

# Mais sobre modelagem

Prof. Caio Azevedo

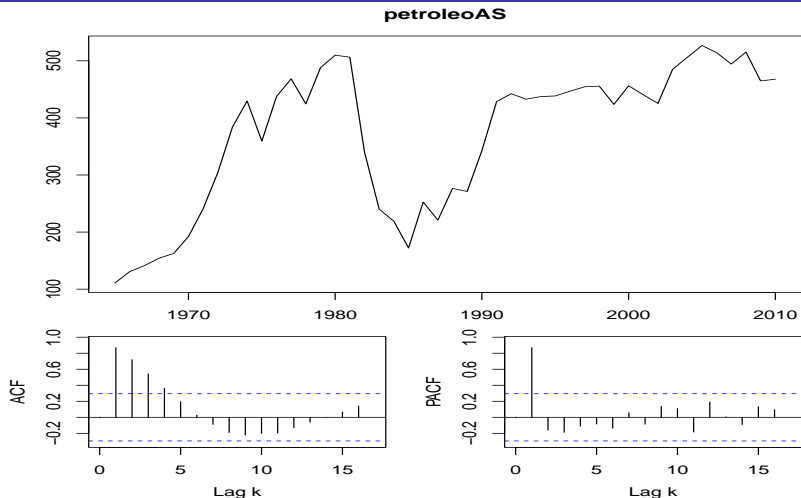
# Introdução

- Discutiremos um pouco mais sobre modelagem de ST, limitando-nos a classe de modelos ARIMA.
- Essencialmente usaremos as metodologias vistas (gráficos de ST, FAC, FACP, eventualmente suavização, estimação por MV, análise de resíduos padronizados e previsão) usando a função `sarima` do pacote `astsa`.

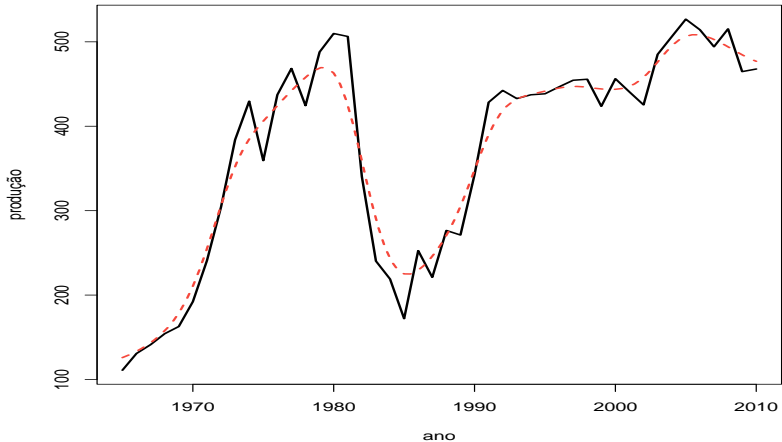
# Produção de petróleo da Arábia Saudita

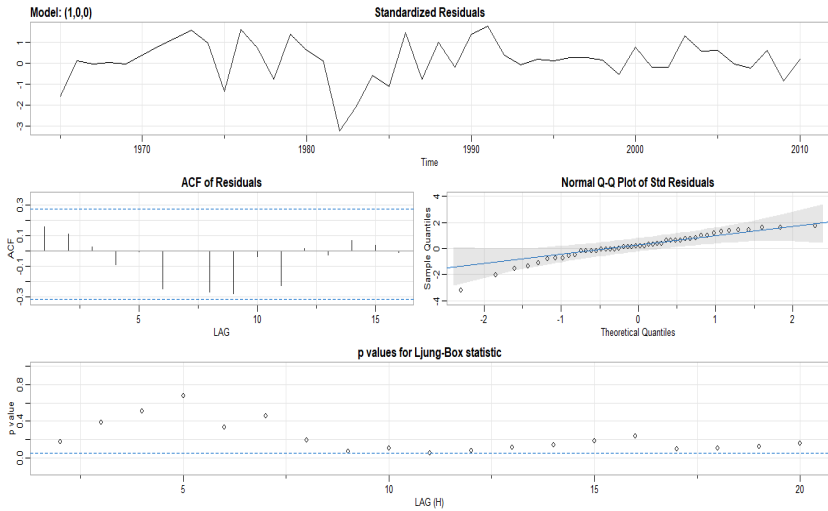
- Já mencionada anteriormente e inicialmente analisada ([aqui](#)).
- Produção anual de petróleo (milhões de toneladas), Arábia Saudita, 1965-2010. Um total de  $n = 46$  observações.
- ST aparentemente estacionária (?) modelável por um processo AR(1).

# Gráficos de ST, FAC e FACP da ST



## Suavização por kernel ( $b=4$ )





# Estatísticas de Comparação de modelos

Modelo	AIC	AICc	BIC
AR1	10,788	10,794	10,907
AR2	10,775	10,788	10,934
AR3	10,785	10,806	10,984

## Ajustes dos modelos

Par.	Est.	EP	IC(95%)	Estat. t	p-valor
$\phi$	0,94	0,05	[0,85 ; 1,04]	19,55	<0,0001
$\mu$	336,97	96,91	[147,04 ; 526,90]	3,48	0,0012

raiz de  $\phi(z)$  :  $|z| \approx 1,061$  (modelo estacionário e causal)

Par.	Est.	EP	IC(95%)	Estat. t	p-valor
$\phi_1$	1,16	0,14	[0,88 ; 1,44]	8,21	<0,0001
$\phi_2$	-0,24	0,15	[-0,53 ; 0,05]	-1,63	0,1100
$\mu$	347,52	76,38	[197,82 ; 497,22]	4,55	<0,0001

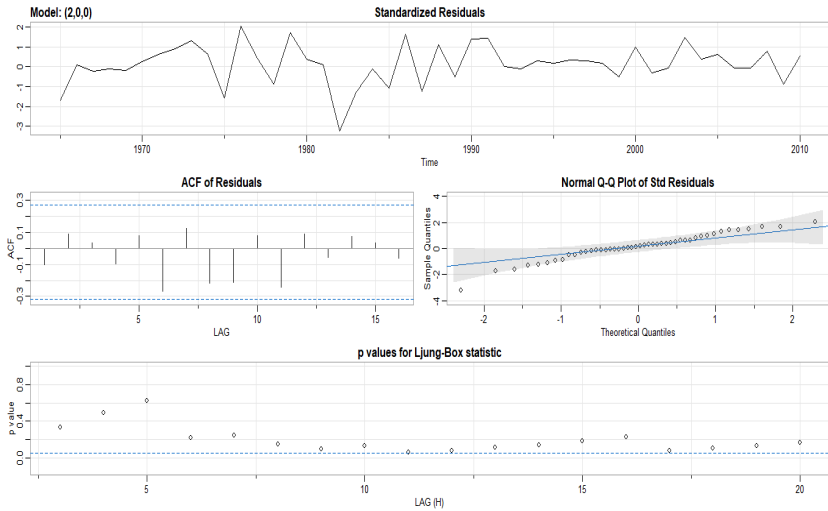
raízes de  $\phi(z) \approx (1, 122; 3, 723)$  (modelo estacionário e causal)

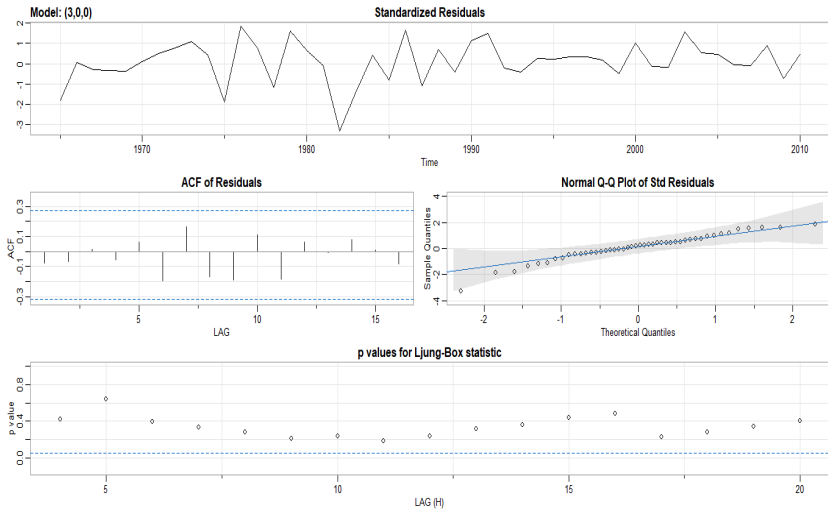


## Ajustes dos modelos

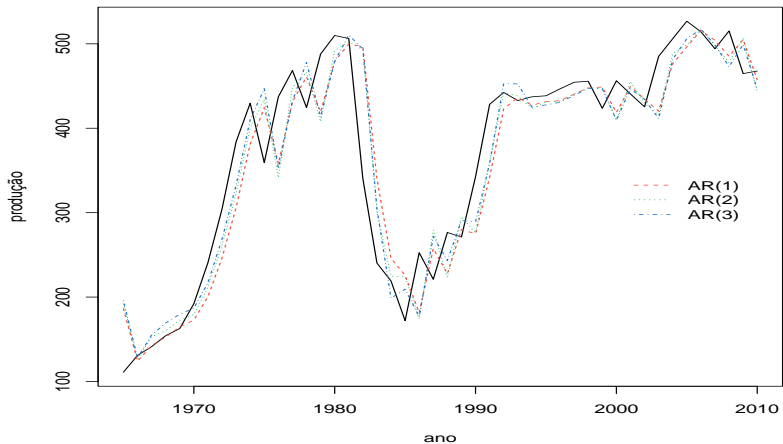
Par.	Est.	EP	IC(95%)	Estat. t	p-valor
$\phi_1$	1,11	0,14	[0,83 ; 1,39]	7,80	<0,0001
$\phi_2$	-0,03	0,22	[-0,46 ; 0,40]	-0,13	0,8995
$\phi_3$	-0,19	0,15	[-0,48 ; 0,11]	-1,26	0,2147
$\mu$	353,41	61,21	[233,45 ; 473,38]	5,77	<0,0001

raízes de  $\phi(z) \approx (1,359; 1,359; 2,854)$  (modelo estacionário e causal)

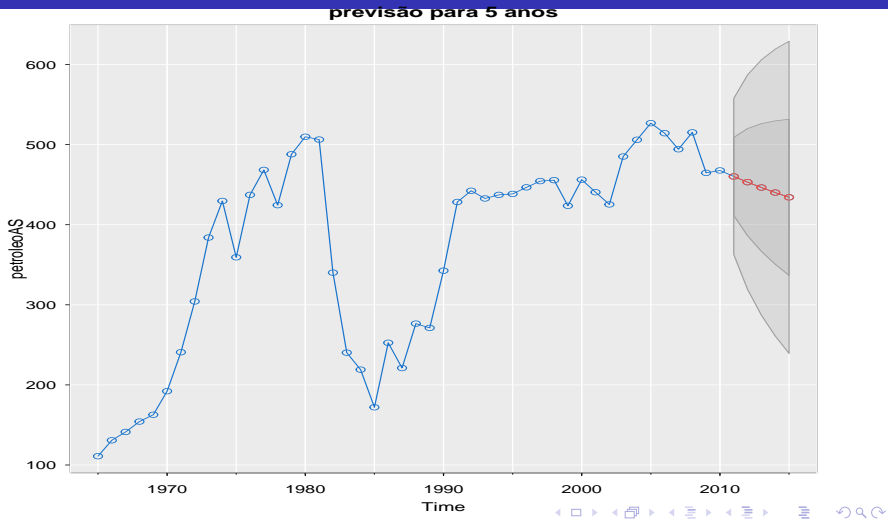




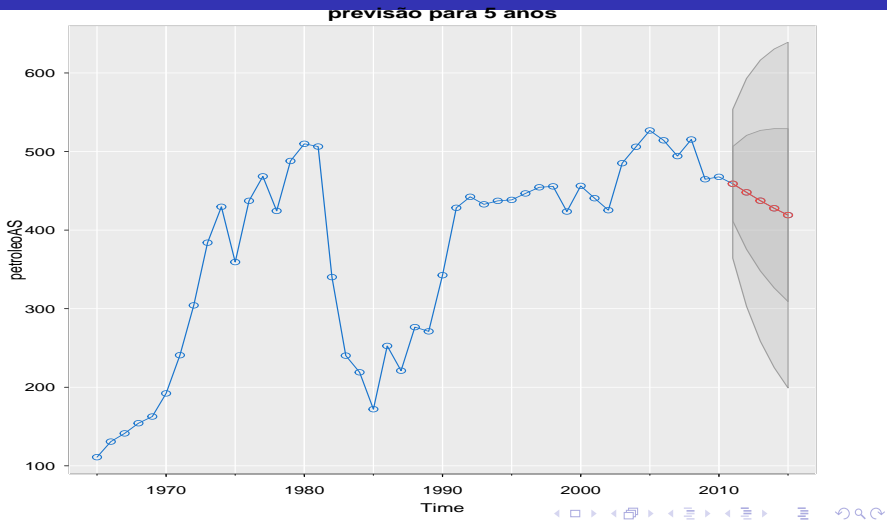
# Previsão para valores observados



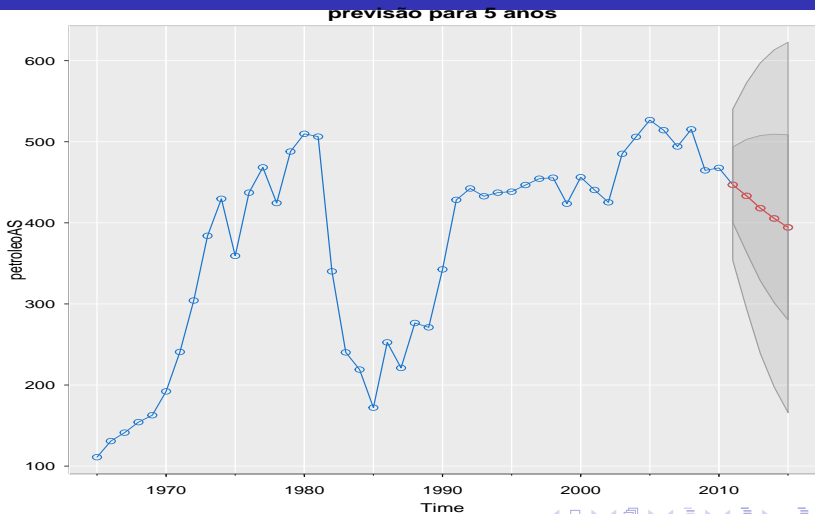
# Previsão para valores futuros AR(1)



# Previsão para valores futuros AR(2)



# Previsão para valores futuros AR(3)



# Comentários

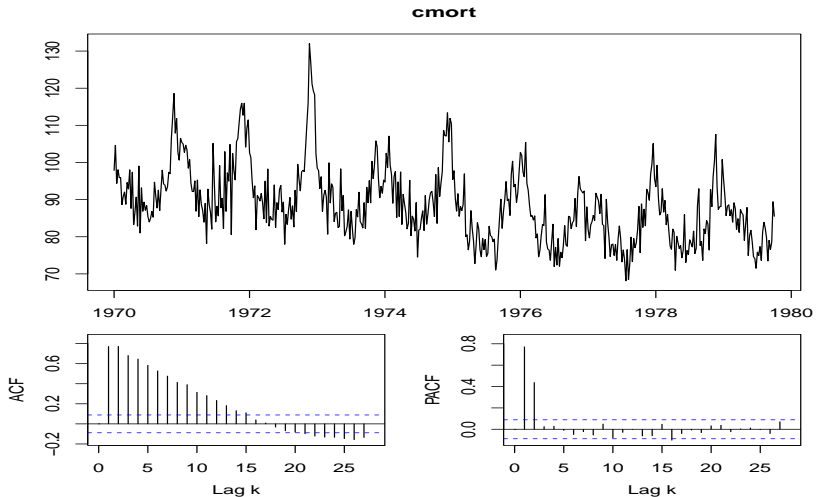
- Nenhum dos três modelos se ajustou bem aos dados: 1) Assimetria negativa dos resíduos e comportamento de não ruído branco. O modelo AR(3) apresentou o “melhor” ajuste.
- Provavelmente por isso, as estatísticas de comparação de modelo apresentaram resultado contraditório com as estimativas. Enquanto que a primeira indica o modelo AR(2), a segunda indica o modelo AR(1).
- Sugestão: ajustar um modelo skew-t ou skew-normal com estrutura (AR(1), AR(2), AR(3)).



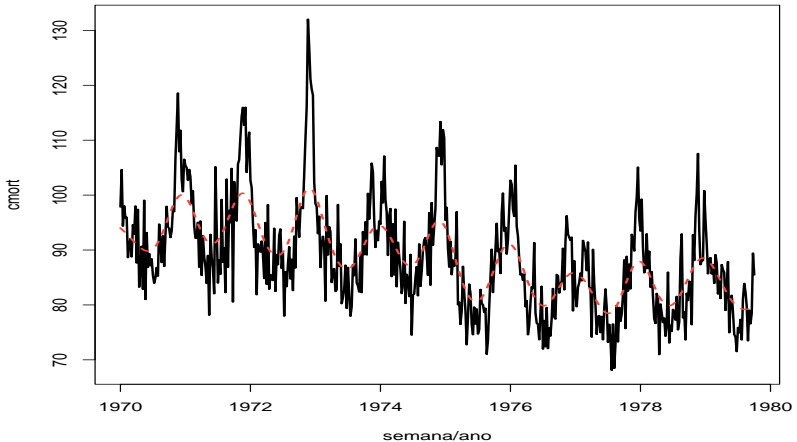
# Exemplo: número de mortes por questões cardiovasculares

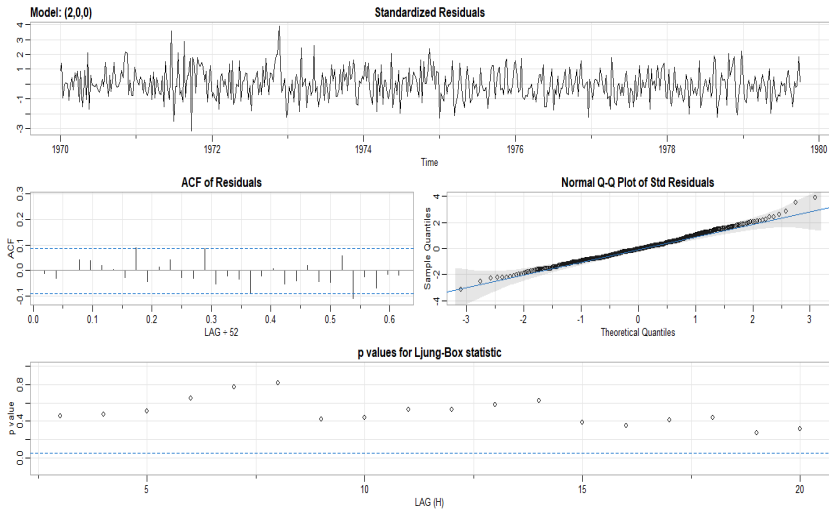
- Média semanal de mortes por questões cardiovasculares, no condado de Los Angeles, no período de 1970-1979
- A ST parece ser estacionária (com sazonalidade).
- Um modelo  $AR(p)$  parece ser apropriado.

# Gráficos de ST, FAC e FACP da ST

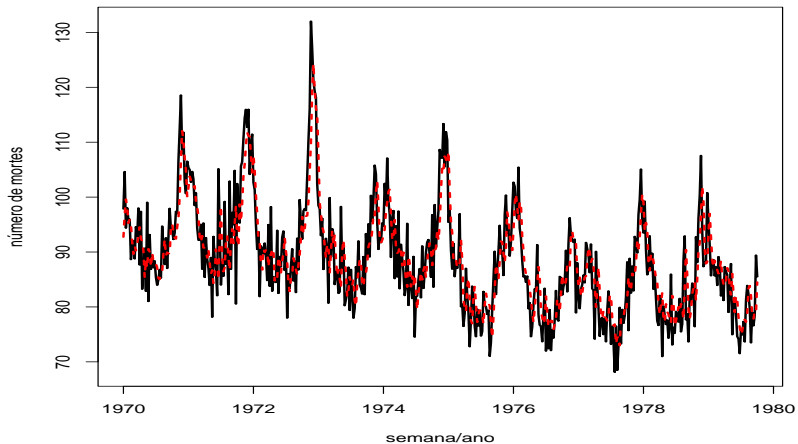


# Suavização por kernel ( $b=4$ )

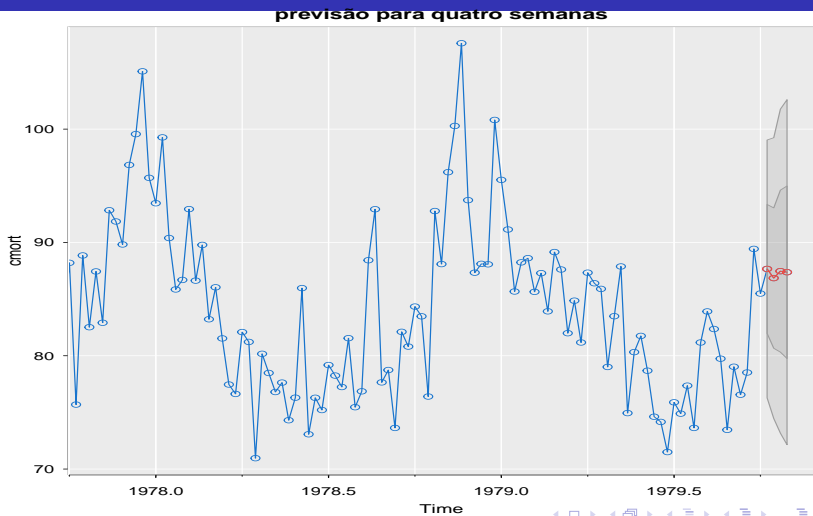




# Previsão para valores observados



# Previsão para valores futuros



# Ajuste dos modelos

Modelo	AIC	AICc	BIC
AR(1)	6,547	6,547	6,572
AR(2)	6,334	6,334	6,367
AR(3)	6,337	6,337	6,378

# Estimativas

Par.	Est.	EP	IC(95%)	Estat. t	p-valor
$\phi_1$	0,43	0,04	0,35 ; 0,51	10,84	<0,0001
$\phi_2$	0,44	0,04	0,36 ; 0,52	11,12	<0,0001
$\mu$	88,85	1,94	85,05 ; 92,66	45,78	<0,0001

raízes de  $\phi(z) \approx (1,094; 2,066)$  (modelo estacionário e causal)

Par.	Est.	EP	IC(95%)	Estat. t	p-valor
$\phi_1$	0,42	0,04	[0,33 ; 0,51]	9,47	<0,0001
$\phi_1$	0,43	0,04	[0,34 ; 0,52]	9,76	<0,0001
$\phi_1$	0,03	0,04	[-0,06 ; 0,11]	0,57	0,5673
$\mu$	88,87	1,99	[84,98 ; 92,77]	44,70	<0,0001



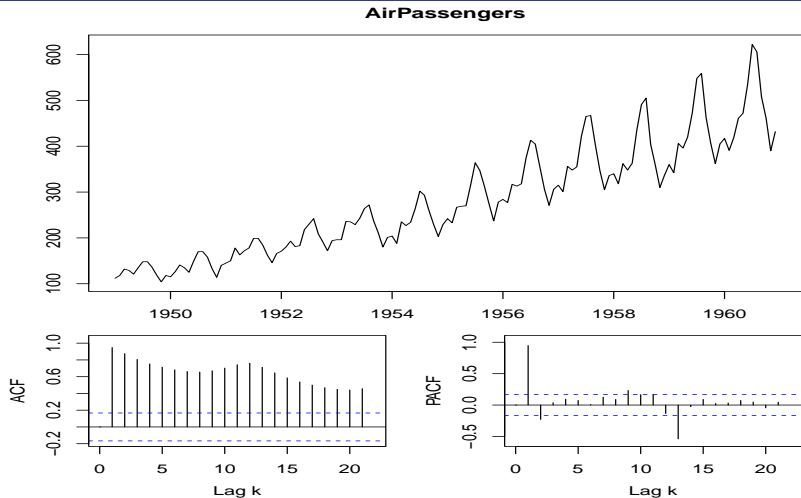
# Comentários

- O modelo se ajustou (análise residual + análise preditiva para os valores observados). Contudo, a previsão para valores futuros parece “destoar” um pouco do que se esperaria.
- O modelo AR(2) se ajustou melhor do que os demais. Inclusive, coeficientes (MA) do modelo ARMA(2,2) se mostraram não significativos.
- Sugestão: ajustar um modelo SARIMA (com sazonalidade + AR, sem a componente MA).

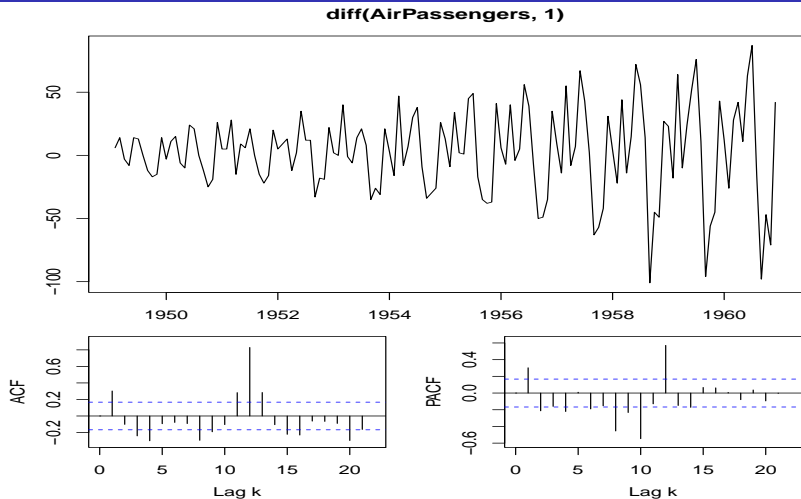
# Número de passageiros em companhias aéreas

- Totais mensais de passageiros de companhias aéreas internacionais, 1949 a 1960.
- Já mencionada anteriormente e inicialmente analisada ([aqui](#)).
- A ST parece não estacionária. Sazonalidade e/ou variabilidade não constante.
- As diferenças de ordem 1 e 2 parecem não levar a ST estacionárias.
- De qualquer forma, vamos ajustar modelos AR(1) e ARIMA(2,1,2).

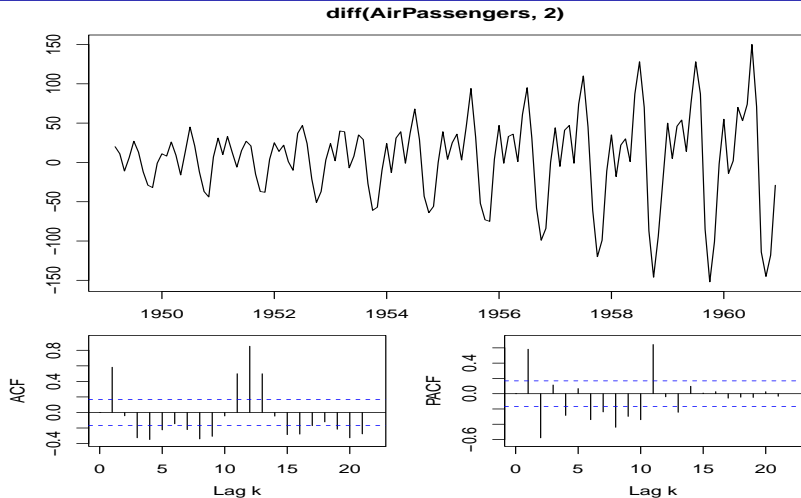
# Gráficos de ST, FAC e FACP da ST

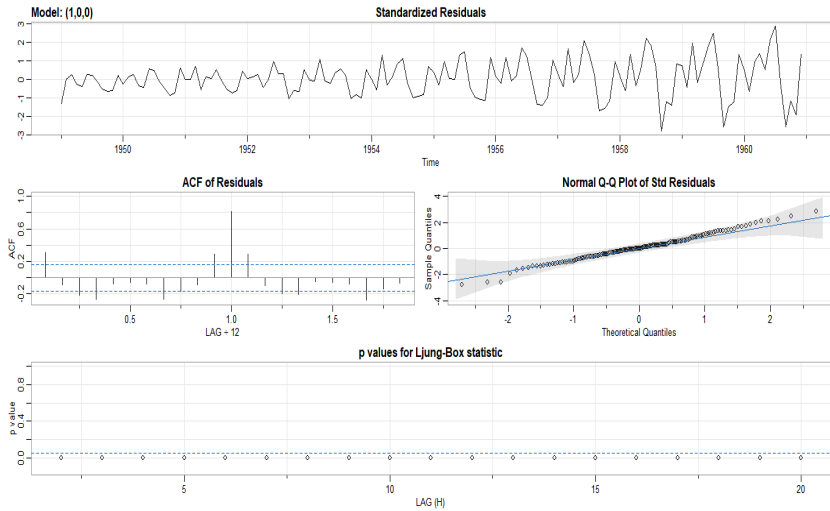


# Gráficos de ST, FAC e FACP da ST ( $\Delta_1$ )

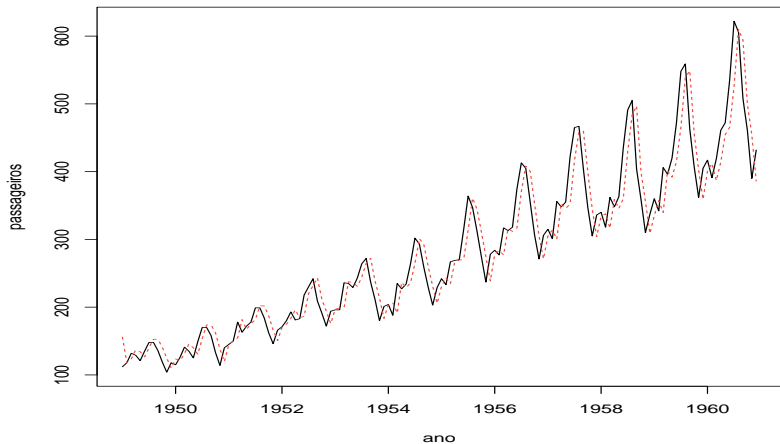


# Gráficos de ST, FAC e FACP da ST ( $\Delta_2$ )

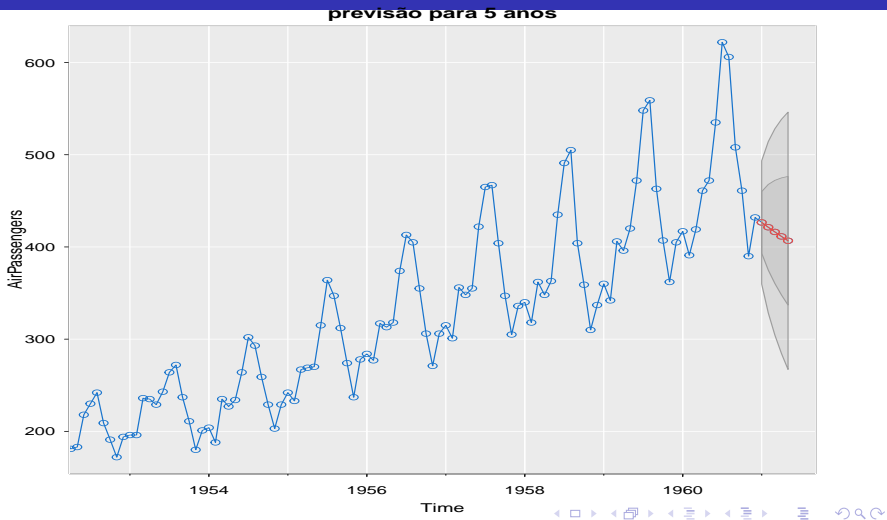




# Previsão para valores observados



# Previsão para valores futuros





# Estimativas

Par.	Est.	EP	IC(95%)	Estat. t	p-valor
$\phi_1$	0,96	0,02	[0,92 ; 1,01]	45,16	< 0,0001
$\mu$	278,46	67,11	[146,92 ; 410,01]	4,15	0,0001

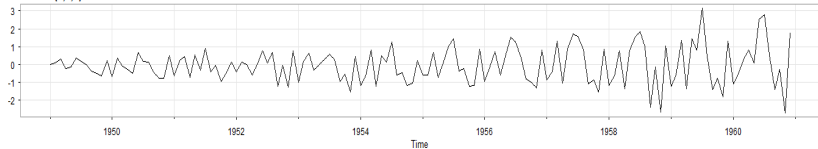
raiz de  $\phi(z)$ :  $|z| \approx 1,037$  (modelo estacionário ??? e causal)

# Comentários

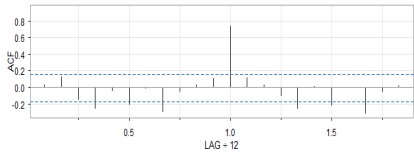
- O modelo não se ajustou bem aos dados (análise residual e predição futura) embora as predições (pontuais) para valores observados tenham se mostrado apropriadas.
- A indicação da estacionariedade, através da raiz de  $\phi(z)$ , que conflita com o esperado, provavelmente, se deve ao mal ajuste do modelo.
- Veremos o ajuste do modelo ARIMA(2,1,2).

Model: (2,1,2)

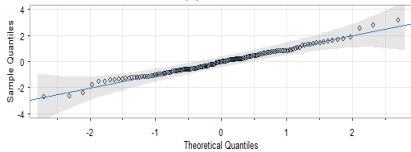
Standardized Residuals



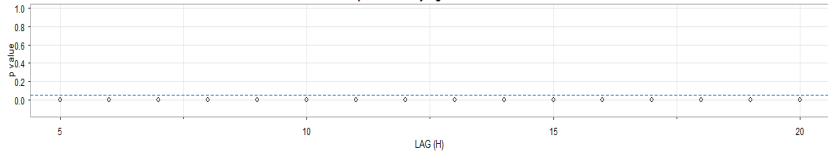
ACF of Residuals



Normal Q-Q Plot of Std Residuals



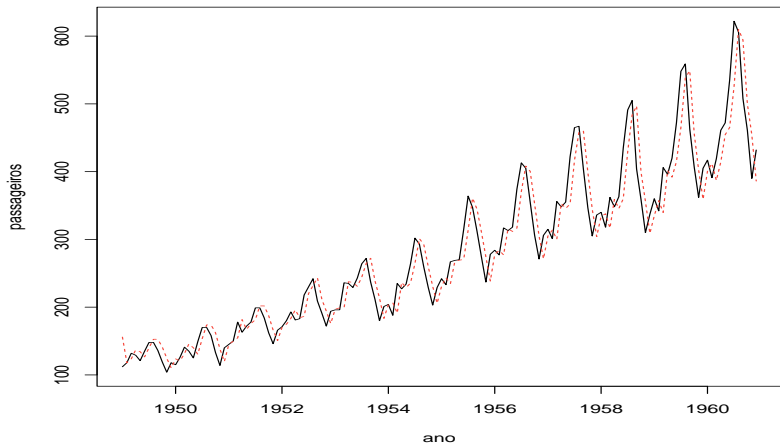
p values for Ljung-Box statistic



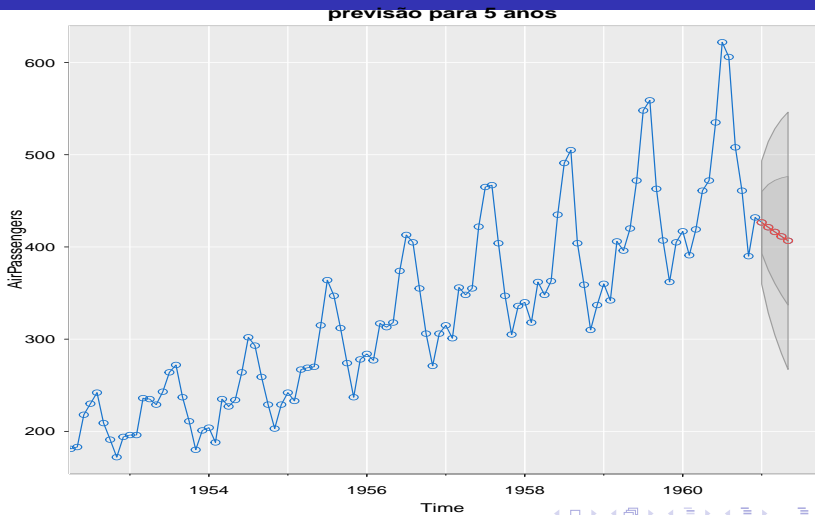
# Comentários

- O modelo não se ajustou bem aos dados (análise residual e predição futura) embora as predições (pontuais) para valores observados tenham se mostrado apropriadas.
- A indicação da estacionariedade, através da raiz de  $\phi(z)$ , que conflita com o esperado, provavelmente, se deve ao mal ajuste do modelo.
- Veremos o ajuste do modelo ARIMA(2,1,2).

# Previsão para valores observados



# Previsão para valores futuros



# Estimativas

Par.	Est.	EP	IC(95%)	Estat. t	p-valor
$\phi_1$	0,38	0,32	[-0,24 ; 1,00]	1,19	0,2372
$\phi_2$	0,10	0,28	[-0,45 ; 0,66]	0,37	0,7136
$\theta_1$	-0,23	0,30	[-0,82; 0,35]	-0,79	0,4325
$\theta_2$	-0,77	0,30	[-1,35 ; -0,18]	-2,57	0,0111
$\mu$	2,65	0,18	[2,29 ; 3,01]	14,41	<0,0001

raízes  $-\phi(z) \approx (1,280; 2,467)$  (processo estacionário ??? e causal) e  $\theta(z) \approx (1,072; 3,933)$  (processo invertível)

# Comentários

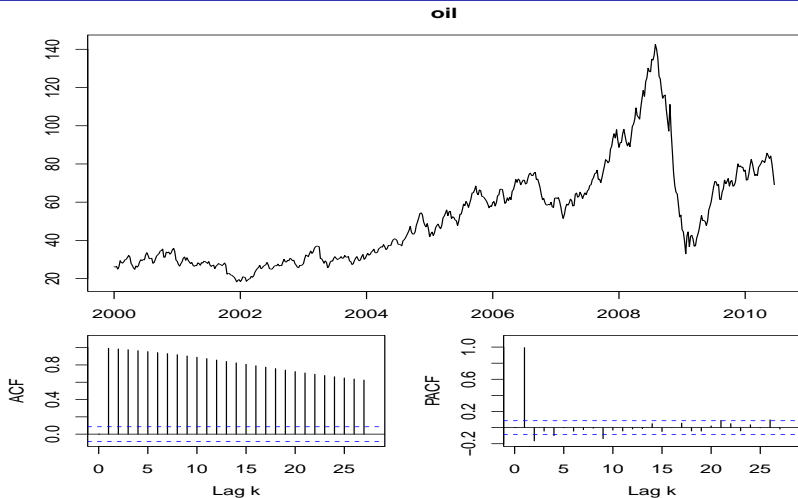
- Essencialmente, o ajuste do modelo  $ARIMA(2,1,2)$  se mostrou tão ruim quanto o do  $AR(1)$ .
- Sugestões: ajustar modelos  $SARIMA(p,d,q)$  e/ou modelos de volatilidade estocástica.



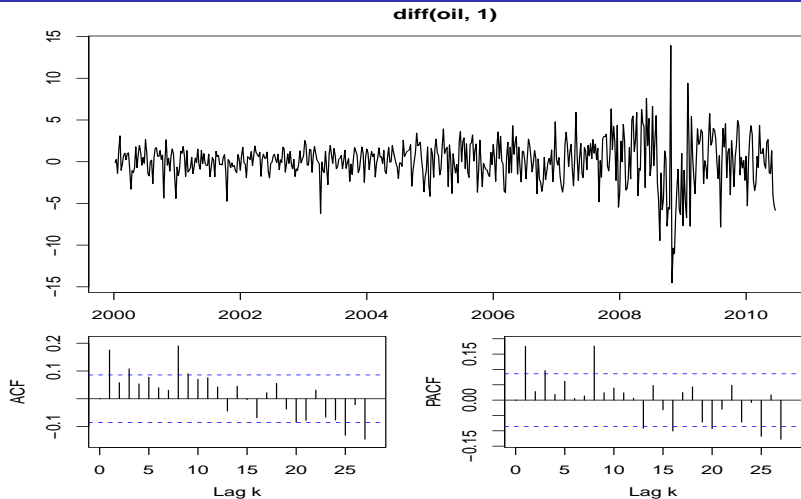
# Preço do (mundial) do petróleo

- Petróleo bruto, preço spot WTI FOB (em dólares por barril), dados semanais de 2000 a meados de 2010.
- ST aparentemente não estacionária, modelável por um processo ARIMA(p,q,d).

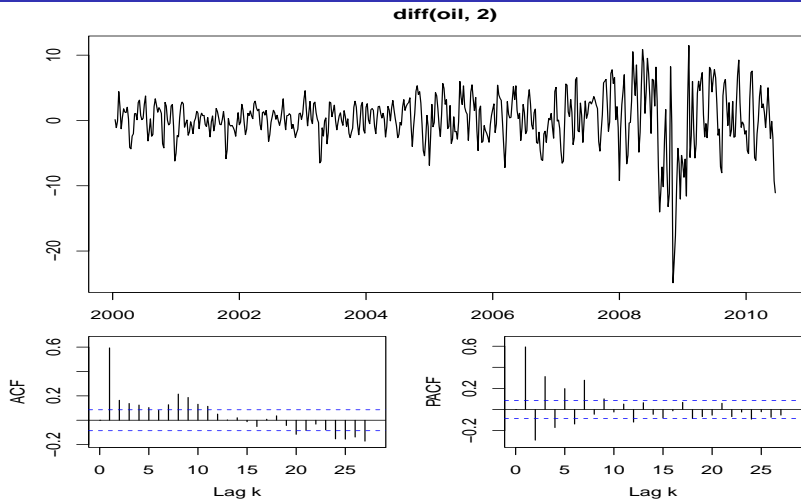
# Gráficos de ST, FAC e FACP da ST



# Gráficos de ST, FAC e FACP da ST ( $\Delta_1$ )

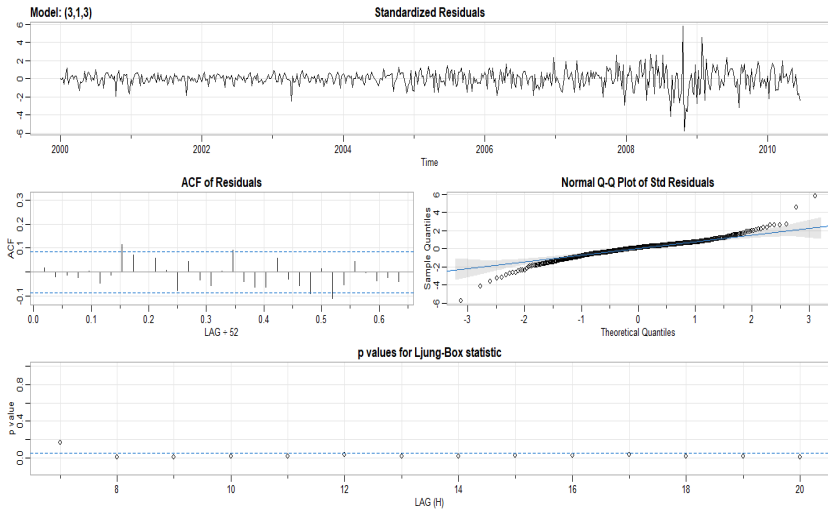


# Gráficos de ST, FAC e FACP da ST ( $\Delta_2$ )



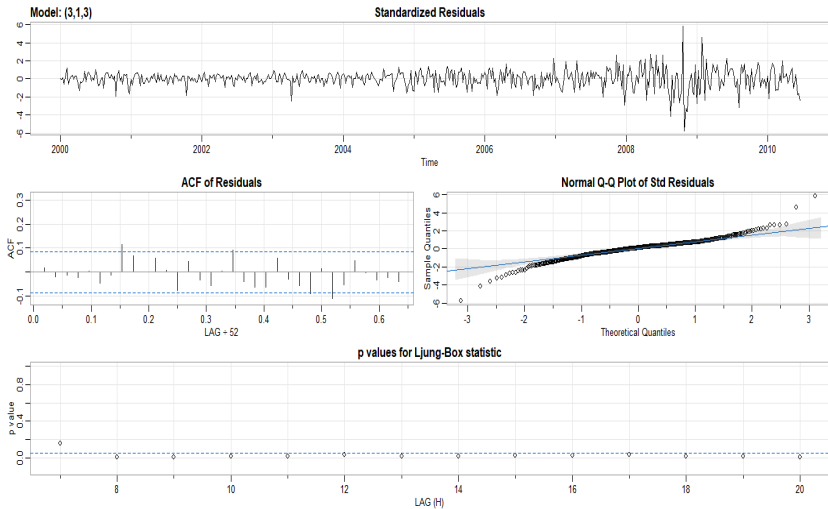
# Suavização por kernel ( $b=4$ )





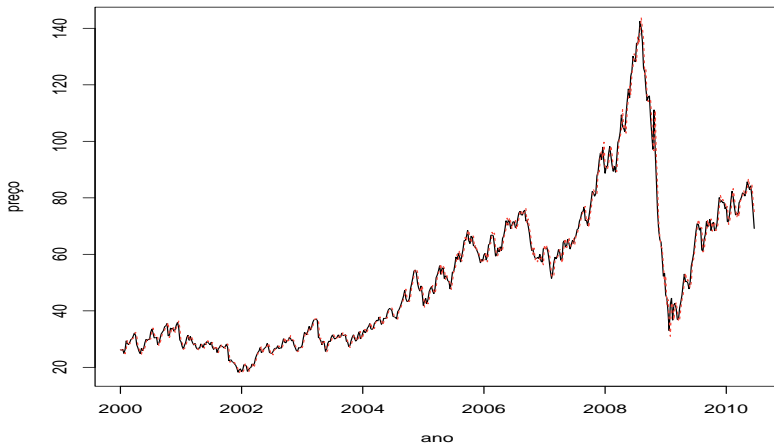
## Estimativas (com intercepto)

Par.	Est.	EP	IC(95%)	Estat. t	p-valor
$\phi_1$	-0,88	0,06	[-1,00 ; -0,76]	-14,20	<0,0001
$\phi_2$	0,64	0,09	[0,46 ; 0,82]	7,06	<0,0001
$\phi_3$	0,82	0,06	[0,71 ; 0,93]	14,46	<0,0001
$\theta_1$	1,04	0,07	[0,91 ; 1,18]	15,31	<0,0001
$\theta_2$	-0,47	0,12	[-0,70 ; -0,23]	-3,88	0,0001
$\theta_3$	-0,78	0,07	[-0,92 ; -0,65]	-11,59	<0,0001
$c$	0,06	0,20	[-0,33 ; 0,45]	0,31	0,7603

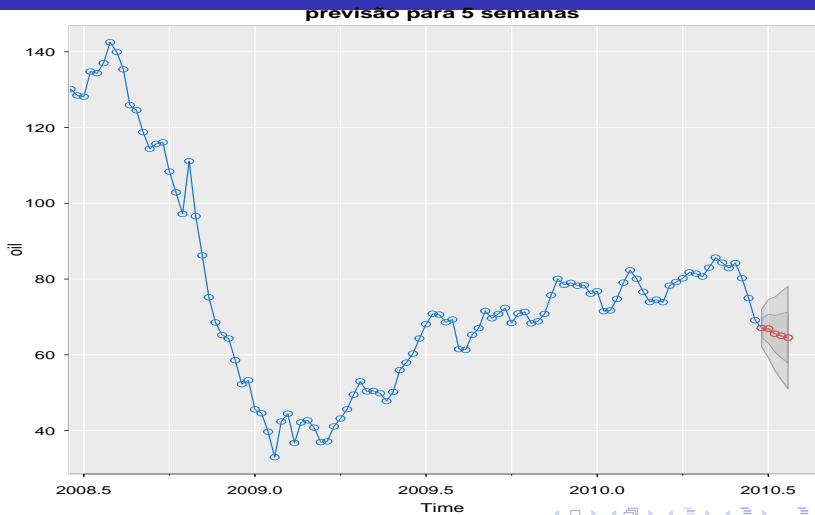




# Previsão para valores observados



# Previsão para valores futuros



# Estimativas

Par.	Est.	EP	IC(95%)	Estat. t	p-valor
$\phi_1$	-0,88	0,06	[-1,00 ; -0,76]	-14,42	<0,0001
$\phi_2$	0,64	0,09	[0,47 ; 0,82]	7,22	<0,0001
$\phi_3$	0,82	0,06	[0,71 ; 0,93]	14,64	<0,0001
$\theta_1$	1,04	0,07	[0,91 ; 1,17]	15,53	<0,0001
$\theta_2$	-0,47	0,12	[-0,70 ; -0,24]	-3,96	0,0001
$\theta_3$	-0,79	0,07	[-0,92 ; -0,65]	-11,76	<0,0001

raízes:  $-\phi(z) \approx (1,128; 1,042; 1,042)$  (processo estacionário e causal) e

$\theta(z) \approx (> 1,001; > 1,001; 4,750)$  (processo invertível)

# Comentários

- O modelo não se ajustou bem aos dados (possivelmente, pela distribuição dos erros).
- Contudo, em termos de previsão, o modelo se comportou bem.
- Sugestão: ajustar um modelo semelhante, mas supondo distribuição  $t$  de Student, para os erros.