Macroeconometria – Séries de tempo

FAUSTO JOSÉ ARAÚJO VIEIRA

Aula 3

17 de abril a 22 de maio de 2018

RESUMO – MODELO ARMA(AULA ANTERIOR)

Revisão: processo de uma série ARMA(p,q)

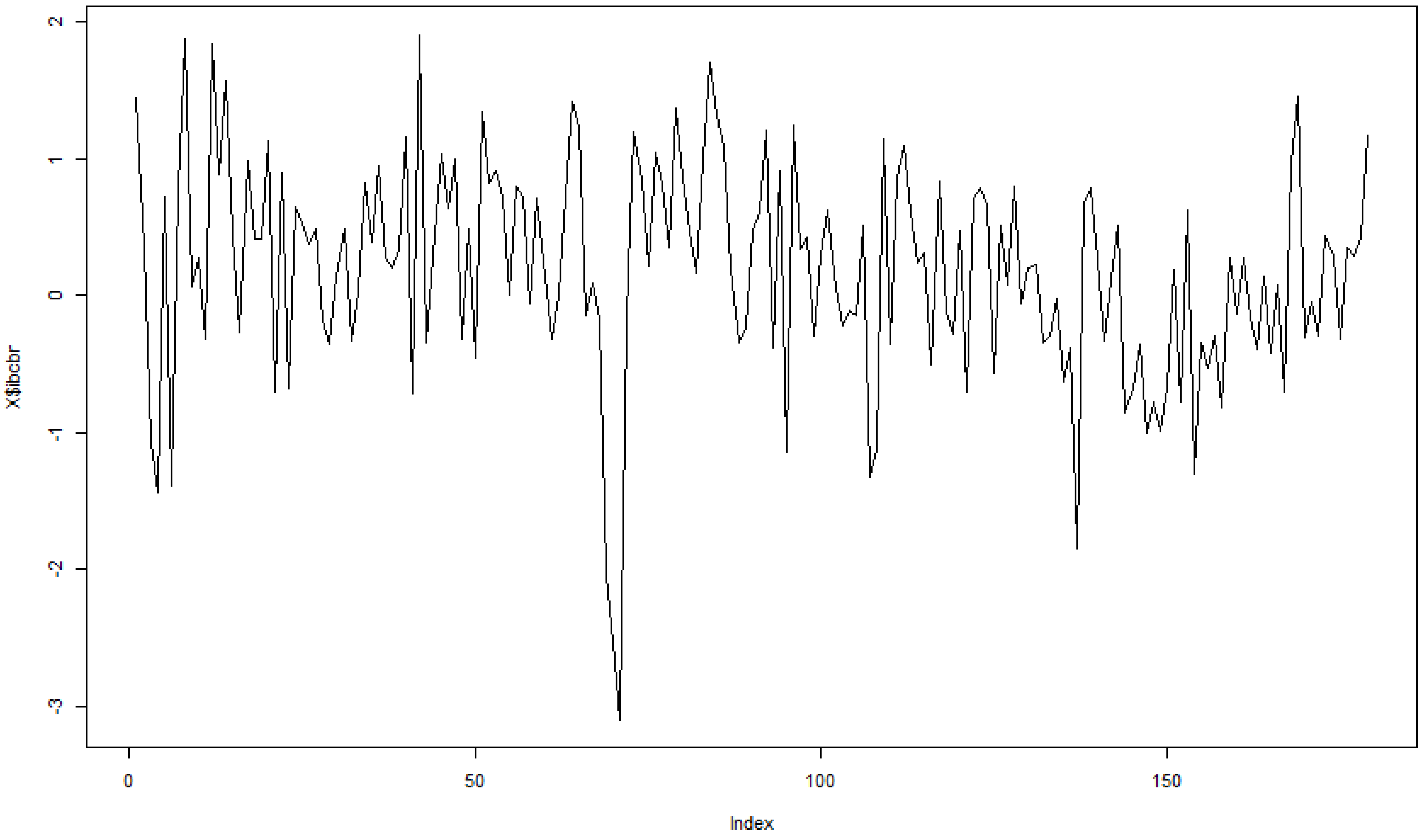
1. Identificar as ordens p e q do modelo.
2. Estimá-lo.
3. Verificar se os resíduos estimados são um ruído branco.
   - Sim, próximo passo;
   - Não, volte ao passo 1.
4. Projetar!

- Se uma série for considerada não estacionária, deve ser diferenciada.

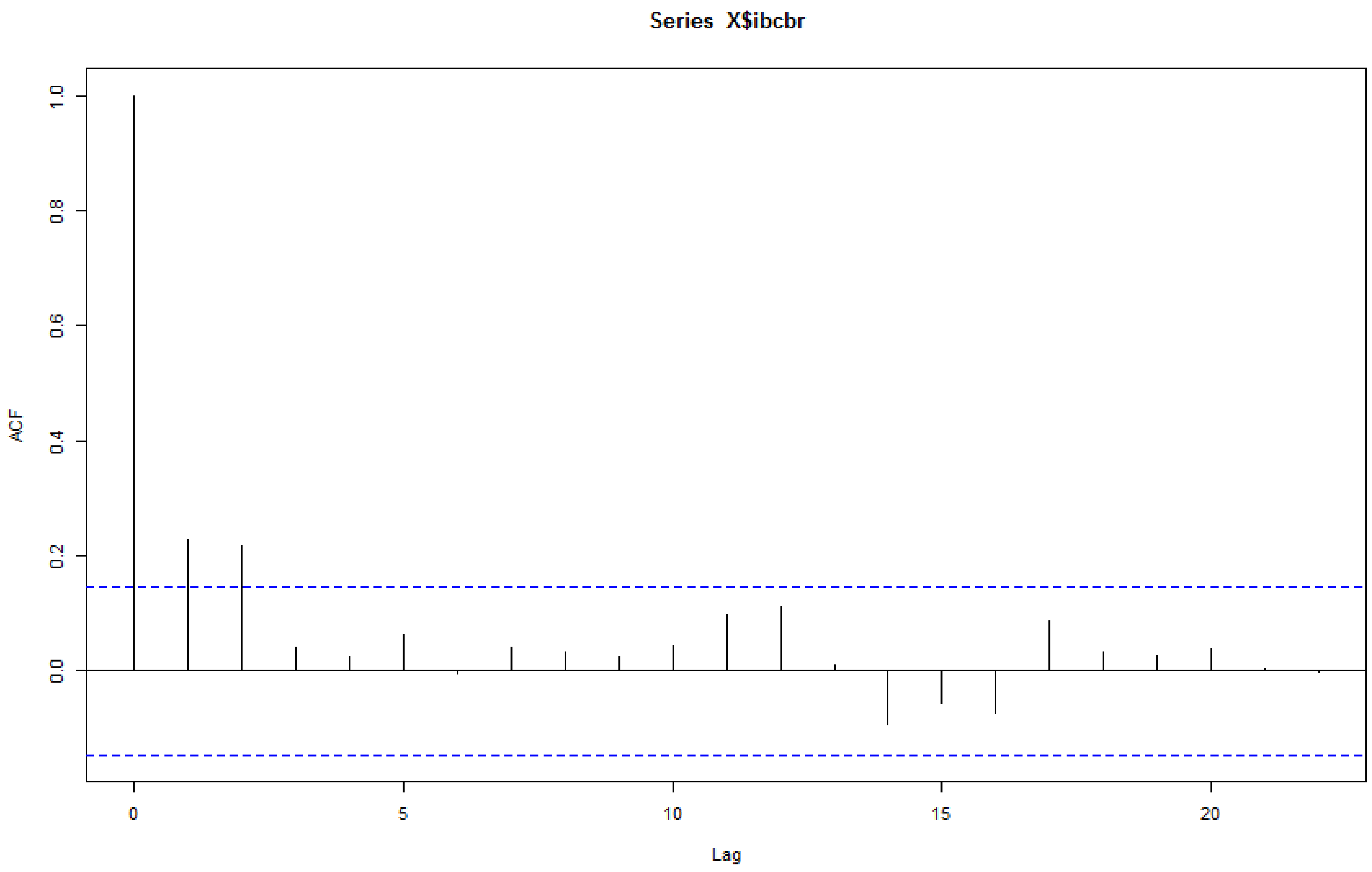
Como escolher o modelo ARMA(p,q)

Modelo	FAC	FACP
$AR\ (p)$	Decai	Truncada na defasagem p
$MA\ (q)$	Truncada na defasagem q	Decai
$ARMA\ (p,q)$	Decai se $j > q$	Decai se $j > p$

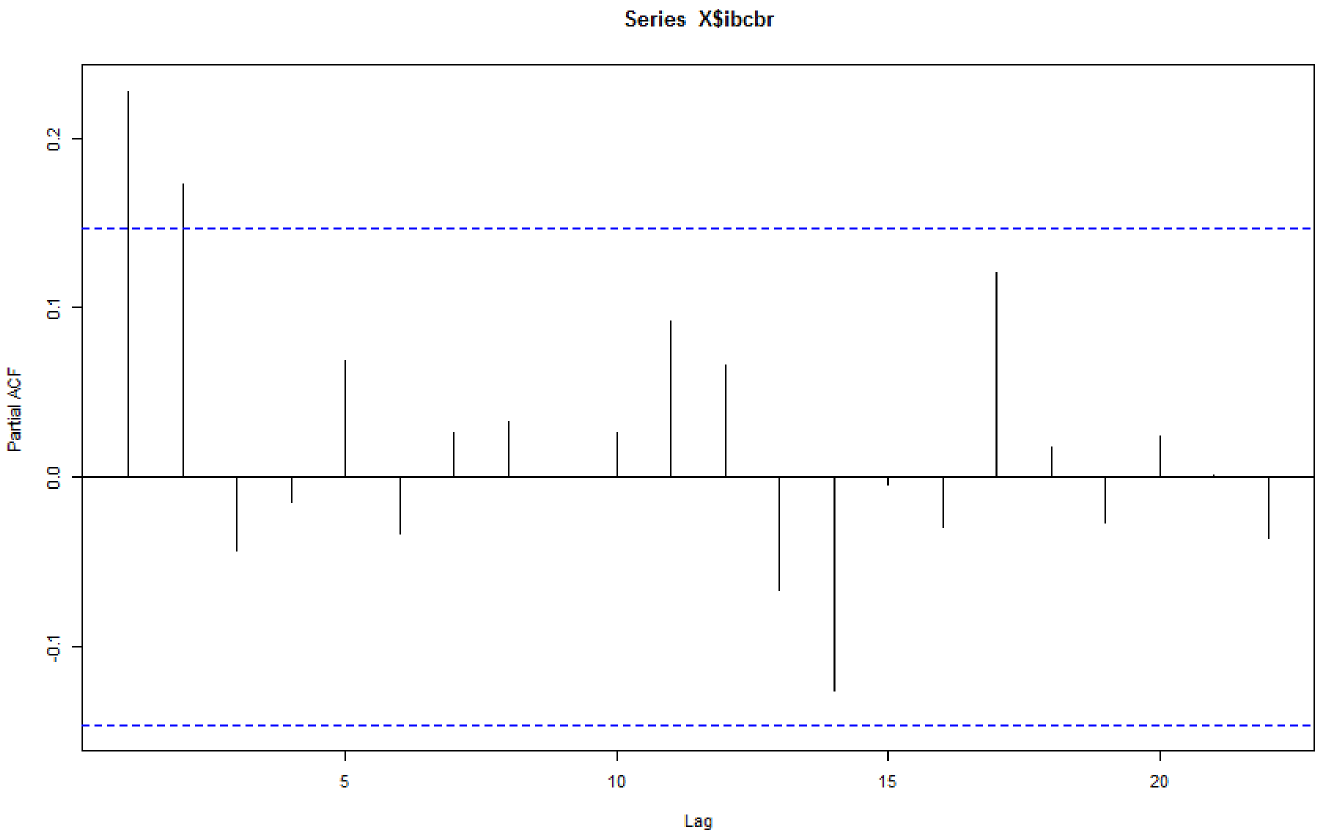
IBC-br



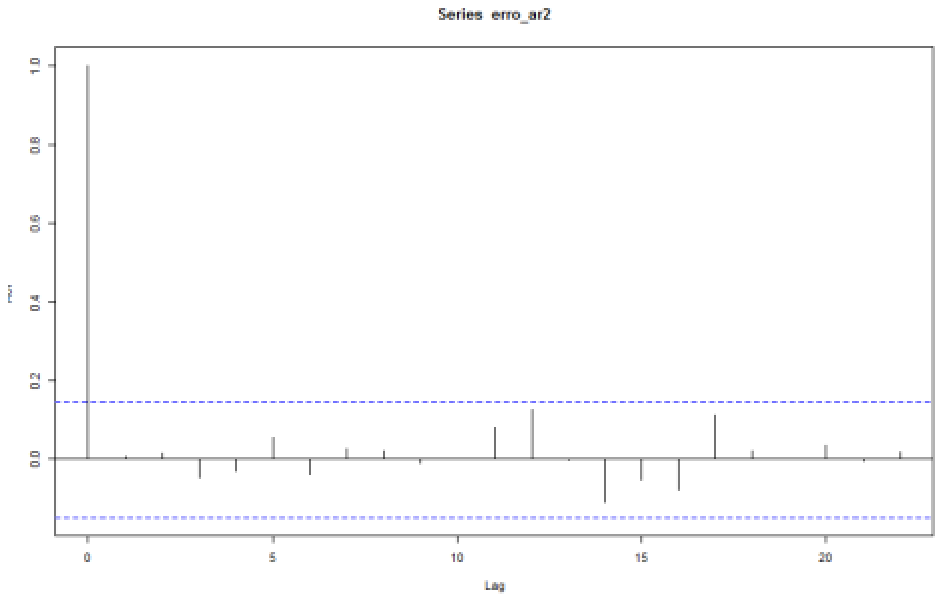
Função de Autocorrelação

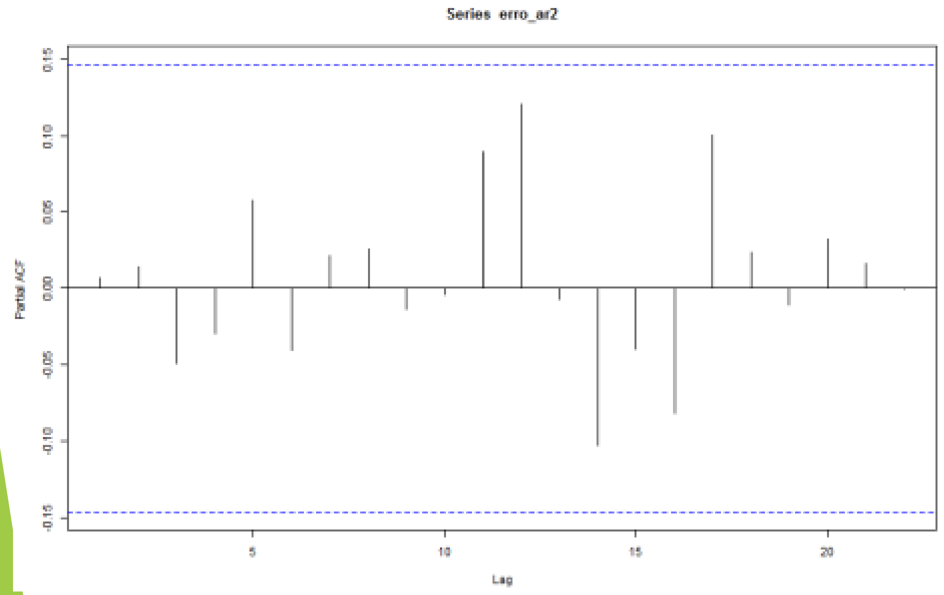


Função de Autocorrelação Parcial



Análise dos resíduos (IBC-br)





m	Qm	pvalue
1	0.01	0.9305878
2	0.04	0.8363149
3	0.48	0.4872894
4	0.65	0.7236980
5	1.22	0.7484525
6	1.49	0.8283449
7	1.61	0.8999974
8	1.69	0.9461123
9	1.71	0.9739582
10	1.71	0.9885560
11	2.95	0.9664136
12	5.95	0.8191315
13	5.95	0.8764191
14	8.21	0.7682411
15	8.80	0.7875653

Critério de informação

- Escolher o modelo baseado nas seguintes estatísticas:

1. Schwarz ou BIC- Baysian Information Criterion;

$$BIC\,(p,q) = \ln \hat{\sigma}^2 + n\frac{\ln T}{T}$$

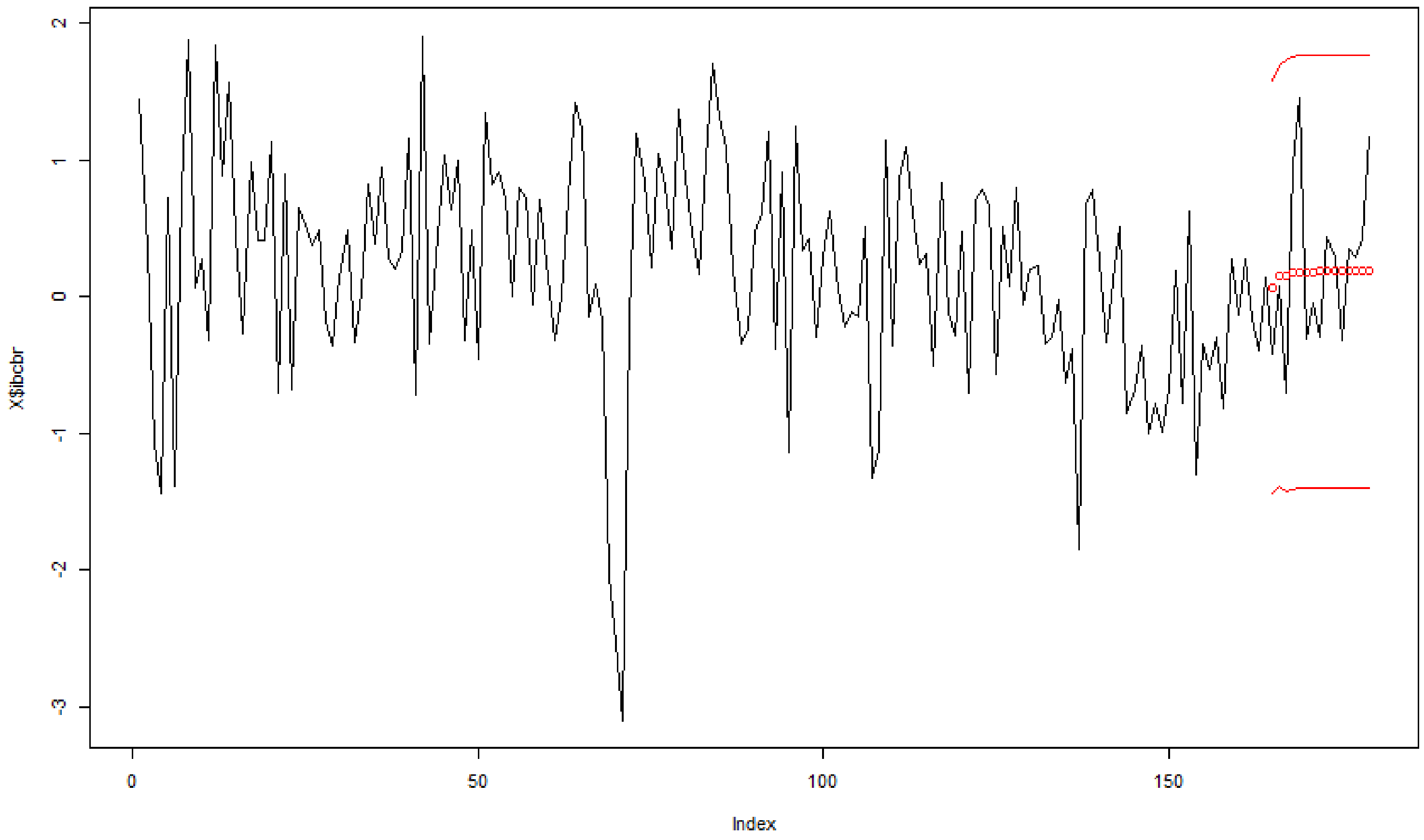
2. Akaike ou AIC – Akaike Information Criterion;

$$AIC\,(p,q) = \ln \hat{\sigma}^2 + n\frac{2}{T}$$

3. Hanna-Quinn, HQ;

$$HQ\,(p,q) = \ln \hat{\sigma}^2 + n\frac{2}{T}\ln \ln T$$

Projetar



RESUMO DA AULA

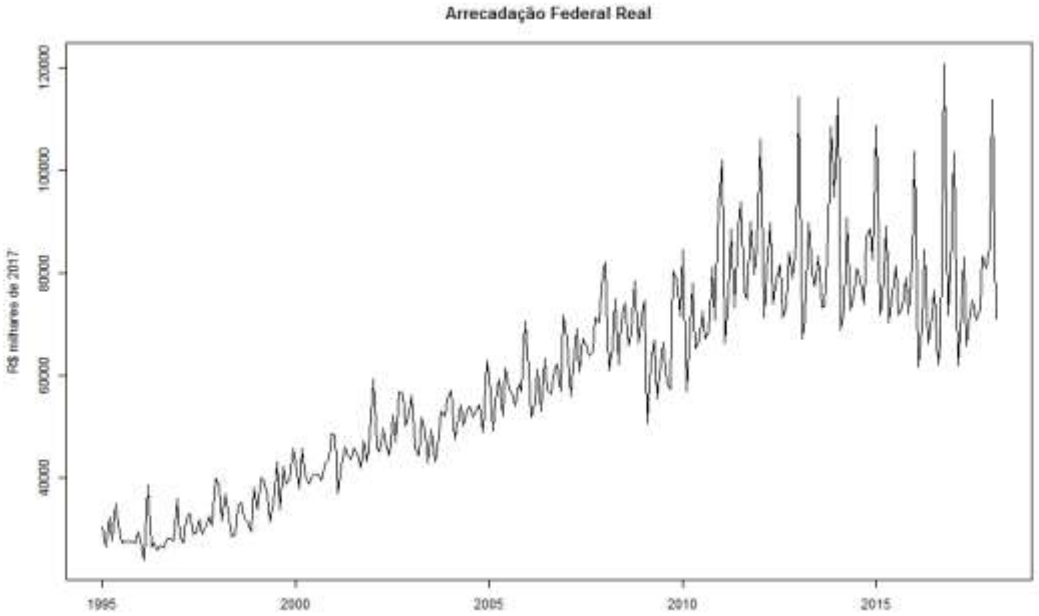
Resumo

1. Médias móveis
2. Modelos com variável binária
3. EWMA
4. SARIMA
5. X-12
6. Tramo-Seats
7. Modelo arima + variáveis explicativas
8. Aplicalção:

MODELOS DE AJUSTE SAZONAL

Sazonalidade e suavização: visão tradicional

- As séries são ajustadas por algoritmos determinísticos.
- Ignora-se a modelagem de componentes estocásticos porventura existentes.
- Alisamento e dessazonalização procuram expurgar fatores que geram perturbações sistemáticas na série, para ter uma ideia mais precisa da tendência que ela segue.
- A figura a seguir mostra um padrão sazonal pela existência de picos e vales igualmente espaçados ao longo do tempo.



Sazonalidade e suavização

- O processo estocástico é produto de quatro fatores:

$$y_t = C_t \times S_t \times T_t \times U_t$$

C_t é um componente de ciclo de longo prazo;
S_t é um componente sazonal;
T_t é um componente de tendência;
U_t é um componente irregular.

- Objetivo: estimar S_t e, em seguida, expurgar esse termo da série y_t, para fins de previsão.

Média móvel tradicional

- Hipótese: componente sazonal fixo para subperíodos iguais, por isso, ignora-se uma possível dinâmica desse componente.
- Calcula-se a média móvel da série y_t definida da seguinte forma:
- Centrada:

$$x_t = \begin{cases} \dfrac{(0{,}5y_{t+6} + y_{t+5} + \cdots + y_t + \cdots y_{t-5} + 0{,}5y_{t-6})}{12}, \text{ se a série é mensal;} \\ \dfrac{(0{,}5y_{t+2} + y_{t+1} + y_t + y_{t-1} + 0{,}5y_{t-2})}{4}, \text{ se a série é trimestral.} \end{cases}$$

Média móvel tradicional

- Filtro elimina a sazonalidade e o componente irregular, tal que :

$$x_t = C_t \times T_t$$

- Isolando o componente sazonal e irregular:

$$z_t \equiv \frac{y_t}{x_t} = \frac{C_t \times S_t \times T_t \times U_t}{C_t \times T_t} = S_t \times U_t$$

- Deseja-se expurgar ou anular o componente irregular de z_t. Para isso, obtém-se a média de z_t nos períodos similares dentro do ano. Supondo que o ano seja dividido em q períodos, encontra-se:

$$\bar{v}_j = \frac{\sum_{m=0}^{\left[\frac{T-j}{q}\right]} z_{mq+j}}{\left[\frac{T-j}{q}\right] + 1}, \quad j = 1, 2, \ldots, q$$

- Em que [.] representa um número interio.

Média móvel tradicional

Se os dados são mensais, $q = 12$. Se houver 100 observações, então $T = 100$. Para a observação constante do primeiro mês da série, $j = 1$:

$$\left[\frac{T-j}{q}\right] = \left[\frac{100-1}{12}\right] = [8,25] = 8.$$

Logo $m = 0, 1, 2, \ldots, 8$, de modo que $mq + j = 1, 13, 25, \ldots, 97$.
No caso em que $j = 5$,

$$\left[\frac{T-j}{q}\right] = \left[\frac{100-5}{12}\right] = [7,91666] = 7.$$

Logo $m = 0, 1, 2, \ldots, 7$, de modo que $mq + j = 5, 17, 29, \ldots, 89$.

Média móvel tradicional

O índice $\overline{v}_j$ deve ser normalizado. Há duas maneiras de proceder.

- Multiplica-se cada índice pelo normalizador $\frac{12}{\sum_{j=1}^{q} \overline{v}_j}$, tal que a soma de $v_j = \overline{v}_j \frac{12}{\sum_{j=1}^{q} \overline{v}_j}$ resultará em 12.
- Para séries com componentes multiplicativos, a normalização deve ser tal que $\prod_{j=1}^{q} v_j = 1$. Para isso:

$$v_j = \frac{\overline{v}_j}{\sqrt[q]{\prod_{j=1}^{q} \overline{v}_j}},$$

isto é, o índice original $\overline{v}_j$ é normalizado pela média geométrica dos índices.

Média móvel tradicional

O índice v_j corresponde ao S_t da composição original da série y_t e é denominado como fator de escala.
Com v_j pode-se dizer que a série y é $(s_j - 1)\%$ maior do que a série isenta de sazonalidade.
O passo final é dividir a série original, y_t pelo índice correspondente ao mês a que pertence t. Formalmente, isso é feito obtendo-se:

$$s_{mq+j} = \frac{y_{mq+j}}{v_j}, m = 0, 1, 2, \ldots, \left[\frac{T-j}{q}\right], j = 1, 2, \ldots, q.$$

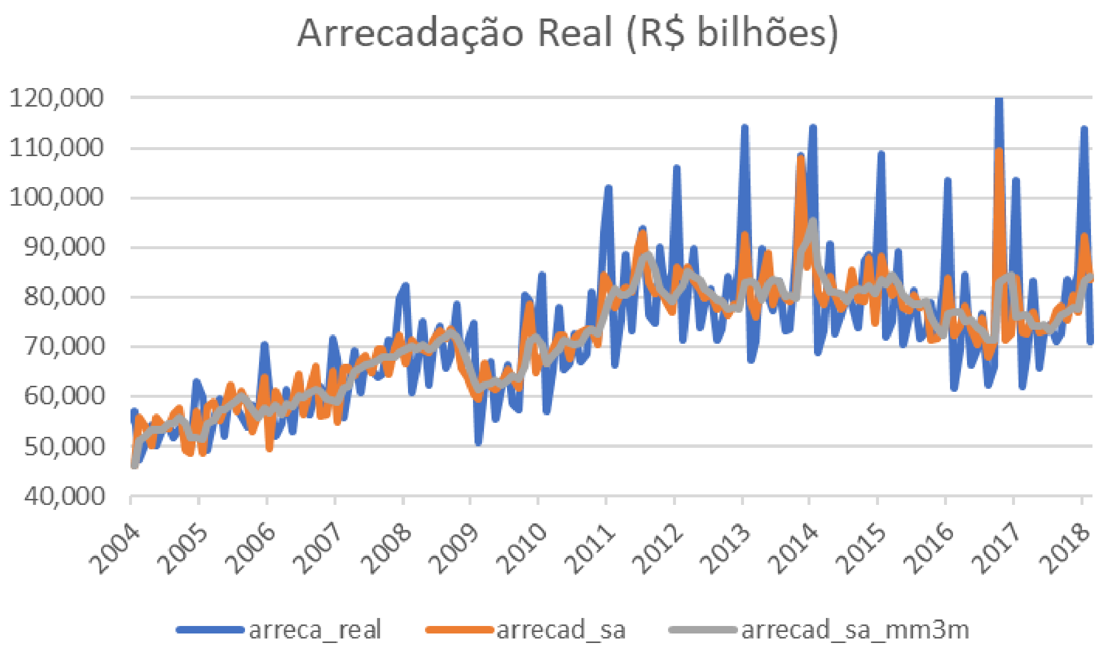
Média móvel tradicional

- O índice v_j corresponde ao S_t da composição original da série y_t e é denominado como fator de escala.
- Com v_j pode-se dizer que a série y é $(s_j - 1)\%$ maior do que a série isenta de sazonalidade.
- O passo final é dividir a série original, y_t pelo índice correspondente ao mês a que pertence t. Formalmente, isso é feito obtendo-se:

$$s_{mq+j} = \frac{y_{mq+j}}{v_j}, m = 0, 1, 2, \ldots, \left[\frac{T-j}{q}\right], j = 1, 2, \ldots, q.$$

Média móvel tradicional - Exemplo

- Exercício no Excel: aula3_1_exercio excel.xlsm
- Arrecadação federal da RFB deflacionada.
- Fazer o ajuste no próprio Excel.



- Média móvel é errado? Sinaliza a tendência

Sazonalidade – Variável binária

- Sazonalidade determinística S_t pode ser escrita como uma função de variáveis binárias sazonais;
- Seja a frequência sazonal:

  s=4 para trimestral;

  s=12 para mensal;

- Variável D_{jt} assume valores igual a 1 se j é igual ao mês/trimestre referente do ano:

  Exemplo: D_{1t}=1 se o mês é janeiro, D_{1t}=0 para os outros meses.

- Estimação é por dada MQO e a regressão pode ser desenhada da seguinte forma: $y_t = \sum_{i=1}^{s} \gamma_i D_{it} + e_t$ ou $y_t = \alpha + \sum_{i=1}^{s-1} \beta_i D_{it} + e_t$
- A diferença é se há constante ou não (multicolinearidade).

Sazonalidade – Variável binária

- Baixando os dados.
- setwd("C:/Users/ytcfa/Desktop/aula/series de tempo/aula3")
- list.files()
- X<-read.csv("aula3_dados.csv",sep=";", dec=".", head=TRUE)
- X$data<-as.Date(X$data,'%d/%b/%y')
- head(X)

Sazonalidade – Variável binária

- Instalando o pacote para estimação sazonal com variável binária

```
install.packages("uroot")
require(uroot)
```

- Tornando o IPCA em uma série temporal

```
ipca_ts<-ts(X$ipca,start = c(1995,1),frequency=12)
ipca_ts1<-ts(X$ipca[121:278],start = c(2005,1),frequency=12)
```

- Criando as variáveis binárias:

```
sd<-seasonal.dummies(ipca_ts)
sd1<-seasonal.dummies(ipca_ts1)
```

Sazonalidade – Variável binária

- Regressão com as variáveis binárias:

```
reg1<-lm(ipca_ts~sd)
summary(reg1)
```

- Regressão com as variáveis binárias sem constante:

```
reg2<-lm(ipca_ts~sd-1)
summary(reg2)
```

- Regressão com as variáveis binárias sem constante:

```
reg3<-lm(ipca_ts1~sd1)
summary(reg3)
```

Sazonalidade – Variável binária

- Fomar de estimar a série ajustada sazonalmente:
- Retirando a média:

```
ipca_ts3=ipca_ts-mean(ipca_ts)
```

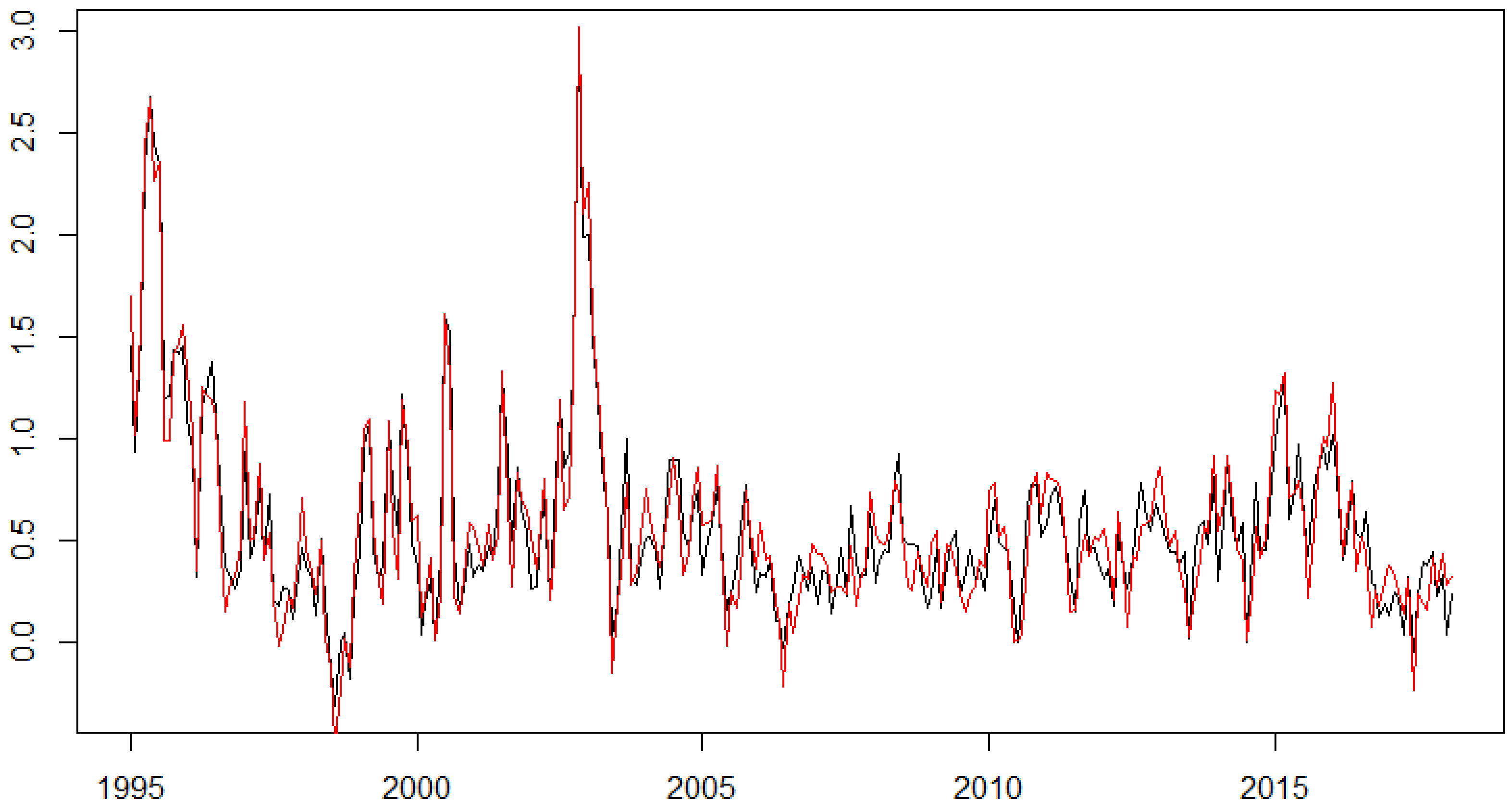
- Estimando a equação:

```
reg3<-lm(ipca_ts3~sd-1sum)
summary(reg3)
```

- Resultado

```
plot(predict(reg3),type="l") #fatores sazonais
ipca_sa=ts(resid(reg3)+mean(ipca_ts),start=c(1995,1),frequency=12)
plot(ipca_sa)
points(ipca_ts,type="l",col="red")
```

Sazonalidade – Variável binária



EWMA – suavização

- Médias móveis exponencialmente ponderadas;
- Consiste em obter uma previsão de forma adaptativa ponderando as observações passadas, desde que ela não tenha tendência nem sazonalidade. A série suavizada x_t é obtida da seguinte forma:

$$\begin{aligned} x_t &= \alpha y_t + \alpha(1-\alpha) y_{t-1} + \alpha(1-\alpha)^2 y_{t-2} + \alpha(1-\alpha)^3 y_{t-3} + \cdots = \\ &= \alpha \sum_{i=0}^{\infty} (1-\alpha)^i y_{t-i}. \end{aligned}$$

- O coeficiente α está entre 0 e 1 e indica a importância das informações mais recentes na definição de x_t. Esses pesos decaem exponencialmente e somam 1:

$$\alpha \sum_{i=0}^{\infty} (1-\alpha)^i = \alpha \times \frac{1}{1-(1-\alpha)} = 1$$

EWMA – suavização

- Retrocedendo x_t para x_{t+1} e multiplicando o resultado por (1-α):

$$(1-\alpha)\,x_{t-1} = \alpha\,(1-\alpha)\,y_{t-1} + \alpha\,(1-\alpha)^2\,y_{t-2} + \alpha\,(1-\alpha)^3\,y_{t-3} +$$

- Subtraindo esse resultado de x_t, obtemos:

$$x_t = \alpha y_t + (1-\alpha)\,x_{t-1}.$$

- O problema nesse caso é encontrar o valor inicial, x_0, para a série suavizada;

- Uma possível saída é a média *y* usando as n observações iniciais para.

EWMA – previsão

- Em termos de previsão, observe que

$$\begin{aligned} x_{t+1} &= \alpha y_{t+1} + (1-\alpha)\, x_t \Longrightarrow \\ x_{t+1} &= x_t + \alpha\left(y_{t+1} - x_t\right). \end{aligned}$$

- Nesse padrão:

$$\begin{aligned} x_{t+2} &= x_{t+1} + \alpha\left(y_{t+2} - x_{t+1}\right) = \\ &= x_t + \alpha\left[\left(y_{t+2} - x_{t+1}\right) + \left(y_{t+1} - x_t\right)\right]. \end{aligned}$$

- Logo:

$$x_{t+k} = x_t + \alpha \sum_{j=0}^{k-1} \left(y_{t+j+1} - x_{t+j}\right).$$

- Numa série estacionária, $\sum_{j=0}^{k-1}\left(y_{t+j+1} - x_{t+j}\right) \simeq 0$, mesmo para valores baixos de k. Considerando simplesmente que $x_{t+k} = x_t$, para todo $k > 0$, a previsão de y_{T+k} será dada por x_T.

Duplo EWMA - previsão

- Para série com tendência linear, sugere-se a dupla suavização, consistindo-se em suavizar a série já suavizada. Isso significa calcular:

$$z_t = \alpha x_t + (1-\alpha)\, z_{t-1}$$

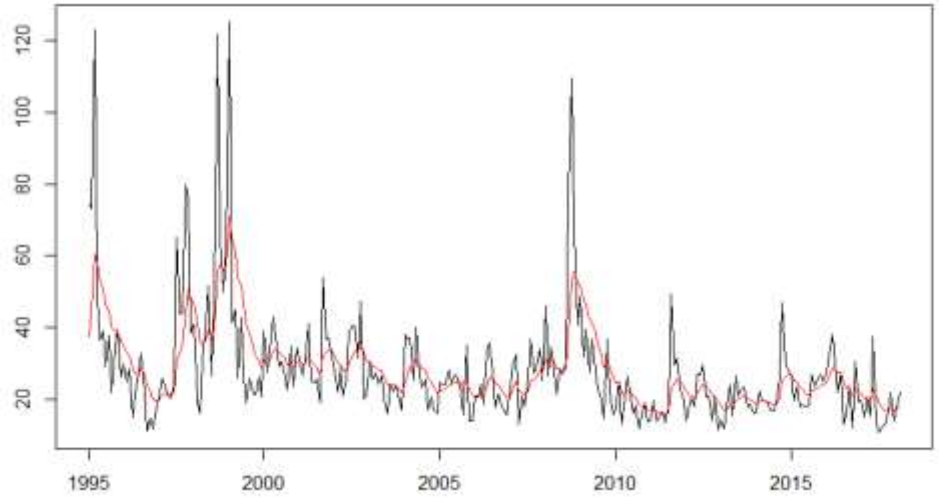
- Nesse caso, é possível construir a previsão da série *y*, prev (y_{t+k}) da seguinte forma:

$$\begin{aligned} prev\left(y_{T+k}\right) &= \left(2+\frac{\alpha k}{1-\alpha}\right) x_T - \left(1+\frac{\alpha k}{1-\alpha}\right) z_T = \\ &= \left(2x_T - z_T\right) + \frac{\alpha k}{1-\alpha}\left(x_T - z_T\right), \end{aligned}$$

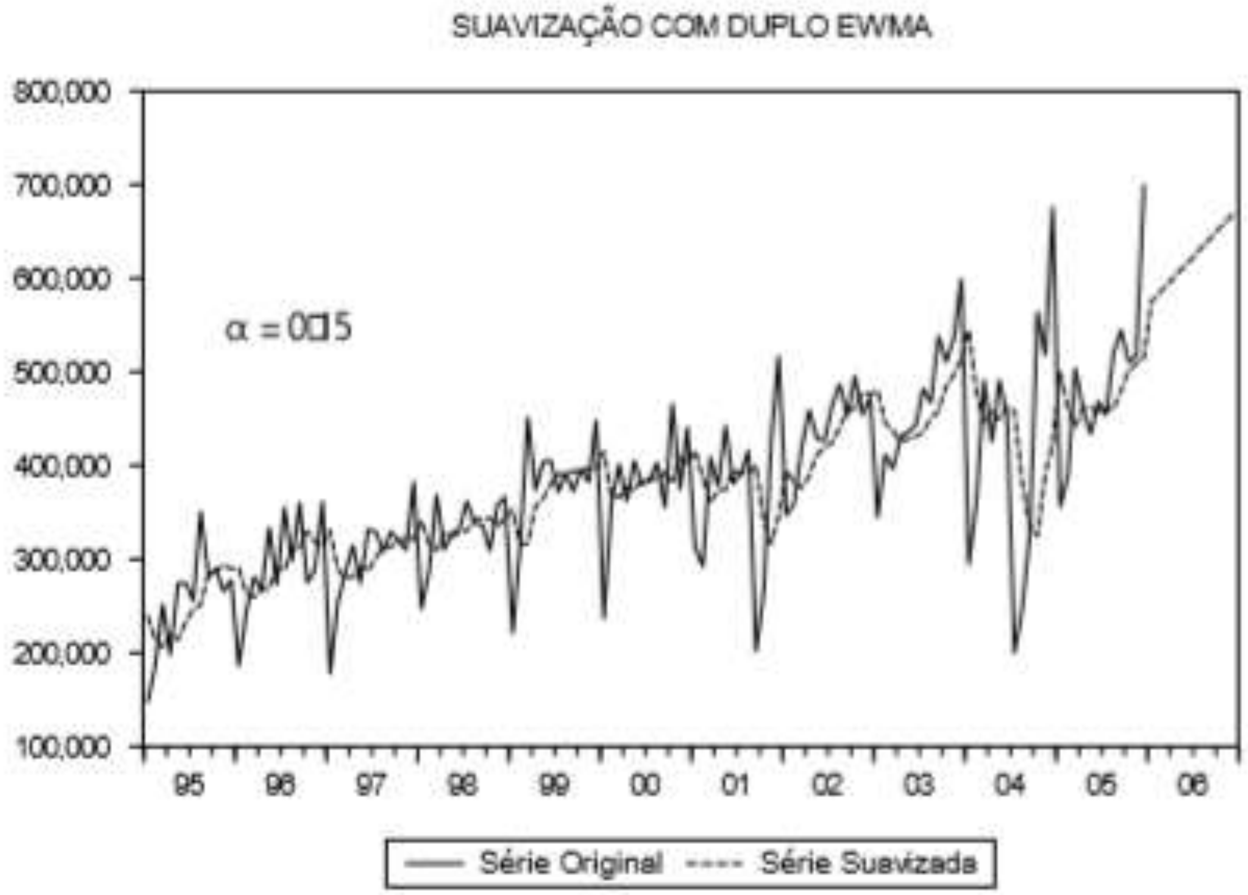
- interpretando-se ($2x_T$ - z_T) como intercepto e $\frac{\alpha}{1-\alpha}\left(x_T - z_T\right)$ como inclinação.

EWMA Suavização - exemplo

- Instalando o pacote
  - install.packages("qcc")
  - require(qcc)
- Primeiro processo de suavização
  - x<-ewma(X$vol_ibov,lambda = 0.2)
  - plot(X$data,X$vol_ibov,type = "l")
  - points(X$data,x$y,type = "l",col="red")

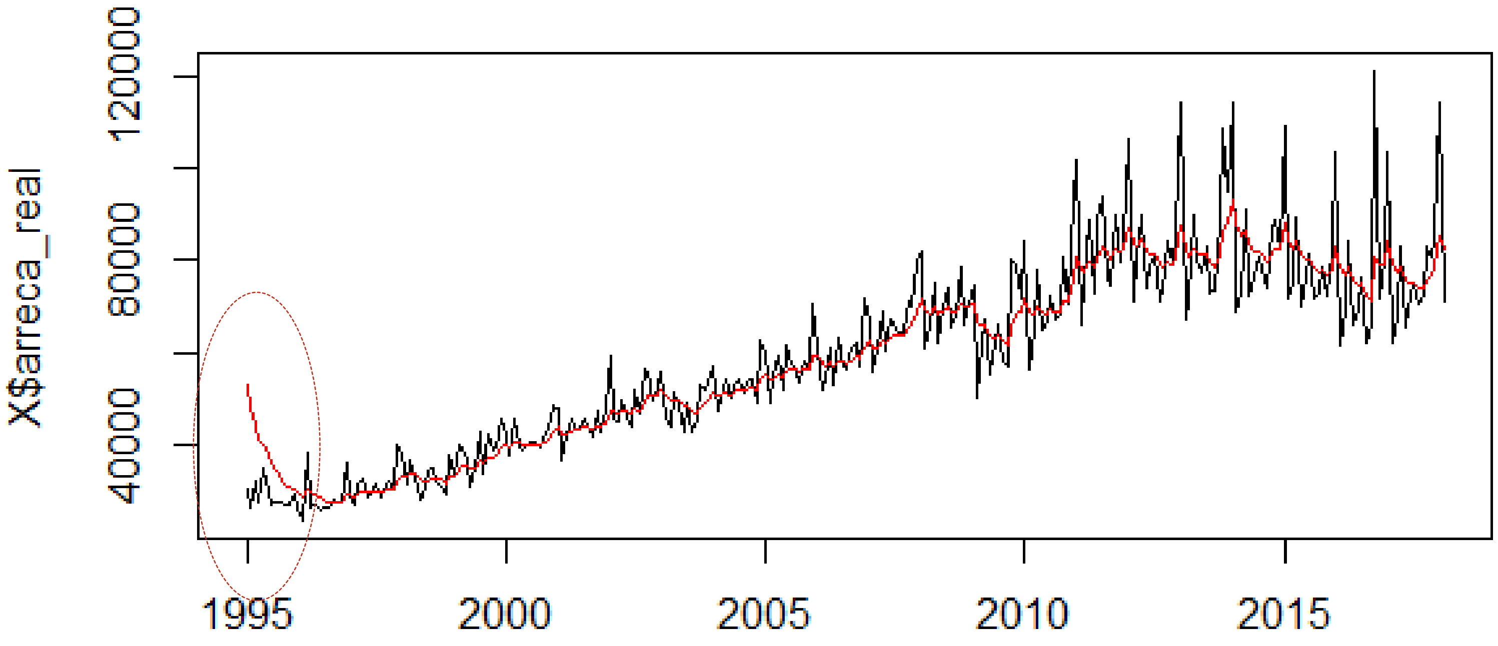


Duplo EWMA – exemplo livro



Duplo EWMA – exemplo prático

- Exercício usando a arrecadação
- Problema no R – não inserir os valores iniciais (usa a média da amostra)



- Mas com o decaimento exponencial, este problema é minimizado com o passar do tempo.

Sazonalidade – ARMA(p,q)(P,Q)$_s$

- Deve-se estimar o modelo e corrigir simultaneamente a sazonalidade em caso de séries univariadas;
  - Muitos modelos estruturais, principalmente os do Banco Central, preferem usar as séries sem ajuste sazonal, colocando *dummies* para captar a sazonalidade;
- É preciso ter cuidado para não dessazonalizar séries sem sazonalidade:
  - Por exemplo: taxas de juros
- Outra característica importante é observar a aderência do modelo para a série em questão;
  - Importante observar se não há sazonalidade nos resíduos;
  - Se as defasagens sazonais são estatisticamente diferentes de zero;
  - Não usar especificações de outras variáveis sem observar a aderência.

Sazonalidade – ARMA(p,q)(P,Q)$_s$

- Usam-se as funções FAC e FACP para identificar a sazonalidade. Suponha

$$y_t = \phi_4 y_{t-4} + \varepsilon_t.$$

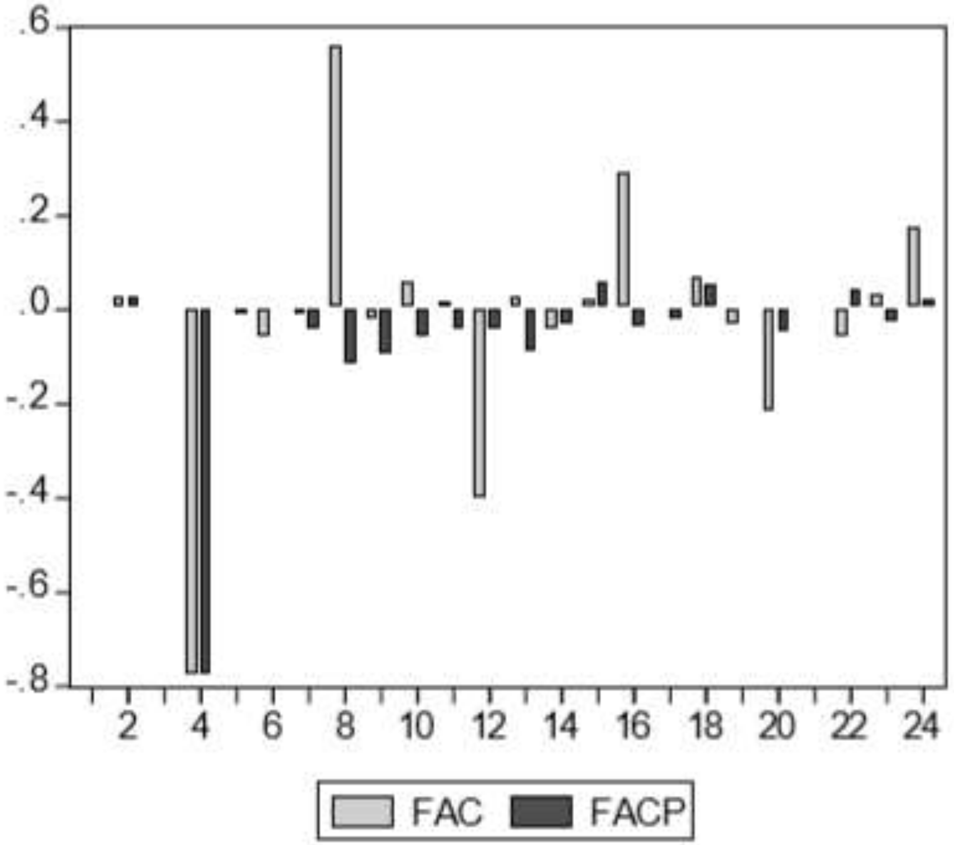
- A FAC desse modeloé tal que: $\rho_i = (\phi_4)^{\frac{i}{4}}$, quando $\frac{i}{4}$ for um número inteiro, exibindo assim um padrão de decaimento exponencial em defasagens múltiplas de 4. A FACP vai apresentar correlação parcial diferente de zero somente na defasagem 4.
- Para:

$$y_t = \varepsilon_t + \theta_4 \varepsilon_{t-4},$$

- a FAC é truncada na defasagem 4, e a FACP tem decaimento sazonal nas defasagens múltiplas de 4.

Sazonalidade – ARMA(p,q)(P,Q)$_s$

- Seguinte modelo: $y_t = \phi_4 y_{t-4} + \varepsilon_t$



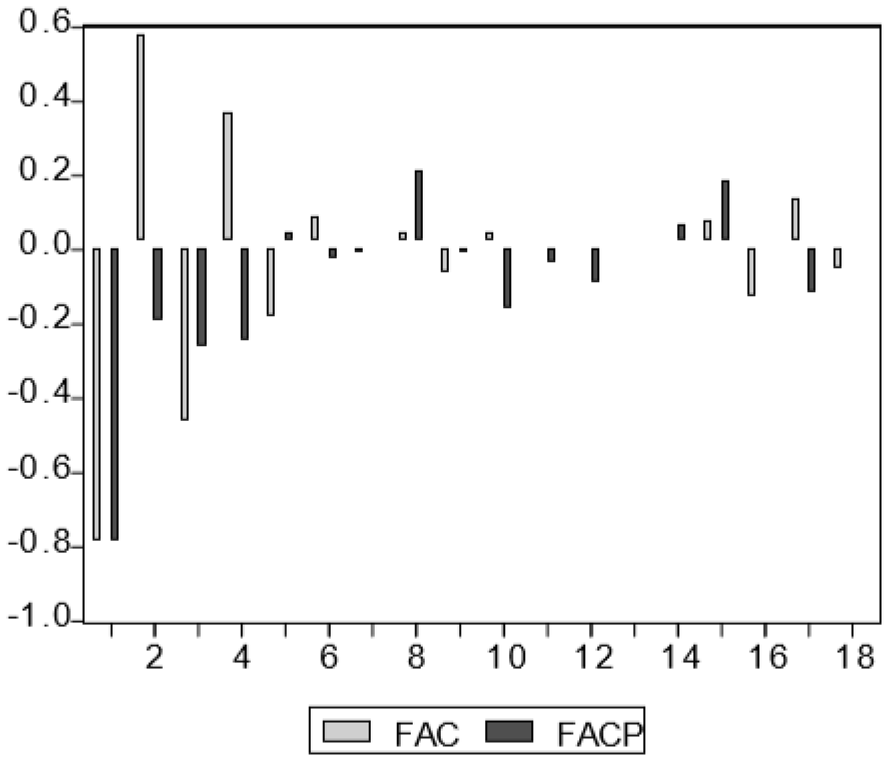
Sazonalidade – ARMA(p,q)(P,Q)$_s$

- Há dois tipos de sazonalidade em séries temporais:
- i) A primeira é a sazonalidade aditiva: um ARMA (1,1) poderá ser sazonal na defasagem 4 via coeficiente autorregressivo ou de médias móveis:

$$
\begin{aligned}
y_t &= \phi_1 y_{t-1} + \phi_4 y_{t-4} + \varepsilon_t + \theta_1 \varepsilon_{t-1} \text{ ou} \\
y_t &= \phi_1 y_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_4 \varepsilon_{t-4}.
\end{aligned}
$$

Sazonalidade – ARMA(p,q)(P,Q)$_s$

- A interação com componentes não sazonais complica a análise.



- No gráfico acima é difícil identificar a sazonalidade de um modelo do tipo ARMA (1,(1,4)):

Sazonalidade – ARMA(p,q)(P,Q)$_s$

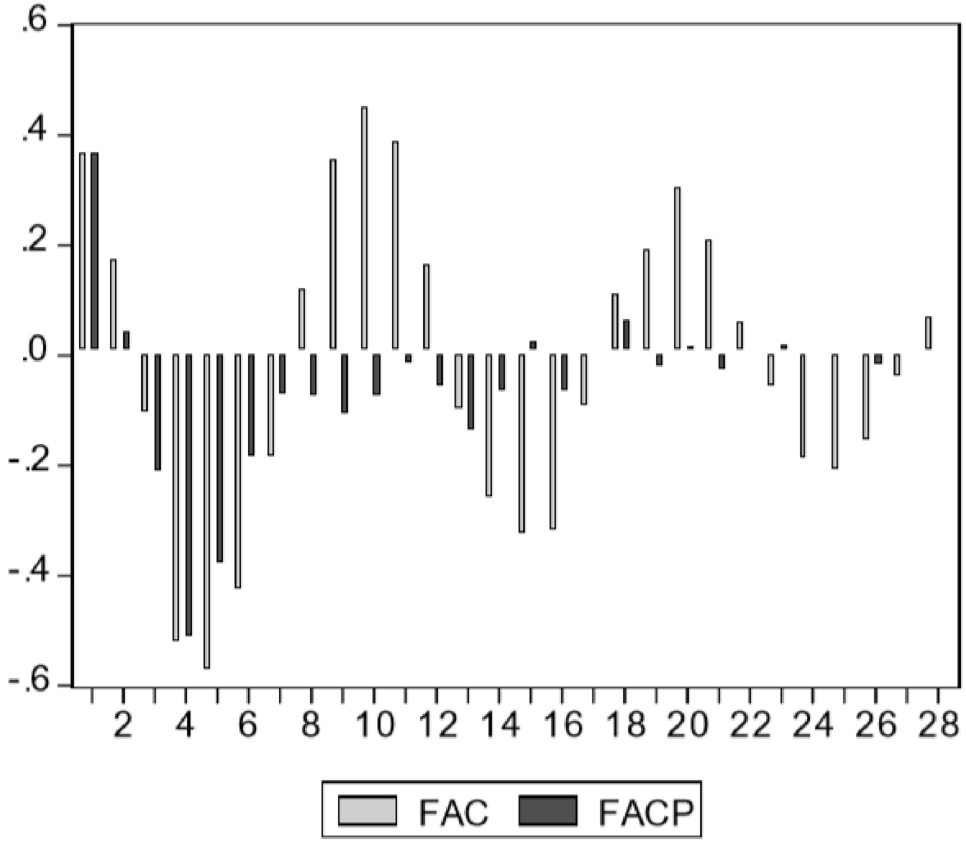
- O outro tipo é a de sazonalidade multiplicativa. Ainda supondo sazonalidade de ordem 4, no caso multiplicativo de um AR e de um MA, teremos respectivamente:

$$
\begin{aligned}
&(1-\phi_1 L)\left(1-\phi_4 L^4\right) y_t = (1+\theta_1 L)\,\varepsilon_t \Longrightarrow \\
&y_t = \phi_1 y_{t-1} + \phi_4 y_{t-4} + \phi_1 \phi_4 y_{t-5} + \varepsilon_t + \theta_1 \varepsilon_t; \\
&(1-\phi_1 L)\, y_t = (1+\theta_4 L)\left(1+\theta_1 L^4\right) \varepsilon_t \Longrightarrow \\
&y_t = \phi_1 y_{t-1} + \varepsilon_t + \theta_1 \varepsilon_t + \theta_4 \varepsilon_{t-4} + \theta_1 \theta_4 \varepsilon_{t-5}.
\end{aligned}
$$

Sazonalidade – ARMA(p,q)(P,Q)$_s$

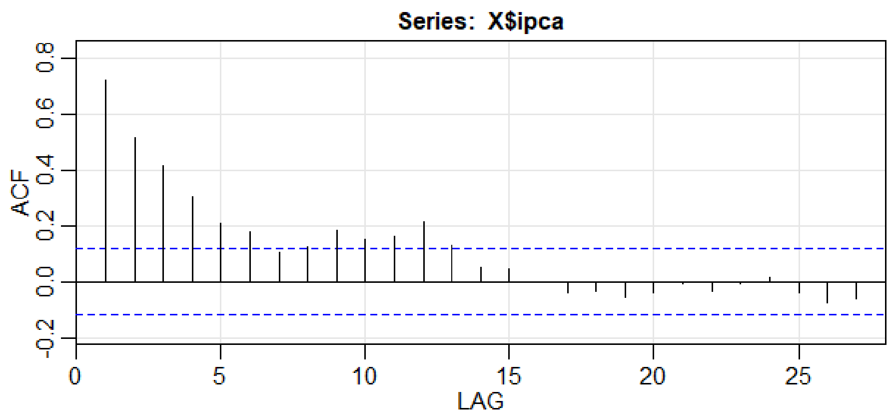
- A Figura a seguir mostra um ARMA(p,q)(P,Q)$_s$. Por exemplo:

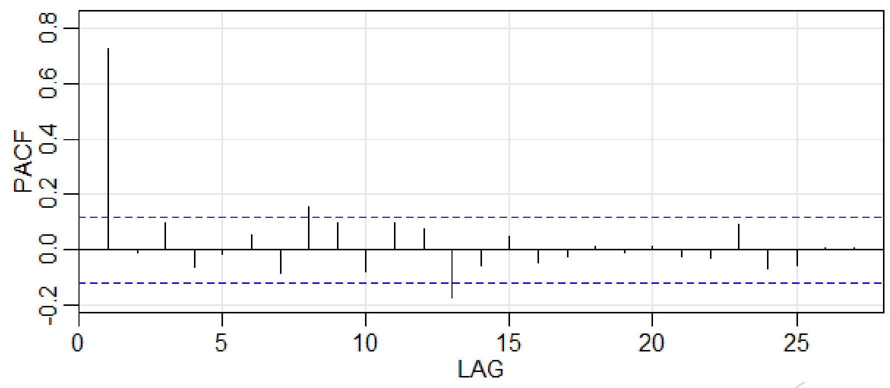
$$ARMA\,(2,1)\,(1,2)_{12} \quad : \quad \left(1-\phi_1 L-\phi_2 L^2\right)\left(1-\phi_{12} L^{12}\right) y_t = \left(1-\theta_1 L\right)\left(1-\theta_{12} L^{12}-\theta_{24} L^{24}\right) \varepsilon_t.$$



Sazonalidade – ARMA(p,q)(P,Q)$_s$

- Instalar o pacote econométrico
  - install.packages("astsa")
  - require(astsa)
- ACF e o PACF da variação do IPCA
  - Acf2(X$ipca)



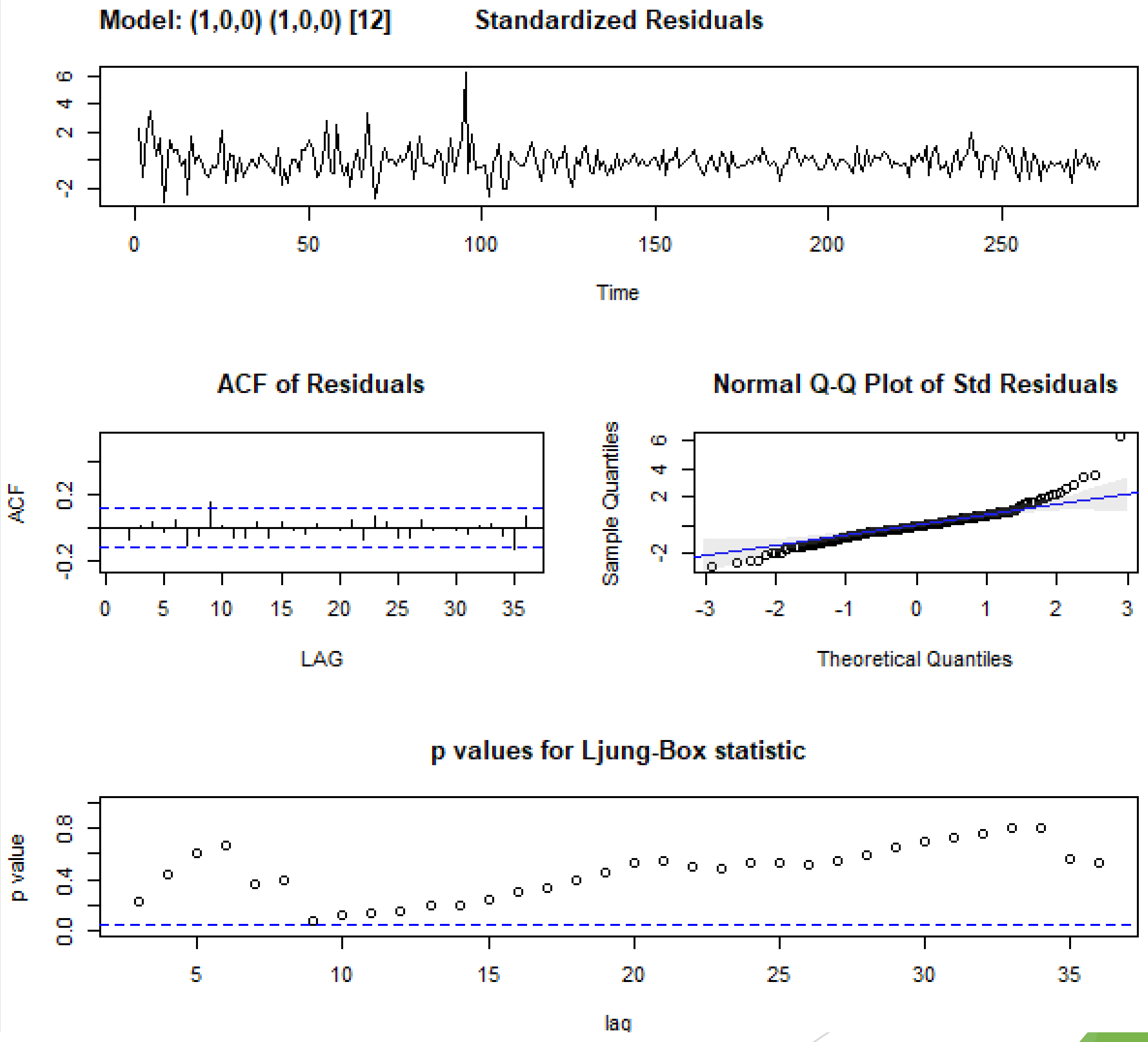


Sazonalidade – ARMA(p,q)(P,Q)$_s$

- Quais modelos possíveis?
  - ARMA(1,0)(1,0)$_{12}$
  - ARMA(1,0)(1,1)$_{12}$
  - ARMA(1,0)(0,1)$_s$
- Comandos para o sarima:
  - x10<-sarima(X$ipca,1,0,0,1,0,0,12)
  - x11<-sarima(X$ipca,1,0,0,1,0,1,12)
  - x01<-sarima(X$ipca,1,0,0,0,0,1,12)

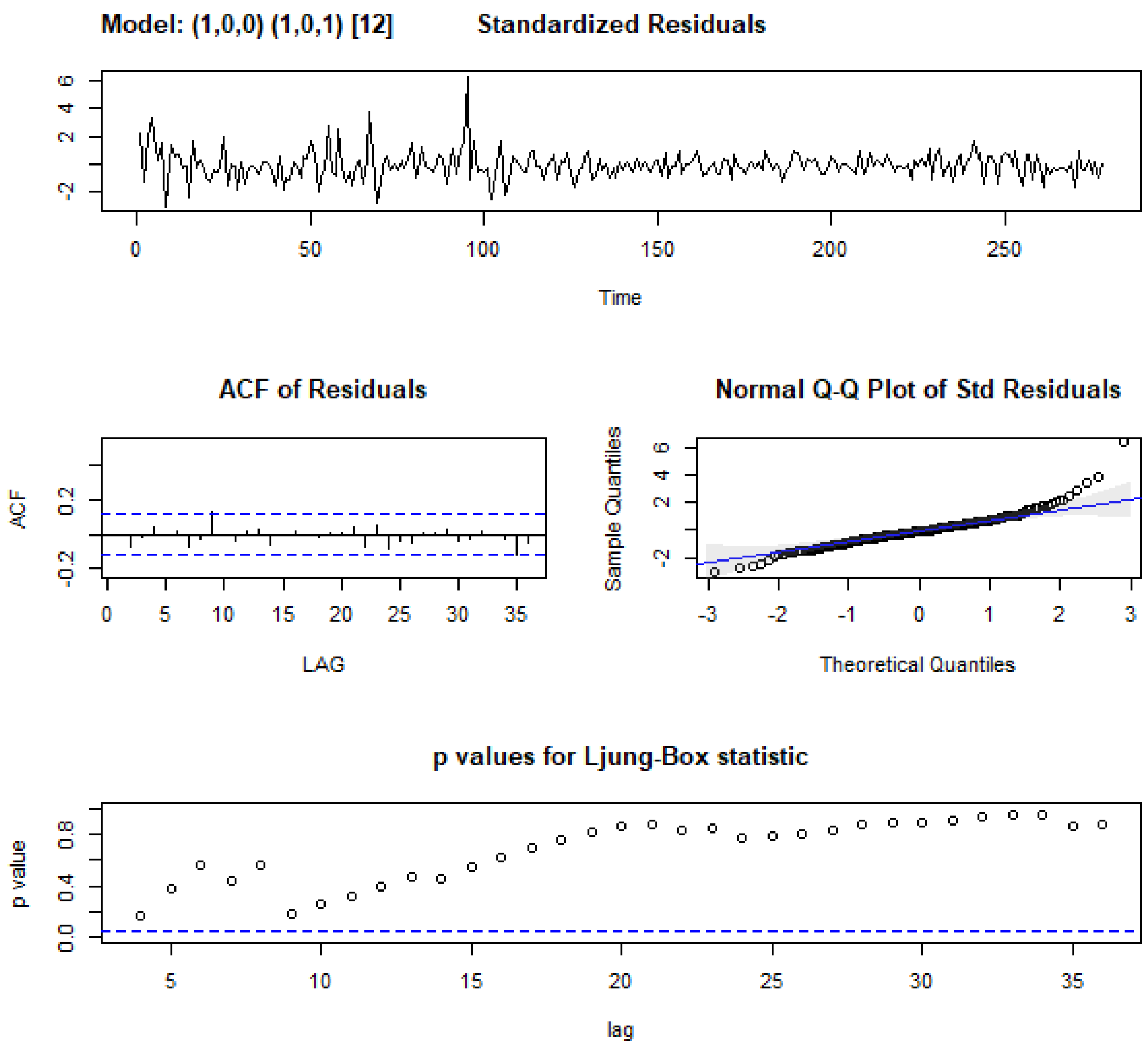
Sazonalidade – ARMA(p,q)(P,Q)$_s$

```
> x10$ttable
      Estimate     SE t.value p.value
ar1     0.7380 0.0406 18.1795       0
sar1    0.2997 0.0613  4.8858       0
xmean   0.5953 0.0973  6.1147       0
```



Sazonalidade – ARMA(p,q)(P,Q)$_s$

```
> x10$ttable
       Estimate     SE t.value p.value
ar1      0.7519 0.0399 18.8258  0.0000
sar1     0.8343 0.1754  4.7576  0.0000
sma1    -0.6409 0.2571 -2.4931  0.0133
xmean    0.6096 0.1398  4.3606  0.0000
```

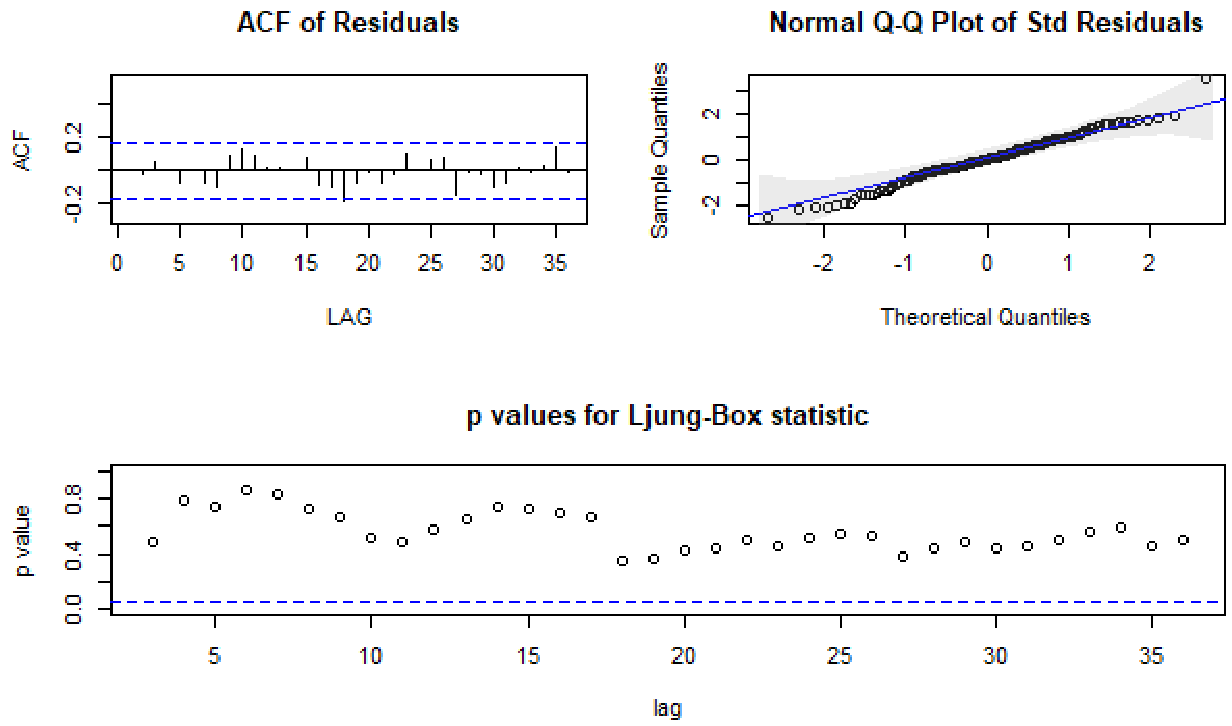


Sazonalidade – ARMA(p,q)(P,Q)$_s$

- Quais os principais problemas que os testes apontam?
  - Normalidade – caudas mais gordas
  - Heterocedasticidade
  - Possível autocorrelação na defasagem 9
- Diagnóstico:
  - Quebra de padrão – diferenças relevantes dos resíduos 0-100 e 100-278;
- Soluções possíveis:
  - Dummies: i) correção do período anterior a 100; ii) para meses com inflação significativamente diferente do padrão histórico;
  - Markov-switch
  - Quebra estrutural

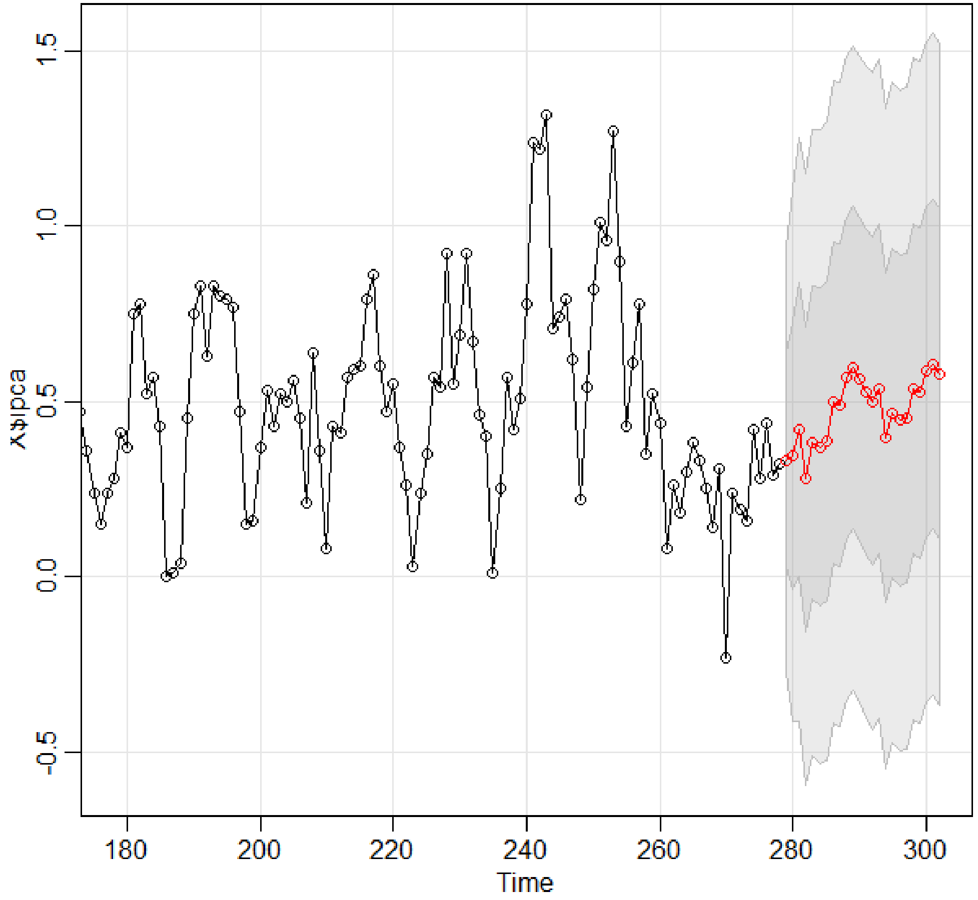
Sazonalidade – ARMA(p,q)(P,Q)$_s$

- Dividindo a amostra
  - x10<-sarima(X$ipca[139:278],1,0,0,0,0,1,12)
- Defasagem 9 não é significante
- Erros normais;



Sazonalidade – ARMA(p,q)(P,Q)$_s$

- Projeção
  - for1<-sarima.for(X$ipca,24,1,0,0,1,0,1,12)



Sazonalidade – ARMA(p,q)(P,Q)$_s$

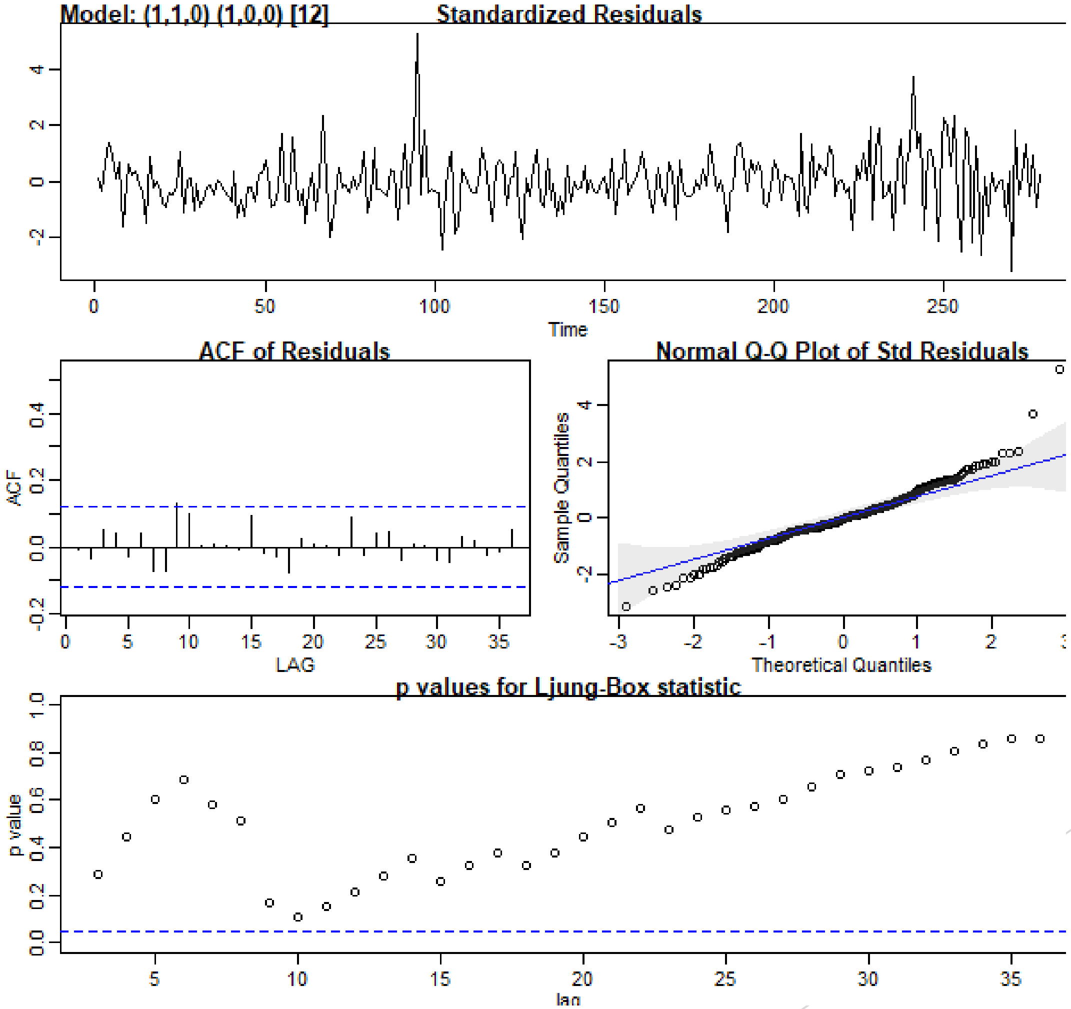
- Comparando com a estimação do período mais curto:
  - for2<-sarima.for(X$ipca[139:278],24,1,0,0,1,0,1,12)

```
> for1$pred
Time Series:
Start = 279
End = 302
Frequency = 1
 [1] 0.3322563 0.3455570 0.4196429 0.2776883 0.3820451 0.3701279 0.3884798 0.4972985
 [9] 0.4902366 0.5667308 0.5947132 0.5646205 0.5260661 0.5004835 0.5347141 0.3955392
[17] 0.4670118 0.4453429 0.4518369 0.5359966 0.5251193 0.5851913 0.6057188 0.5784922
> for2$pred
Time Series:
Start = 141
End = 164
Frequency = 1
 [1] 0.3693451 0.3167601 0.3658594 0.2525477 0.3820670 0.3629170 0.3992754 0.4739448
 [9] 0.4450288 0.4792847 0.4293596 0.4326355 0.4403133 0.4470303 0.4497519 0.4544402
[17] 0.4534323 0.4548505 0.4547998 0.4537484 0.4546162 0.4541200 0.4552424 0.4552577
```

Sazonalidade – ARMA(p,q)(P,Q)$_s$

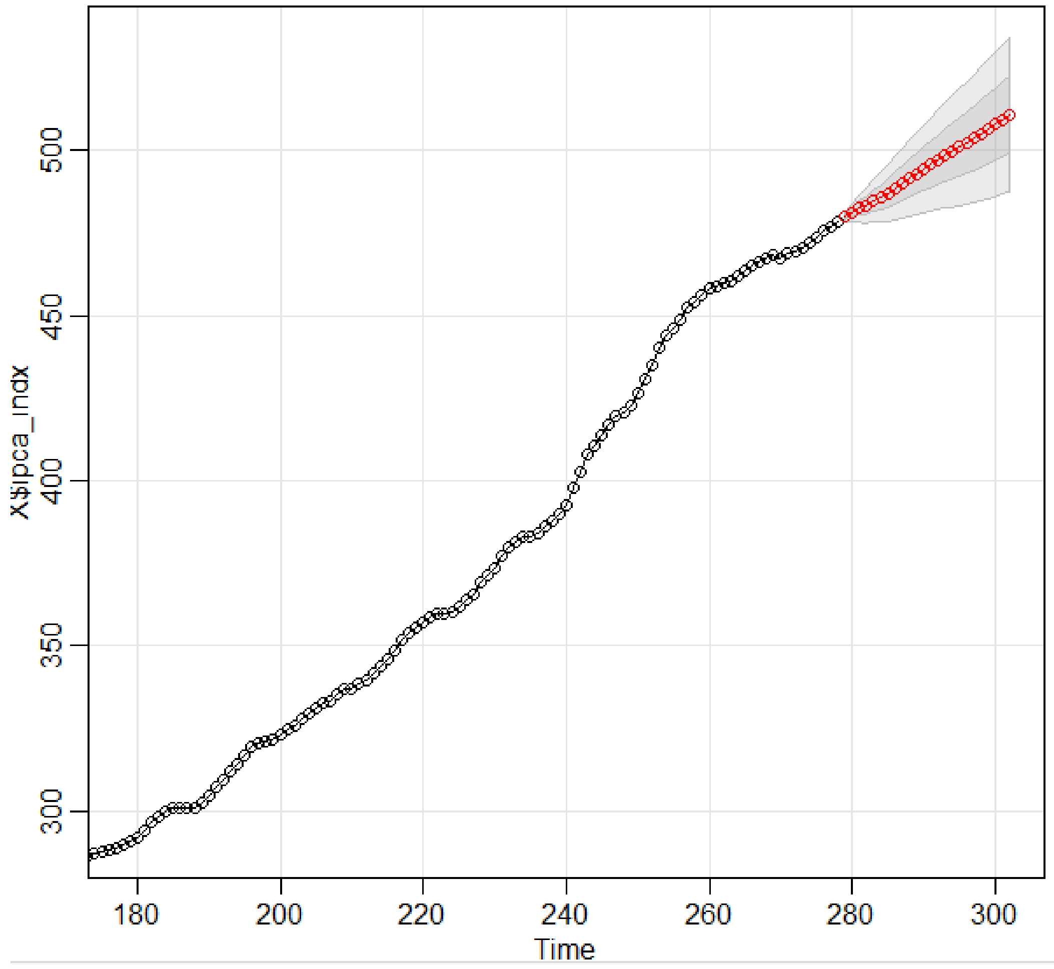
- Estimando um SARIMA $(1,1,0)(1,0,0)_{12}$:
  - x10i<-sarima(X$ipca_indx,1,1,0,1,0,0,12)

```
> x10i$ttable
         Estimate     SE t.value p.value
ar1        0.7032 0.0427 16.4849       0
sar1       0.2686 0.0602  4.4600       0
constant   1.3675 0.1836  7.4489       0
```



Sazonalidade – ARMA(p,q)(P,Q)$_s$

- Projetando um SARIMA $(1,1,0)(1,0,0)_{12}$:
  - sarima.for(X$ipca_indx,24,1,1,0,1,0,0,12)



Census X13: Teoria

- O X13 foi desenvolvido pelo U.S. Census Bureau com o apoio do Banco de España.
- O programa, criado em julho de 2012, é a junção de outros dois programas de ajuste sazonal X12-ARIMA (Findley et al, 1998) e TRAMO/SEATS (Gómez & Maravall, 1996).
- Além de dessazonalizar séries, o programa oferece diversos diagnósticos com o intuito de avaliar a qualidade do ajuste sazonal
- A análise gráfica de uma série temporal permite visualizar suas características para uma boa modelagem, por exemplo: seu padrão sazonal, quebras estruturais, possíveis outliers, se há necessidade (e possibilidade) de usar transformação logarítmica nos dados para estabilizar a variância, etc.

Census X13: Teoria

- Formas de decomposição:
- Legenda: O – original; TC – Tendência/ciclo; S – Sazonal; I - Irregular
- Aditivo: $O_t = TC_t + S_t + I_t$ e $SA_t = O_t - S_t$

Quando usar: i) sazonalidade estável entre os meses; ii) independência dos componentes – tendência não há elevação da sazonalidade;

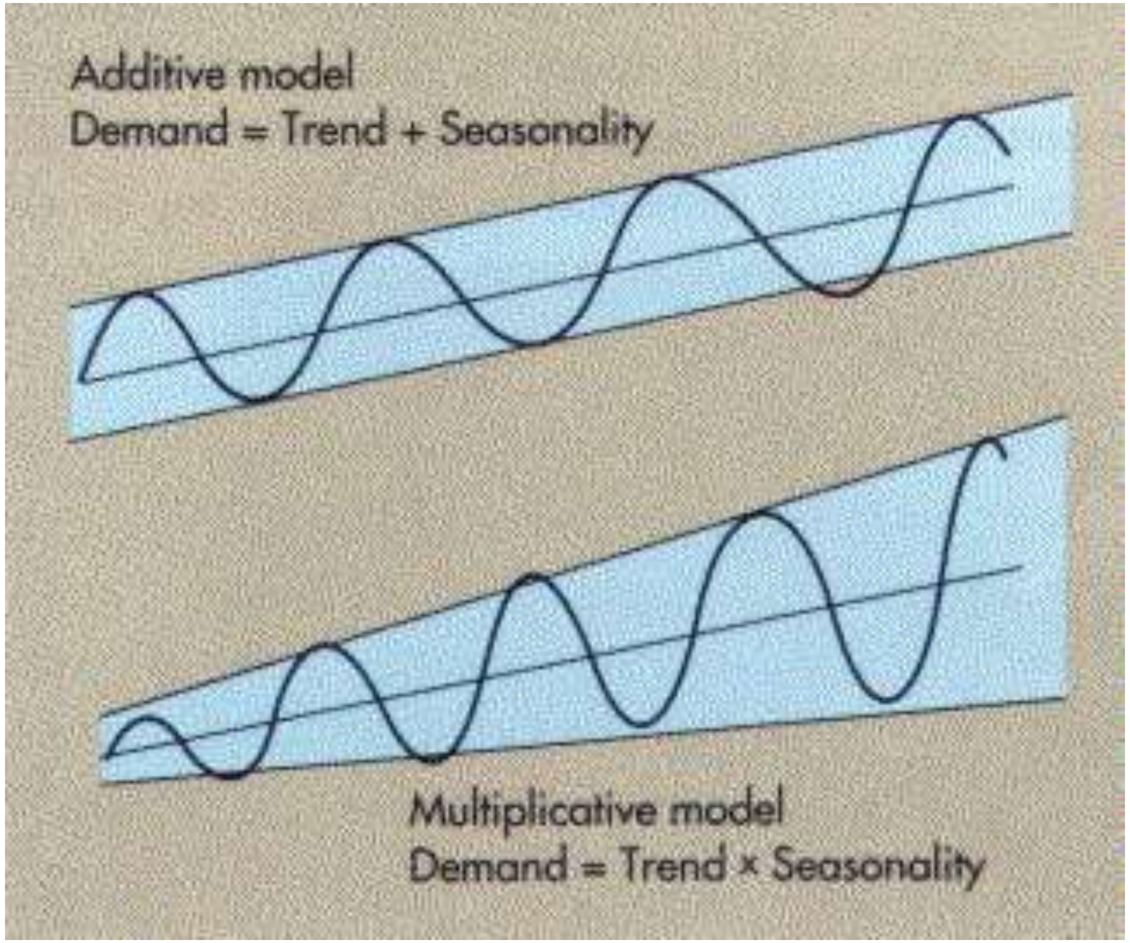
- Multiplicativo $O_t = TC_t \times S_t \times I_t$ e $SA_t = O_t/S_t,$

Quando usar: i) quando a amplitude da sazonalidade não é regular; ii) sazonalidade relacionada com a tendência;

- Pseudo-aditivo: $O_t = TC_t \times (S + I - 1)$

Quando usar: os componentes sazonalidade e irregular são independentes, mas se relacionam com a tendência.

Census X13: Teoria

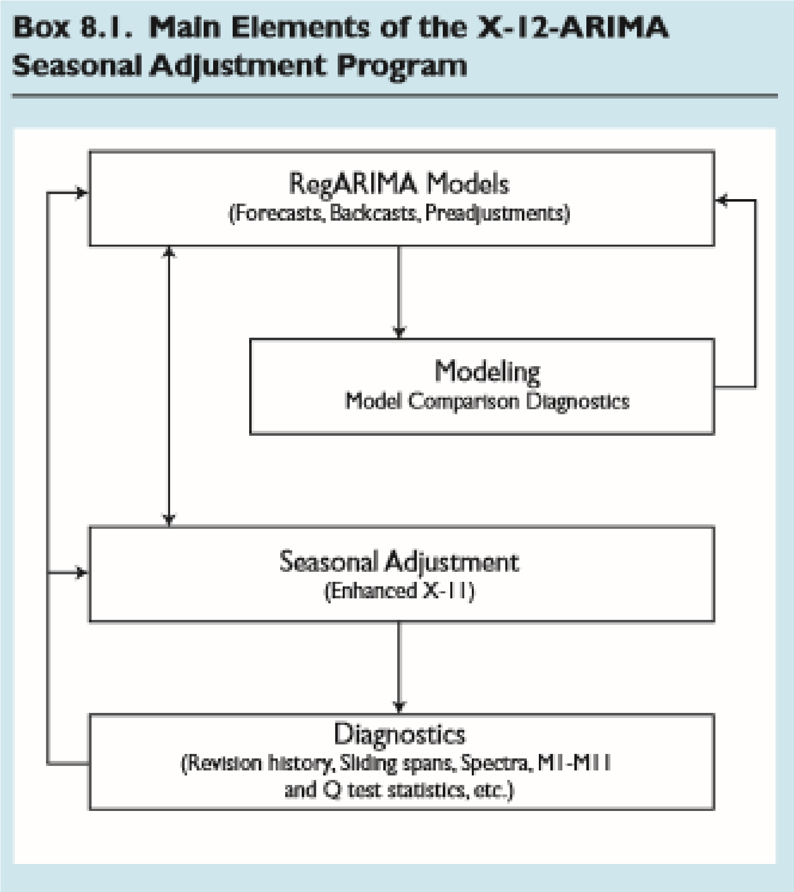


Census X13: Teoria

- Não há solução única para o modelo de ajuste sazonal;
- Sempre há revisão dos dados passados quando são divulgadas novas informações;
- Um bom ajuste dos dados, significa que os resíduos não tem sazonalidade;

Census X13: Teoria X-12

- O primeiro passo é ajustar as séries de efeitos de *outliers*, mudança de nível e eventos irregulares (calendário);
- Segundo passo é filtrar a variável pré ajustada (X-11)
- Terceiro passo: vários diagnósticos de qualidade para aprovação do resultado final.



Census X13: Teoria X-12 (filtro)

- Estimação Inicial (trimestral)
  - Média móvel centrada: $T_t^1 = 1/8 X_{t-2} + 1/4 X_{t-1} + 1/4 X_t + 1/4 X_{t+1} + 1/8 X_{t+2}$
  - Razão inicial SI: a série original é divida pela série T^1 resultando em saz. e irregular.
  - Fator sazonal inicial (média móvel): $\hat{S}_t^1 = 1/9 SI_{t-8} + 2/9 SI_{t-4} + 3/9 SI_t + 2/9 SI_{t+4} + 1/9 SI_{t+8}$
  - Normalização: $S_t^1 = \dfrac{\hat{S}_t^1}{\frac{1}{8}\hat{S}_{t-2}^1 + \frac{1}{4}\hat{S}_{t-1}^1 + \frac{1}{4}\hat{S}_t^1 + \frac{1}{4}\hat{S}_{t+1}^1 + \frac{1}{8}\hat{S}_{t+2}^1}$
  - Estimação inicial da série sazonal: $A_t^1 = X_t / S_t^1 = T_t \cdot S_t \cdot I_t / S_t = T_t^1 \cdot I_t$.
- Estimativas revisitadas
  - Estima-se T^2 pela média móvel de Henderson da estimação A^1
  - Refaz os passos acima, com os valores revistados.

Census X13: Teoria X-12 (diagnóstico)

- Testes principais devem ser observados: i) F-test para a presença de sazonalidade; ii) M- e Q-testes
- Exemplo:

D8.A F-Tests for Seasonality

Test for the Presence of Seasonality Assuming Stability

	Sum of Squares	Degrees of Freedom	Mean Square	F-Value
Between quarters	809.1996	3	269.73319	43.946**
Residual	497.1645	81	6.13783	
Total	1306.3640	84		

Seasonality present at the 0.1 percent level.

Nonparametric Test for the Presence of Seasonality Assuming Stability

Kruskal-Wallis Statistic	Degrees of Freedom	Probability Level
53.2410	3	0.000%

Seasonality present at the 1 percent level.

- Kruskall-Wallis: teste não paramétrico para testar se a variância de duas séries são estatisticamente diferentes.

Census X13: Teoria X-12 (diagnóstico)

The Q-test statistic at the bottom is a weighted average of the M-test statistics.

F 3. Monitoring and Quality Assessment Statistics
All the measures below are in the range from 0 to 3 with an acceptance region from 0 to 1.

Statistics	Weight in Q	Value
1. The relative contribution of the irregular component over a one-quarter span (from Table F 2.B).	13	M1 = 0.245
2. The relative contribution of the irregular component to the stationary portion of the variance (from Table F 2.F).	13	M2 = 0.037
3. The amount of quarter-to-quarter change in the irregular component compared with the amount of quarter-to-quarter change in the trend-cycle (from Table F2.H).	10	M3 = 0.048
4. The amount of auto-correlation in the irregular as described by the average duration of run (Table F 2.D).	5	M4 = 0.875
5. The number of quarters it takes the change in the trend-cycle to surpass the amount of change in the irregular (from Table F 2.E).	11	M5 = 0.200
6. The amount of year-to-year change in the irregular compared with the amount of year-to-year change in the seasonal (from Table F 2.H).	10	M6 = 0.972
7. The amount of moving seasonality present relative to the amount of stable seasonality (from Table F 2.I).	16	M7 = 0.378
8. The size of the fluctuations in the seasonal component throughout the whole series.	7	M8 = 1.472
9. The average linear movement in the seasonal component throughout the whole series.	7	M9 = 0.240
10. Same as 8, calculated for recent years only.	4	M10= 1.935
11. Same as 9, calculated for recent years only.	4	M11= 1.935

ACCEPTED at the level 0.52 Check the three above measures that failed. Q (without M2) = 0.59 ACCEPTED.

Tramo-Seats: Teoria

- TRAMO (Time Series Regression with ARIMA Noise, Missing Observations, and Outliers) – é um programa para estimação, projeção e interpolação dos modelos ARIMA.
- SEATS (Signal Extraction in ARIMA Time Series) – decompõe a série temporal em componentes não observáveis como tendência, sazonalidade e irregular.
- Processo semelhante ao X-12
  - Ajustar as séries, escolher o modelo e verificar os resultados dos resíduos.
- Diferente do X-12, este é um teste paramétrico e pode ser analisado estatisticamente.
- Primeiro processo: ajustar as séries
  - Efeitos calendário – feriados específicos e recorrentes
  - Detectar outliers

Tramo-Seats: Teoria

- Escolha de modelo

Predefined specifications in TRAMO/SEATS

Name	Explanation
RSA0	level, Airline model
RSA1	log/level, outlier detection, airline model
RSA2	log/level, working days, Easter, outlier detection, airline model
RSA3	log/level, outlier detection, automatic model identification
RSA4	log/level, working days, Easter, outlier detection, automatic model identification
RSA5	log/level, trading days, Easter, outlier detection, automatic model identification

- A menos que ele encontre um modelo na identificação automática, ele usar o Airline model.
- Normalmente o model escolhido no RSA0 é um (0,1,1)(0,1,1)

Tramo-Seats: Teoria

- Escolha do modelo

Options	Set	Definition
Transformation→ Function	Auto	Program will test for log/level specification
Calendar effects→ Trading days →Type	Predefined	
Details→ Trading days	td2	Working-day and leap year
Pretest	True	Program tests whether the effect is significant or not.
Easter effect → option Duration.	Pretest 6	Program tests whether the effect is significant or not. Default duration of Easter
Arima modelling → Automatic modelling → Enabled	True	The program automatically identifies the orders of ARIMA model
Outlier detection → Enabled	True	The program automatically detects the outliers

Tramo-Seats: Teoria

- Diagnóstico – o que é necessário:
  - Propriedades do modelo escolhido e os componentes;

    annual totals (somatório é igual a zero); normalidade; definition – relação entre os componentes (tendência, irregular e sazonal);

  - Número e tipos de *outliers*;

    número de *outliers;*

  - Estabilidade do componente sazonal;

    Estabilidade do modelo (coeficientes); *Sliding spans* (mudança no componente sazonal devido os *outliers*);

  - Falta de sazonalidade nos resíduos e não há efeito calendário;

    Friedman e Kruskall-Wallis teste;

Census X13: Teoria

- Análise dos resíduos:
  - QS é a estatística que teste a hipótese de sazonalidade residual.
  - Analisa-se a série original (transformada, quando for o caso) e um período mais curto (EV adj – *extreme value adjusted*).

```
QS statistic for seasonality:
Original Series                              77.73  (P-Value =     0.0000)
Original Series (EV adj)                     77.73  (P-Value =     0.0000)
Residuals                                     0.01  (P-Value =     0.9953)
Seasonally Adjusted Series                    0.00  (P-Value =     1.0000)
Seasonally Adjusted Series (EV adj)           0.00  (P-Value =     1.0000)
Irregular Series                              0.00  (P-Value =     1.0000)
Irregular Series (EV adj)                     0.00  (P-Value =     1.0000)
```

- É calculado se há autocorrelação das séries.

Census X13: R-Studio

- Instalando o X13
  - install.packages(“seasonal")
  - require(seasonal)
- Transformando o dataframe em série de tempo
  - arrecad=ts(X$arreca_real,start=c(1995,1),frequency=12)
- Assitente gráfico #auxilia a escrever a função do modelo
  - install.packages("seasonalview")
  - require(seasonalview)

Census X13: R-Studio – Seats

- Especificando o modelo a ser estimado
  - mdlseats=seas(x=arrecad,transform.function = "log")
- Resumo da estimação
  - summary(mdlseats)
- Estatísticas de adequação do modelo
  - qs(mdlseats)
- Série ajustada sazonalmente
  - mdlseats_sa=mdlseats$series$s11
  - plot(mdlseats_sa)
- Fatores sazonais
  - plot(mdlseats$series$s18

Name	Small	Description of table
trend	s12	final SEATS trend component
seasonal	s10	final SEATS seasonal component
irregular	s13	final SEATS irregular component
seasonaladj	s11	final SEATS seasonal adjustment component
transitory	s14	final SEATS transitory component
adjustfac	s16	final SEATS combined adjustment factors
adjustmentratio	s18	final SEATS adjustment ratio

Census X13: R-Studio – Seats

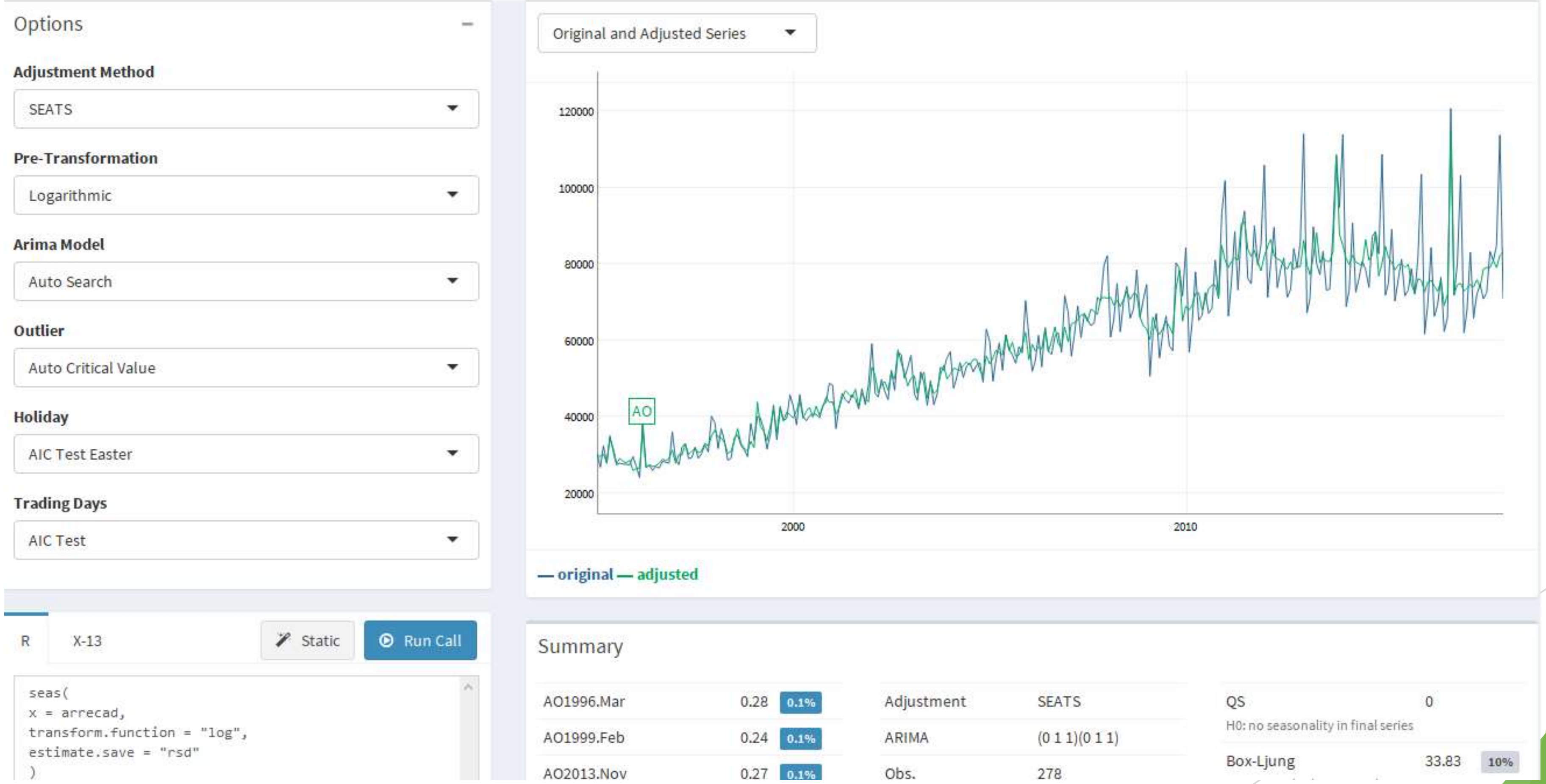
- Outras opções:
- Cinco melhores modelos:
  - mdlseats$fivebestmdl
- Salvar e apresentar os resíduos
  - mdlseats _e=mdlseats$series$rsd
  - plot(mdlseats _e)
- Apresenta os coeficientes da equação
  - mdlseatsestreg
  - mdlseatsestarima

Census X13: R-Studio – X11

- Especificando o modelo a ser estimado
  - mdlseats=seas(x=arrecad,transform.function = "log",x11="")
- Resumo da estimação
  - summary(mdlseats)
- Análise dos resíduos
  - teste<-udg(mdlx11)
  - Kruskal-Wallis – teste$f2.kw
  - Teste F – teste$f2.f
  - teste$f3.m01 ... 12
  - Teste$f3.q
- Comandos são semelhantes ao Seats

Census X13: R-Studio – Interface gráfica

- Visualizar graficamente e alterações do modelo
- view(mdlseats)



Resumo – métodos de ajuste sazonal

Método	Positivo	Negativo
Média móvel tradicional	Fácil de executar, pode ser feito no excel	Pressupõe fatores sazonais constantes
Variável binária	Estimado por MQO e adicinar a equações com variáveis explicativas	Fatores constantes e a hipótese de não correlação das variáveis binárias com alguma variável omitida
Suavização - EWMA	Fácil execução	Deterimnar o valor de suavização e os valores iniciais
ARMA sazonal	Fatores sazonais variáveis	Não há um modelo único
X-12	Fatores sazonais variáveis e opção de escolha automática	Testes são *ad-hoc*
Tramo	Fatores sazonais variáveis e análise estatística	pressupõe normalidade dos resíduos

Exercício

- Calcular o CAGED usando o modelo com variáveis binárias;
- Fazer o ajuste sazonal do saldo da balança comercial brasileira usando o modelo ARMA sazonal;
- Calcular a série de exportação de bens brasileiros usando o X-12 e o TRAMO-SEATS;
  - Analisar os resultados!

ADICIONANDO VARIÁVEIS EXPLICATIVAS

Modelos com defasagens degeneradas

- Incluir no modelo:
  - Defasagens – garantindo que o resíduo seja um RB;
  - *Dummies* sazonais, caso a série não seja ajustada sazonalmente;
  - Variáveis explicativas que tenham poder preditivo da variável dependente.
- Modelo geral:

$y_t = \alpha + \delta_1 y_{t-j} + \delta_2 y_{t-k} + \beta X_t + \gamma D_t + e_t$ para j e k >0

- Importante: resíduo é um RB!

Modelos com defasagens degeneradas

- Exemplo:
- Instalando o pacote

```
install.packages(“stats") #pacote para na estimação e projeção
require(stats)
install.packages(“dyn") #pacote para na estimação e projeção
require(dyn)
```

- Transformando os dataframes em séries temporais

```
caged=ts(X$caged[97:276],start = c(2003,1),frequency = 12)
pim=ts(X$pim_yoy[97:276],start = c(2003,1),frequency = 12)
```

- Criando as *dummies* temporais

```
dados=cbind(pim,seasonal.dummies(caged))
```

Modelos com defasagens degeneradas

- Primeira estimação:

  x<-arima(caged,order=c(1,0,0),xreg=dados,include.mean = FALSE)

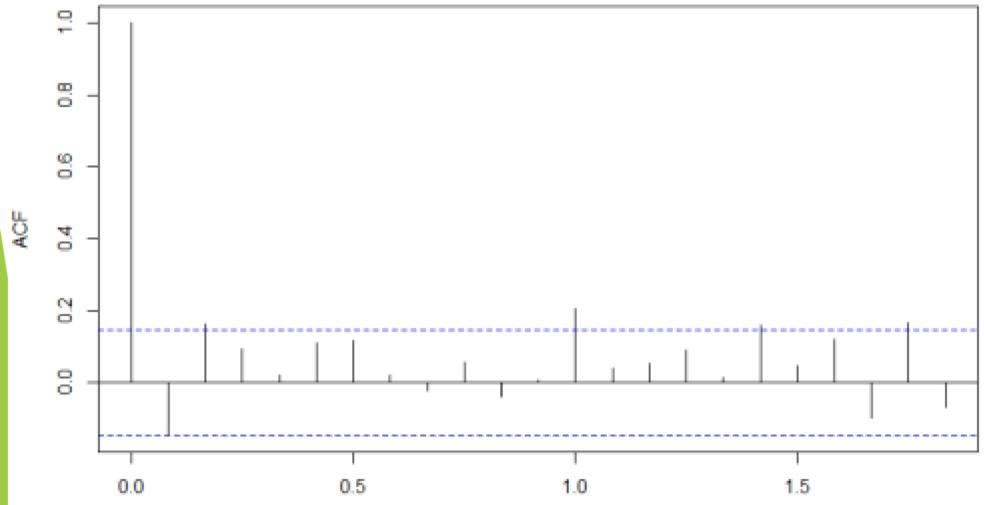
- Resultado da estimação

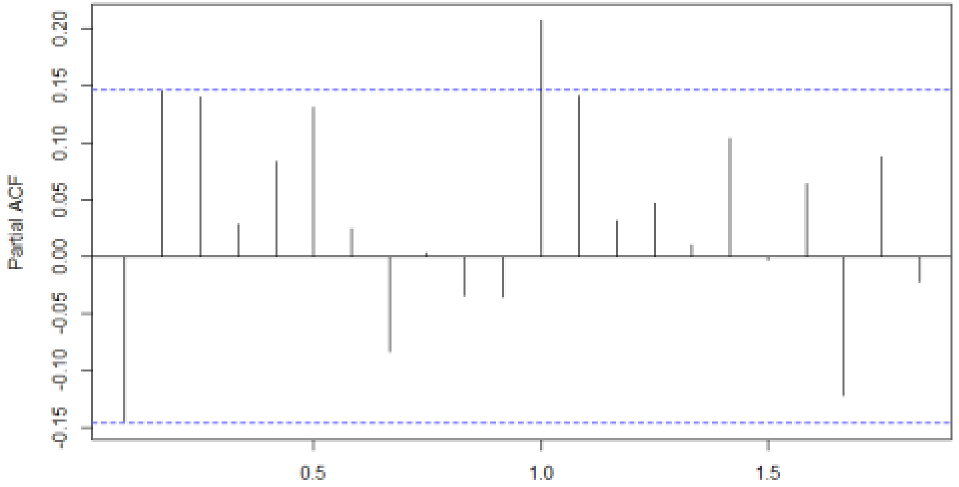
  x<enter>

- Verificando os resíduos

  acf(x$residuals)

  acf(erro); pacf(erro)





Modelos com defasagens degeneradas

- Segunda estimação:

  x1=arima(caged,order=c(2,0,0),seasonal=c(1,0,0), xreg=dados,include.mean = FALSE)

  Ou x1<-dyn$lm(caged~lag(caged,-1)+lag(caged,-2)+lag(caged,-12) +lag(pim,-0)+sd-1)

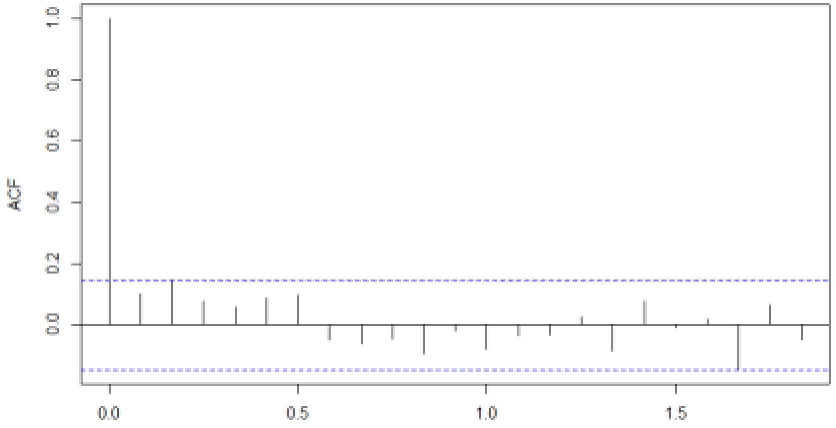
- Resultado da estimação

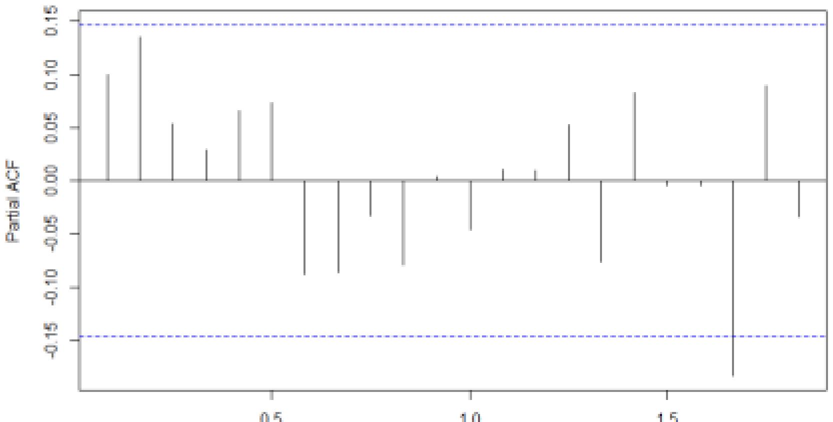
  summary(x1)

- Verificando os resíduos

  erro1=resid(x1)

  acf(erro1); pacf(erro1)





Artigo – Ang et al (2006)

- DO MACRO VARIABLES, ASSET MARKETS OR SURVEYS FORECAST INFLATION BETTER?
- Surveys do! We examine the forecasting power of four alternative methods of forecasting U.S. inflation out-of-sample: time series ARIMA models; regressions using real activity measures motivated from the Phillips curve; term structure models (…); and survey-based measures.
- NBER Working Paper No. 11538
- Andrew Ang, Geert Bekaert and Min Wei
- Four different measures of inflation; The first three are consumer price index (CPI) measures: i) including CPI-U for All Urban Consumers, All Items (*PUNEW*); ii) CPI for All Urban Consumers, All Items Less Shelter (*PUXHS*) and iii) CPI for All Urban Consumers, All Items Less Food and Energy (*PUXX*), which is also called core CPI; iv) Personal Consumption Expenditure deflator (*PCE*);
- The sample period is 1952:Q2 to 2002:Q4 for PUNEW and PUXHS, 1958:Q2 to 2002:Q4 for PUXX, and 1960:Q2 to 2002:Q4 for PCE.

Artigo – Ang et al (2006)

- Modelos
  - ARIMA: $\pi_{t+1} = \mu + \phi\pi_t + \psi\varepsilon_t + \varepsilon_{t+1}$
  - Curva de Phillips: $\pi_{t+4,4} = \alpha + \beta(L)'X_t + \varepsilon_{t+4,4}$
  - Modelo linear da estrutura a termo: $X_{t+1} = \mu + \Phi X_t + \varepsilon_{t+1}$
  - Usando projeções de mercado: $\pi_{t+4,4} = \alpha_1 + \beta_1 f_t^S + \varepsilon_{t+4,4}$
- Metodologia de projeção:

$$\mathrm{E}_t(\pi_{t+4,4}) = \mathrm{E}_t\left(\sum_{i=1}^{4} \pi_{t+i}\right)$$

$$\pi_{t+4,4} = \pi_{t+1} + \pi_{t+2} + \pi_{t+3} + \pi_{t+4}$$

Artigo - Ang et al (2006)

- Resultado

Table 8: Best Models in Forecasting Annual Inflation

	PUNEW		PUXHS		PUXX		PCE	
Panel A: Post-1985 Sample								
Best Time-Series Model	ARMA	1.000	ARMA	1.000	ARMA	1.000	ARMA	1.000*
Best Phillips-Curve Model	PC1	0.979	PC1	1.000	PC8	0.862	PC4	1.027
Best Term-Structure Model	TS7	1.091	MDL2	1.047	TS1	0.945	TS7	1.018
Raw Survey Forecasts	SPF1	0.779*	SPF1	0.819*	SPF1	0.691	SPF1	1.199
	LIV1	0.789	LIV1	0.844	LIV1	0.655*	LIV1	1.082
	MICH1	0.902	MICH1	0.881	MICH1	1.185	MICH1	1.217
Panel B: Post-1995 Sample								
Best Time-Series Model	RGM	0.764*	RGM	0.833*	RW	0.915	ARMA	1.000*
Best Phillips-Curve Model	PC1	0.977	PC1	0.992	PC8	0.767	PC6	1.020
Best Term-Structure Model	VAR	0.913	VAR	0.973	TS6	0.866	TS8	1.025
Raw Survey Forecasts	SPF1	0.861	SPF1	0.914	SPF1	0.699	SPF1	1.250
	LIV1	0.792	LIV1	0.856	LIV1	0.557*	LIV1	1.101
	MICH1	0.862	MICH1	0.937	MICH1	0.822	MICH1	1.338

The table reports the best time-series model, the best OLS Phillips Curve model, the best model using term structure data, along with SPF1, LIV1, and MCH1 forecasts for out-of-sample forecasting of annual inflation at a quarterly frequency. Each entry reports the ratio of the model RMSE to the RMSE of an ARMA(1,1) forecast. Models with the smallest RMSEs are marked with an asterisk.

RGM – Regme-Switching Model

Artigo - Ang et al (2006)

- Aplicar para o IPCA no Brasil.
- Usaremos o período de 2003-2013.
- O restante da série será para a projeção fora da amostra.
- Para esta aula, aplicaremos dois modelos:
  - ARIMA: $\pi_{t+1} = \mu + \phi\pi_t + \psi\varepsilon_t + \varepsilon_{t+1}$
  - Projeção de mercado: $\pi_{t+4,4} = \alpha_1 + \beta_1 f_t^S + \varepsilon_{t+4,4}$
- Os outros modelos, discutiremos nas próximas aulas quando aprendermos sobre filtros de suavização, VAR e componentes principais.

Artigo - Ang et al (2006)

- Baixando os dados.
- list.files()
- Z<-read.csv("modelo_ang_bekaert.csv",sep=";", dec=".", head=TRUE)
- Z$data<-as.Date(Z$data,'%d/%b/%y')
- head(Z)

Artigo - Ang et al (2006)

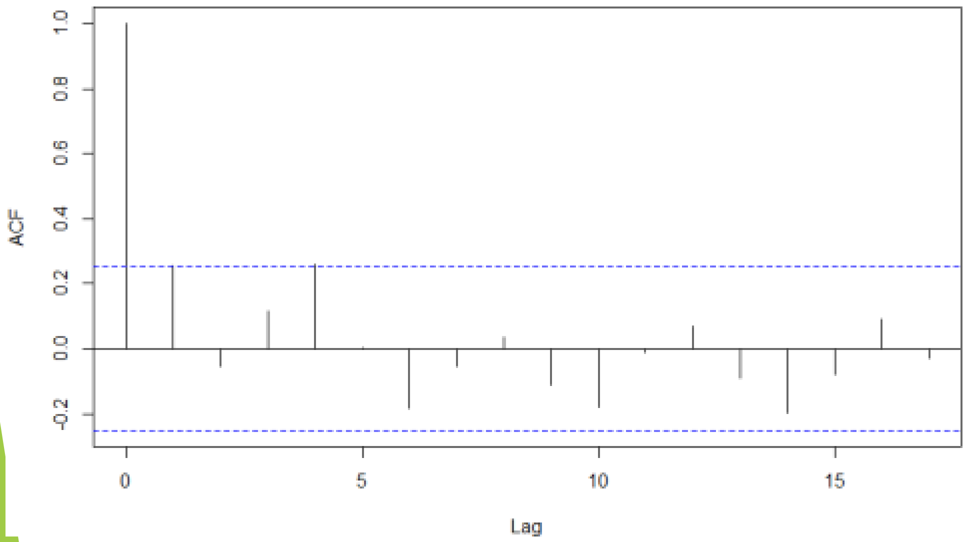
- Modelo ARMA

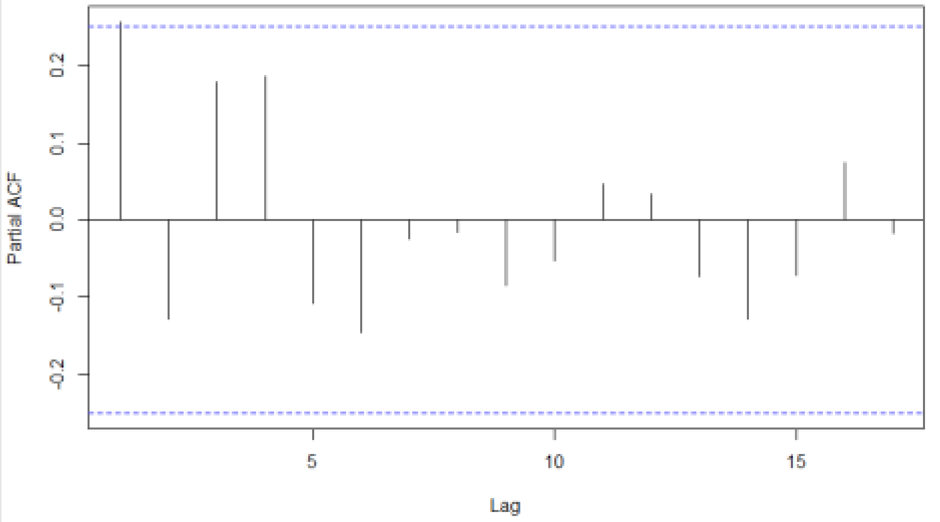
  ipca_trim<-ts(Z$ipca[1:61],start = c(2003,1),frequency = 4) # série temporal

  require(nlme) #pacote para estimar o acf e o pacf

  acf(ipca_trim)

  pacf(ipca_trim)



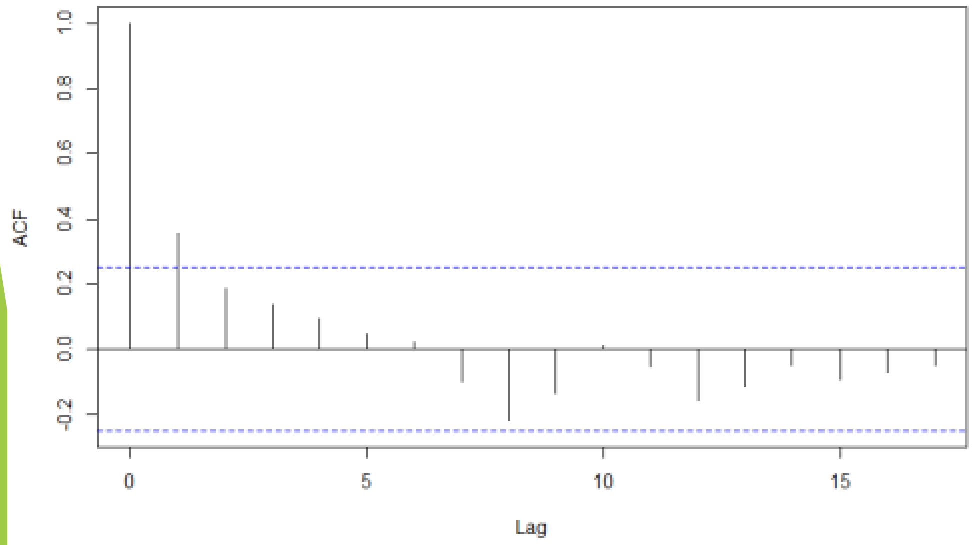


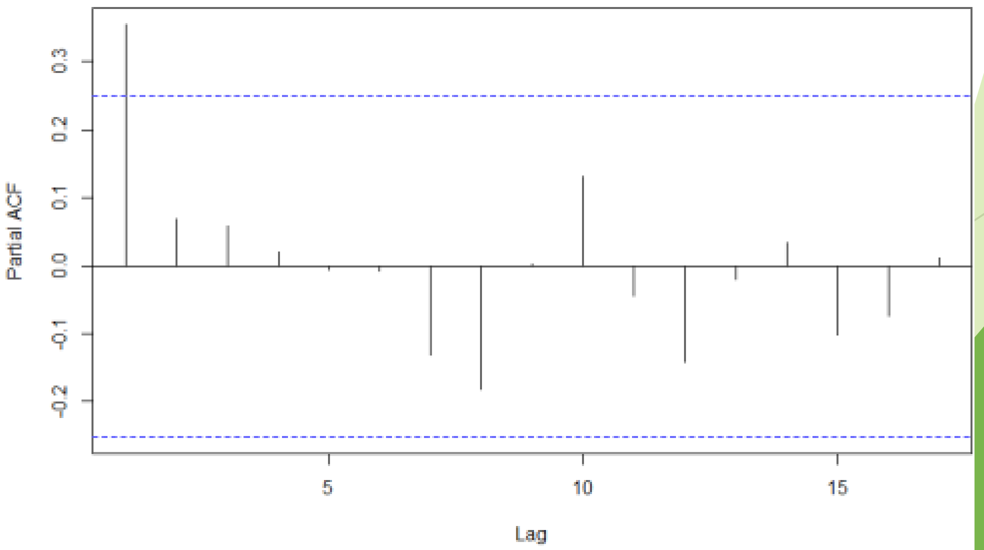
- Soluções: i) SARIMA ou ii) Com *dummies* sazonais (BC)

Artigo - Ang et al (2006)

- Modelo usando *dummies* sazonais:

```
sd_trim<-seasonal.dummies(ipca_trim)

x_trim<-lm(ipca_trim~sd_trim-1)

x_erro<-resid(x_trim)

acf(x_erro)

pacf(x_erro)
```





Artigo - Ang et al (2006)

- Modelo AR(1) com *dummies* sazonais:
- Separando as amostras:
  - ipca_treino<-ts(Z$ipca[1:44],start = c(2003,1),frequency = 4) #período até 2013
- Estimando a equação para o período de "treino":
  - sd_treino<-seasonal.dummies(ipca_treino)
  - x_treino<-arima(ipca_treino,order = c(1,0,0),xreg=sd_treino, include.mean = FALSE)
  - x_treino <enter>
  - acf(x_treino$resid)
  - pacf(x_treino$resid)
  - LjungBoxTest(x_treino$resid)

```
> LjungBoxTest(x_treino$resid)
  m     Qm    pvalue
  1   2.14 0.1435730
  2   2.16 0.3402804
  3   2.42 0.4900405
  4   2.49 0.6458116
  5   2.60 0.7606189
  6   3.98 0.6787439
  7   4.19 0.7577596
  8   6.29 0.6150024
  9   6.29 0.7107722
 10   6.55 0.7669198
 11   6.62 0.8287221
 12   6.66 0.8793832
 13   7.52 0.8734840
 14   7.52 0.9127904
 15   7.71 0.9350440
```

Artigo - Ang et al (2006)

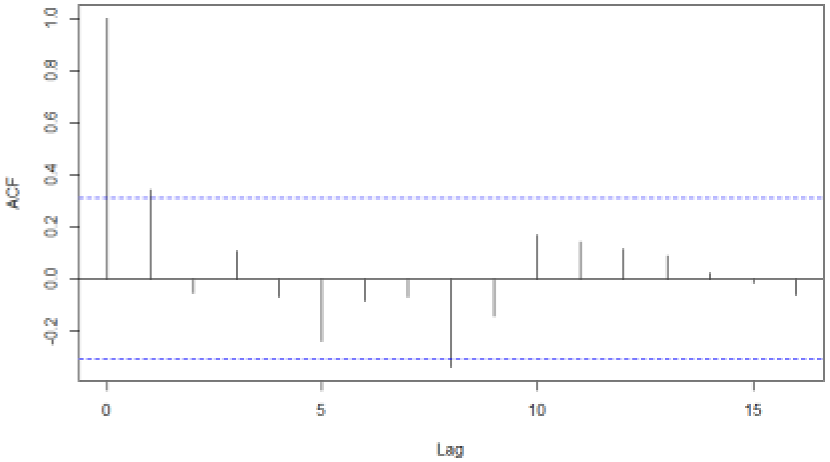
- Projetando para o período de 2014Q1-2018Q1:
- Aumentando o tamanho da variável independente
  - sd_proj<-seasonal.dummies(ipca_proj) #período de 2014Q1 – 2018Q1
- Projetando com o modelo estimado:
  - p_ipca_arma=predict(x_treino,sd_proj,n.ahead=17)
  - p_ipca_arma$pred

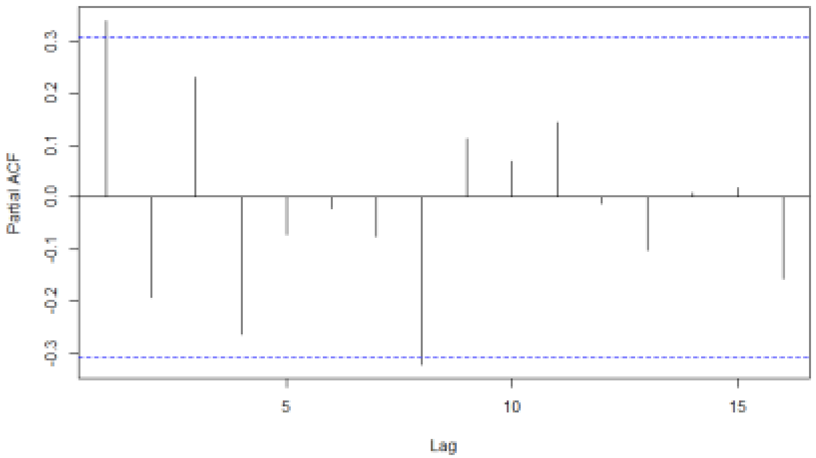
```
> p_ipca_arma$pred
         Qtr1     Qtr2     Qtr3     Qtr4
2014 2.165048 1.310610 1.058445 1.709986
2015 2.019985 1.247483 1.030974 1.698032
2016 2.014783 1.245219 1.029989 1.697603
2017 2.014596 1.245138 1.029954 1.697588
2018 2.014590
```

- Estimando o MSE
  - MSE_arma=mean((p_ipca_arma$pred-ipca_proj)^2)^0.5=0.8376

Artigo - Ang et al (2006)

- Modelo com projeções de mercado:
- Criando um conjunto de variáveis explicativas:
  - ipca_expctajs=ts(Z$ipca_expect[1:40],start = c(2004,1),frequency = 4) – defasagem 1 ano
  - ipca_treino1<-ts(Z$ipca[5:44],start = c(2004,1),frequency = 4)
  - sd_treino1<-seasonal.dummies(ipca_treino1)
  - dados1<-cbind(ipca_expctajs,sd_treino1)
- Estimando a equação:
  - x_expec<-lm(ipca_treino1~dados1-1)
  - acf(resid(x_expec)) e pacf(resid(x_expec))

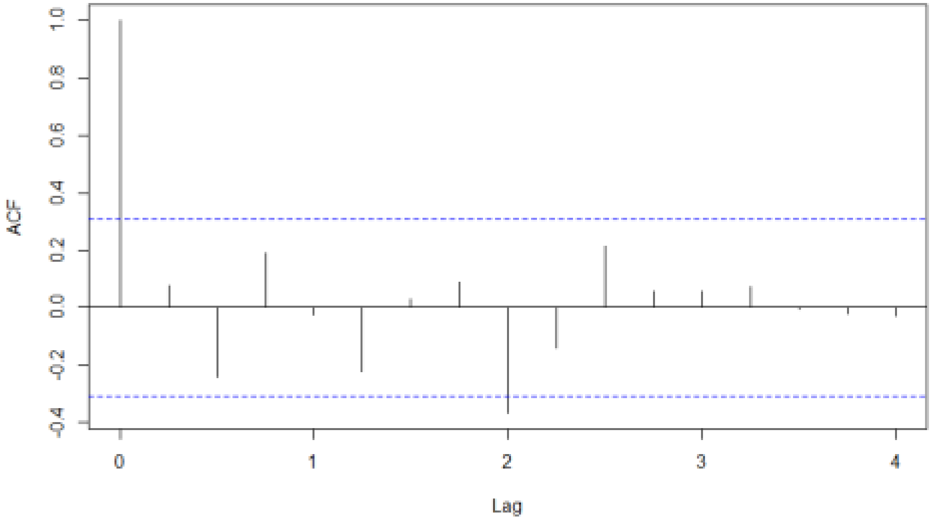


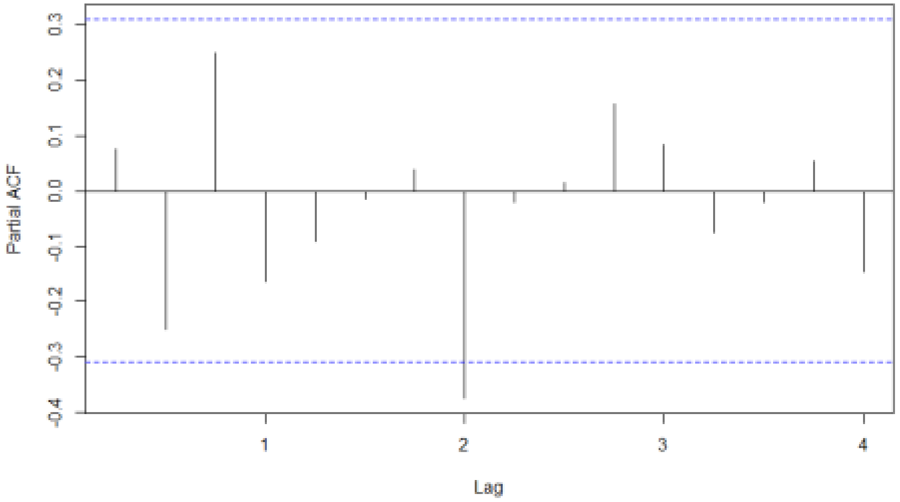


Artigo - Ang et al (2006)

- Estimando a equação com o AR(1)

```
x_treino1<-arima(ipca_treino1,order = c(1,0,0),xreg=dados1, include.mean = FALSE)

acf(x_treino1$residuals)

pacf(x_treino1$residuals)
```





Artigo - Ang et al (2006)

- Criando variáveis para projeção:

  ipca_expcta_proj=ts(Z$ipca_expect[41:57],start = c(2014,1),frequency = 4)

  dados1_proj<-cbind(ipca_expcta_proj,sd_proj)

- Projetando com o modelo estimado:
  - p_ipca_expec=predict(x_treino1, dados1_proj,n.ahead=17)
  - p_ipca_expec$pred

```
> p_ipca_expec$pred
          Qtr1      Qtr2      Qtr3      Qtr4
2014 1.8461263 1.2798458 1.0207529 1.6717144
2015 1.7442329 1.2447309 1.0096535 1.7006681
2016 1.7461521 1.2451335 0.9929848 1.7198984
2017 1.7636873 1.2433647 0.9402737 1.5945716
2018 1.6524167
```

- MSE_arma=mean((p_ipca_expec$pred-ipca_proj)^2)^0.5=0.8415