustamento do modelo ARIMA(1, 1, 0) com constante. Segue-se que o modelo ajustado à série $\ln(\text{ICV})$ é dado por

$$(1 - 0,5073B)(1 - B) \ln(\text{ICV}) = 0,01045 + \hat{a}_t, \quad (8.20)$$

com $\hat{\sigma}_a^2 = 0,0000908$.

Quadro 8.1: Ajustamento de um modelo ARIMA(1,1,0) à série $\ln(\text{ICV})$.				
	Estimate	SE	t-value	p-value
ar1	0.5073	0.0822	6.1694	0
constant	0.0212	0.0018	11.755	0
sigma2 estimated as 9.076e-05				
log likelihood = 365.38				
AIC = -724.75				

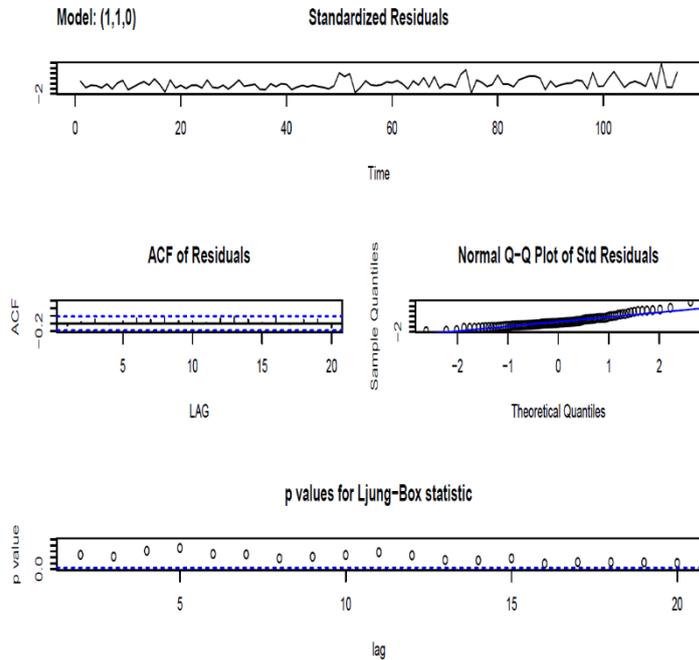


Figura 8.5: Análise residual do modelo ARIMA(1,1,0) ajustado à série $\ln(\text{ICV})$.

Exemplo 8.3. Vamos verificar se o modelo AR(1) com θ_0 (equação 6.15) proposto para a série Umidade é adequado.

O Quadro 8.2 apresenta um resumo do ajustamento feito com a função `sarima` da biblioteca `astsa` do R. Notamos a significância dos parâmetros do modelo (ϕ_1 e μ). Entretanto, ao analisarmos a Figura 8.6, notamos que r_2 está fora do intervalo de confiança, indicando que o modelo não está bem ajustado.

Quadro 8.2: Modelo AR(1) ajustado à série de Umidade.

	Estimate	SE	t-value	p-value
ar1	0.5306	0.0446	11.8902	0
xmean	81.2416	0.7577	107.2238	0

sigma2 estimated as 45.57
log likelihood = -1191.79
AIC= 2389.58

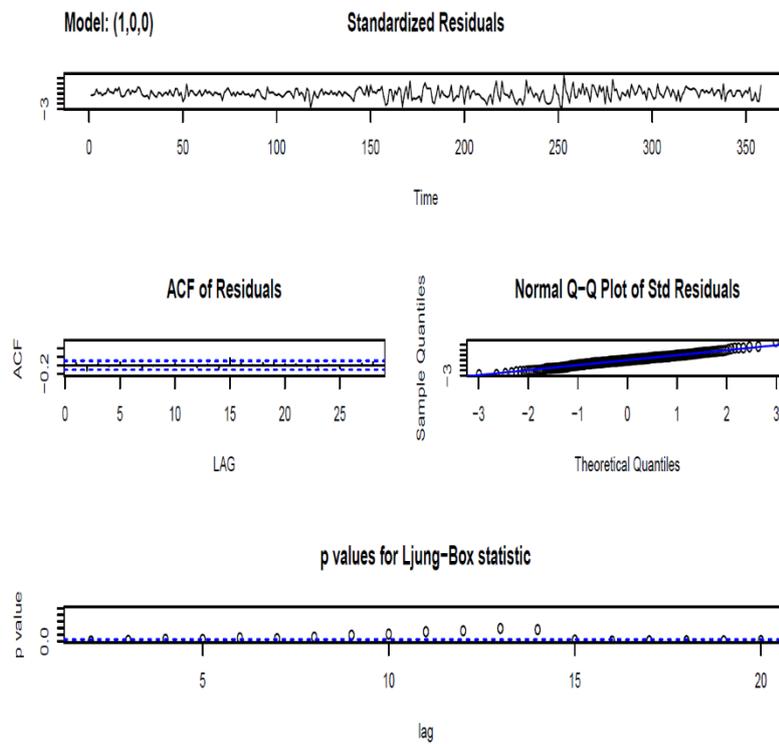


Figura 8.6: Análise residual para o modelo AR(1) ajustado à série de Umidade.

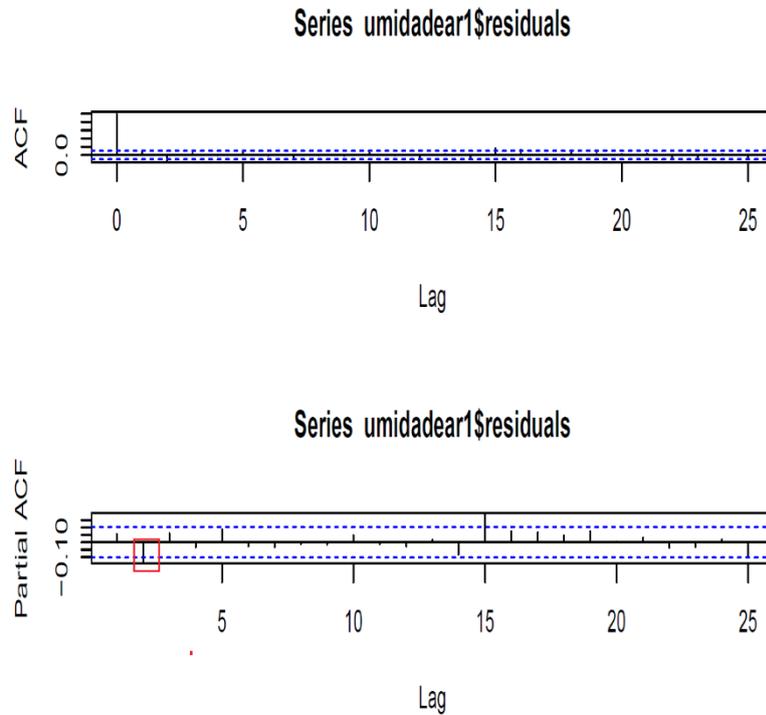


Figura 8.7: Fac e facp dos resíduos do modelo AR(1) ajustado á série Umidade.

Analisando a Figura 8.7, observamos que $\hat{\phi}_{22}$ também está fora do intervalo de confiança, sugerindo um modelo AR(2) para a série Umidade.

O Quadro 8.3 apresenta um resumo do ajustamento de um modelo AR(2). Podemos notar que o p -valor associado ao parâmetro ϕ_2 está muito próximo de 0,05, e por esse motivo, será mantido no modelo.

Para verificarmos a adequação do modelo AR(2), analisamos a Figura 8.8, constatando que o modelo está bem ajustado: a f.a.c. residual apresenta quase todos os valores dentro do intervalo de confiança, com exceção feita a r_{15} e r_{16} , que não foram levadas em conta, pois referem-se a dependências distantes (mais de duas semanas).

Além disso, temos que os valores da estatística $Q(K)$ nos levam a não rejeitar a hipótese de ruído branco para os resíduos. Assim, o modelo ajustado à série Umidade é dado por

$$(1 - 0,5839B + 0,0997B^2)Z_t = 41,8565 + \hat{a}_t. \quad (8.21)$$

em que $\hat{\theta}_0 = (1 - 0,5839 + 0,0997)(81,1488) = 41,8565$ e $\hat{\sigma}_a^2 = 44,82$.

Quadro 8.3: Modelo AR(2) ajustado à série de Umidade.

	Estimate	SE	t-value	p-value
ar1	0.5839	0.0520	11.2281	0.0000
ar2	-0.0997	0.0521	-1.9143	0.0564
xmean	81.1488	0.6780	119.6900	0.0000

sigma2 estimated as 44.82
log likelihood = -1212.06
AIC = 2432.13

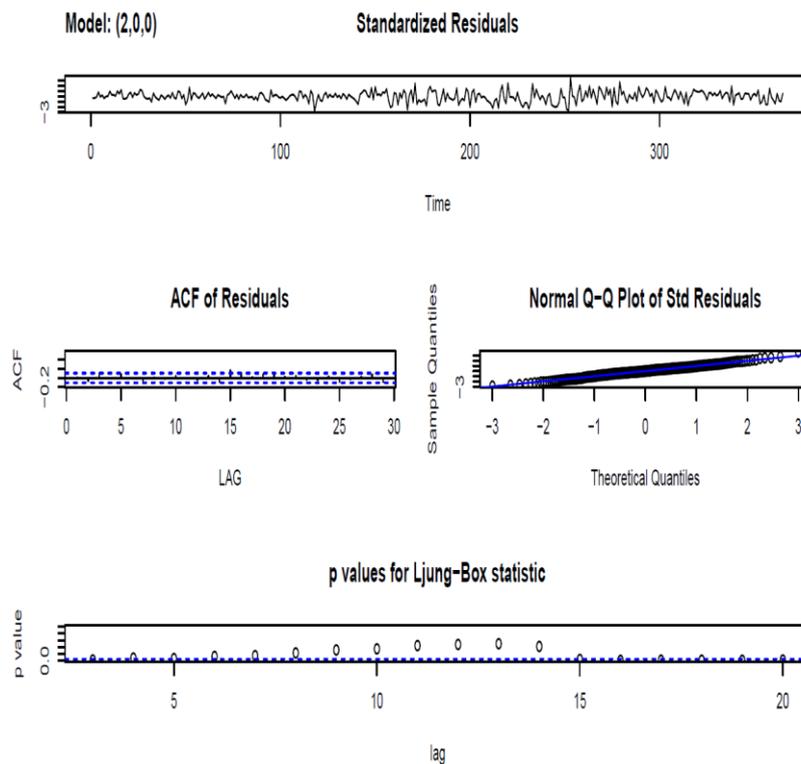


Figura 8.8: Análise residual do modelo AR(2) ajustado

Exemplo 8.4. Vamos analisar o modelo AR(3) com θ_0 (equação 6.17) proposto para a série $\ln(\text{CO})$.

O Quadro 8.4 apresenta o ajustamento feito pela função `sarima` da biblioteca `astsa` do R. Podemos notar a significância de todos os parâmetros.

Quadro 8.4: Modelo AR(3) ajustado à série de ln(CO).

	Estimate	SE	t-value	p-value
ar1	0.7211	0.0520	13.8683	0e+00
ar2	-0.2740	0.0628	-4.3624	0e+00
ar3	0.1741	0.0521	3.3389	9e-04
xmean	1.2235	0.0426	28.6978	0e+00

sigma2 estimated as 0.09429
log likelihood = -85.59
AIC = 181.17

A Figura 8.9 apresenta a análise residual do modelo ajustado. Os valores das autocorrelações r_j , $j \leq 13$ estão dentro do intervalo de confiança, indicando que o modelo está bem ajustado. Como explicado antes, os valores de r_{14} , r_{16} e r_{21} , embora significativos, serão desprezados.

Os p -valores da estatística de Ljung-Box também indicam um bom ajustamento do modelo AR(3), com constante.

Temos, pois, que o modelo ajustado para a série $\ln(\text{CO})$ é dado

$$(1 - 0,721B + 0,274B^3 - 0,174B^3) \ln(\text{CO}) = 0,310 + \hat{a}_t, \quad (8.22)$$

com $\hat{\sigma}_a^2 = 0,0943$ e $\hat{\theta}_0 = (1 - 0,721 + 0,274 - 0,174)/(1,224) = 0,310$.

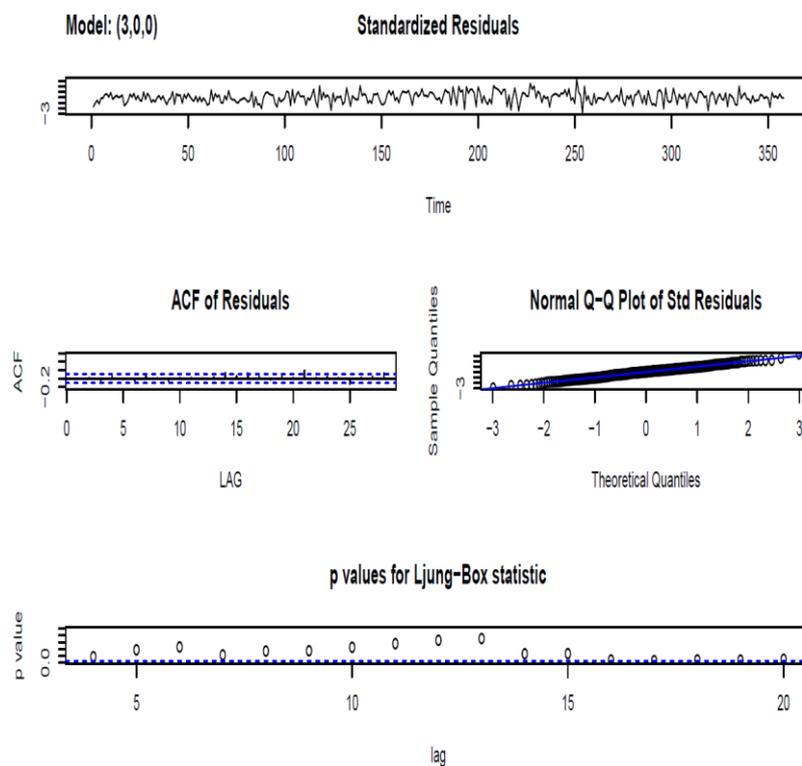


Figura 8.9: Análise residual do modelo AR(3) ajustado à série $\ln(\text{CO})$.