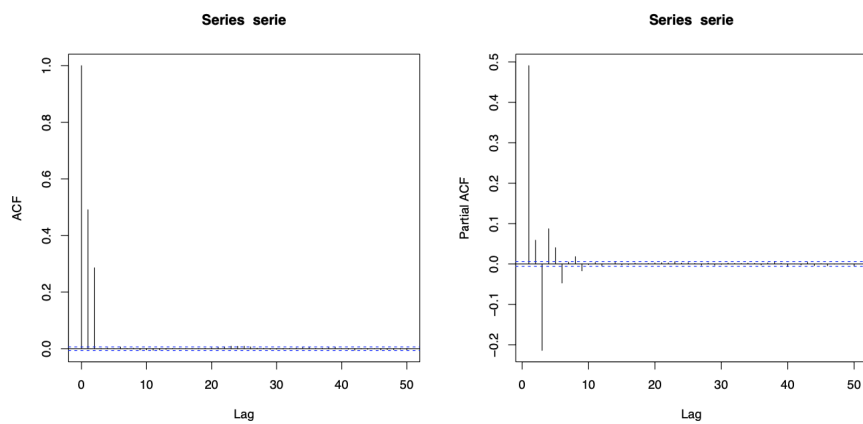


## EAE1223: Econometria III

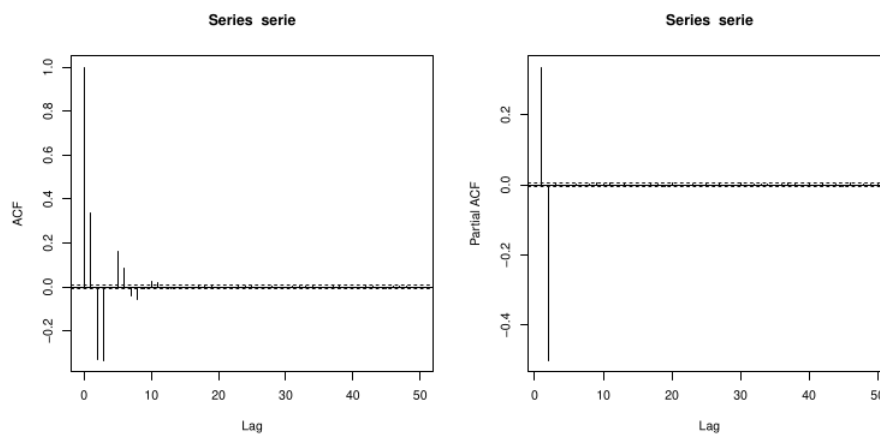
### Exercícios sobre a metodologia de Box-Jenkins

#### Parte 1

- 1 Observando a FAC e FACP, assinale qual processo é mais provável de ter gerado os dados:



- (A) ARMA(5,1)  
(B) AR(3)  
(C) MA(3)  
(D) MA(2)
- 2 Observando a FAC e FACP, assinale qual processo é mais provável de ter gerado os dados. Justifique sua resposta.



- (a) AR(1)
- (b) AR(2)
- (c) MA(1)
- (d) MA(2)
- (e) ARMA(1, 1)

3 Considere o seguinte processo estocástico estacionário:

$$Y_t = \rho_0 Y_{t-1} + \epsilon_t,$$

onde  $\epsilon_t \stackrel{\text{iid}}{\sim} N(0, \sigma_0^2)$ , e  $|\rho_0| < 1$ .

(a) Mostre que:

$$Y_t = \epsilon_t + \sum_{j=1}^{\infty} \rho_0^j \epsilon_{t-j}.$$

(b) Com base no resultado do item anterior, argumente que, para todo  $t > 1$ , a distribuição de  $Y_t$ , condicional a  $Y_1, \dots, Y_{t-1}$ , é Gaussiana com média  $\rho_0 Y_{t-1}$  e variância  $\sigma_0^2$ , isto é:

$$Y_t | Y_1, \dots, Y_{t-1} \sim N(\rho_0 Y_{t-1}, \sigma_0^2).$$

(c) Com base no resultado do item anterior, e no fato de que a densidade de  $Y_2, Y_3, \dots, Y_T$ , condicional a  $Y_1$ , se fatora como:

$$f_{Y_2, \dots, Y_T | Y_1}(Y_2, Y_3, \dots, Y_T | Y_1) = \prod_{t=2}^T f_{Y_t | Y_1, \dots, Y_{t-1}}(Y_t | Y_1, \dots, Y_{t-1});$$

mostre que a densidade de  $Y_2, \dots, Y_T$ , condicional a  $Y_1$  se escreve como:

$$f_{Y_2, \dots, Y_T | Y_1}(Y_2, \dots, Y_T | Y_1) = \frac{1}{\sigma_0^{T-1}} \prod_{t=2}^T \phi\left(\frac{Y_t - \rho_0 Y_{t-1}}{\sigma_0}\right),$$

onde  $\phi(x) = \frac{1}{\sqrt{2\pi}} \exp(-x^2/2)$  é a densidade de uma normal padrão.

(d) Mostre que o estimador  $\hat{\rho}$  de  $\rho_0$  que maximiza a log-verossimilhança de  $Y_2, \dots, Y_T | Y_1$  é idêntico ao estimador condicional de um AR(1) visto em aula, que minimiza:

$$\min_{b \in \mathbb{R}} \sum_{t=2}^T (Y_t - b Y_{t-1})^2$$

## Parte 2

4 Verdadeiro ou falso? Justifique.

O critério MAIC de Ng e Perron difere do critério de informação AIC para avaliação da qualidade preditiva de um modelo, no sentido em que a medida do erro de previsão, no MAIC, é calculada sob a hipótese nula de raiz unitária, enquanto o critério AIC utiliza a soma dos quadrados dos resíduos do modelo irrestrito estimado.

- 5 Você é um pesquisador interessado em projetar uma série  $I(1)$  de interesse. Você decide utilizar a metodologia de Box-Jenkins como ponto de partida, e se encontra na etapa de diagnóstico dos modelos candidatos.
- (a) Um de seus modelos candidatos é um  $ARIMA(1,1,1)$ . Ao analisar os resíduos do modelo estimado, você se depara com o seguinte resultado:

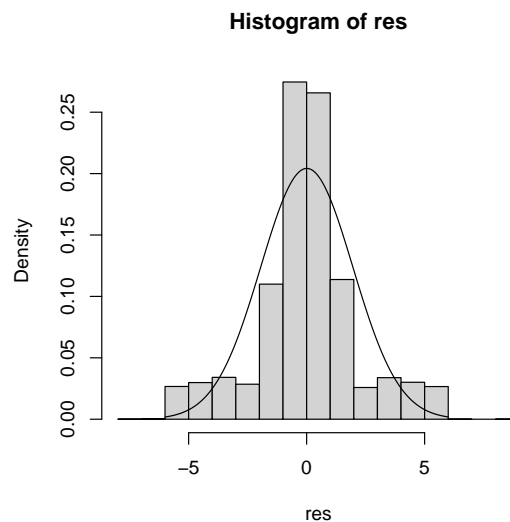
```
[      Ljung-Box test

data:  Residuals from ARIMA(1,1,1)
Q* = 32.957, df = 18, p-value = 0.01689

Model df: 2.    Total lags used: 20
```

Qual é a hipótese nula do teste apresentado acima? E a alternativa? Você rejeita a nula a 5% de significância? Quais as implicações desse resultado, em termos da qualidade preditiva do modelo estimado?

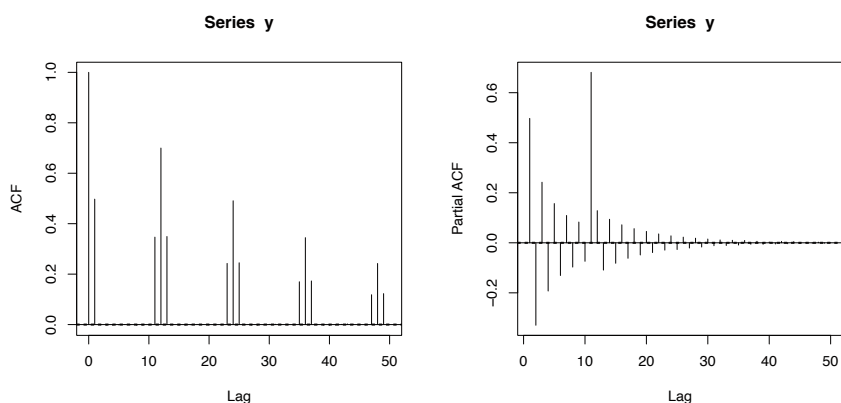
- (b) Você encontra um modelo que parece satisfazer os principais critérios de qualidade preditiva. No entanto, ao rodar o teste de Jarque-Bera, você encontra um  $p$ -valor de 0.000001. Intrigado com o resultado, você constrói um histograma dos resíduos, sobre o qual sobrepõe a distribuição normal cuja média e variância coincidem com a dos resíduos.



Qual é a hipótese nula do teste de Jarque-Bera? E a alternativa? A 5% de significância, você rejeita a hipótese nula? Com base no gráfico acima, indique o motivo suspeito pelo qual você obteve a

referida conclusão no teste. Quais as implicações do resultado do teste para a estimação do seu modelo?

- 6 Para cada uma das séries de tempo analisadas por vocês nos exercícios sobre raiz unitária.
- Divida o conjunto de dados em duas janelas, primeira em segunda, em que a primeira contém 4/5 dos períodos de tempo, e a segunda contém a quinta parte final. *Dica:* use o comando `window` visto em aula.
  - Com base na primeira janela, realize a etapa de identificação da metodologia de Box-Jenkins. Quais são os modelos candidatos? Por quê?
  - Com base na primeira janela, estime os modelos candidatos. Com base nos critérios de diagnóstico, selecione um ou mais modelos com boas métricas. Justifique suas escolhas.
  - Compute as previsões para até um ano fora da primeira janela. Reporte os intervalos de predição associados. Como as previsões se compararam ao que ocorreu na segunda janela? Qual é a interpretação do erro de previsão, nesse caso?
  - Agora, para cada período na segunda janela, compute a previsão um passo à frente, com base nos dados até o período imediatamente anterior, para cada modelo Arima por você selecionado na primeira janela. Calcule o erro quadrático médio, um passo à frente, com base nos erros dessas previsões. Dentre os modelos por você estimados, qual se saiu melhor? Suas conclusões batem com o desempenho calculado no item anterior? Qual a diferença entre as métricas?
- 7 Observando a FAC e FACP da série de tempo *mensal* estacionária abaixo, indique qual processo SARIMA é o mais adequado de haver gerado os dados. Justifique sua resposta.



- 8 Dizemos que uma variável aleatória segue distribuição  $\text{Laplace}(c, d)$  se admite densidade:

$$f(x) = \frac{1}{2} \exp(-|x - c|/d).$$

Nesse caso, a média da variável aleatória é  $c$ , e sua variância é  $2d^2$ .

Mostre que o estimador de máxima verossimilhança do parâmetro  $\rho_0$  de um AR(1) (sem intercepto) com ruído branco  $\epsilon_t \stackrel{\text{iid}}{\sim} \text{Laplace}(0, d_0)$ , que maximiza a log-verossimilhança da distribuição de  $Y_2, \dots, Y_T$  condicional a  $Y_1$ , é idêntico ao estimador  $\hat{\rho}$  que minimiza:

$$\min_{b \in \mathbb{R}} \sum_{t=2}^T |Y_t - bY_{t-1}|$$