Modelagem e Avaliação de Desempenho

Pós Graduação em Engenharia Elétrica - PPGEE
Prof. Carlos Marcelo Pedroso
2018

Séries Temporais

- Considere um processo onde o valor presente de uma série depende dos valores passados.
- Uma série temporal é a amostragem sequencial de uma variável durante um intervalo de tempo, via de regra longo, para possibilitar a identificação de padrões. Pode ser obtida analiticamente, ou também por simulação numérica, ou ainda, por medição experimental.

Séries Temporais

A observação de uma série temporal discreta realizada em instantes de tempo $\tau_1, \tau_2, \ldots, \tau_t, \ldots, \tau_n$ pode ser denotada por $X(\tau_1), X(\tau_2), \ldots, X(\tau_t), \ldots, X(\tau_N)$. Neste capítulo serão consideradas apenas séries temporais discretas, onde as observações são realizadas em um intervalo fixo h. Quando N valores sucessivos da série forem analisados, será escrito $X_1, X_2, \ldots, X_t, \ldots, X_N$ para denotar observações realizadas a intervalos de tempo equidistantes $\tau_0 + h, \tau_0 + 2h, \ldots, \tau_0 + th, \ldots, \tau_0 + Nh$. A esperança será denotada $E(X) = \mu$; a variância será denotada por $V[X] = E\left[(X-\mu)^2\right] = \sigma^2$; a auto-covariância com defasagem k será denotada por $E\left[(X_t - \mu)(X_{t-k} - \mu)\right] = \gamma_k$ e a auto-correlação com defasagem k será denotada por $\rho_k = \gamma_k / \gamma_0$.

Estacionariedade

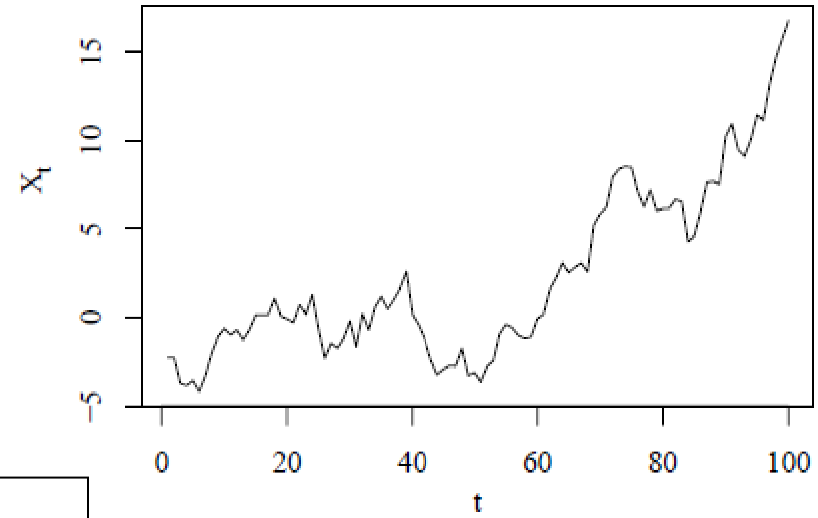
Definição 3.1 Estacionariedade estrita [Willinger e Park 2000]. *X_t é estritamente estacionário se $[X_{t_1}, X_{t_2}, \ldots, X_{t_n}]$ e $[X_{t_1+k}, X_{t_2+k}, ..., X_{t_n+k}]$ possuem a mesma distribuição conjunta para todo $n \in \mathbb{N}$.*

Na estacionariedade estrita, o processo deslocado por k, chamado X_{t+k}, e o primeiro chamado X_t, devem ser equivalentes.

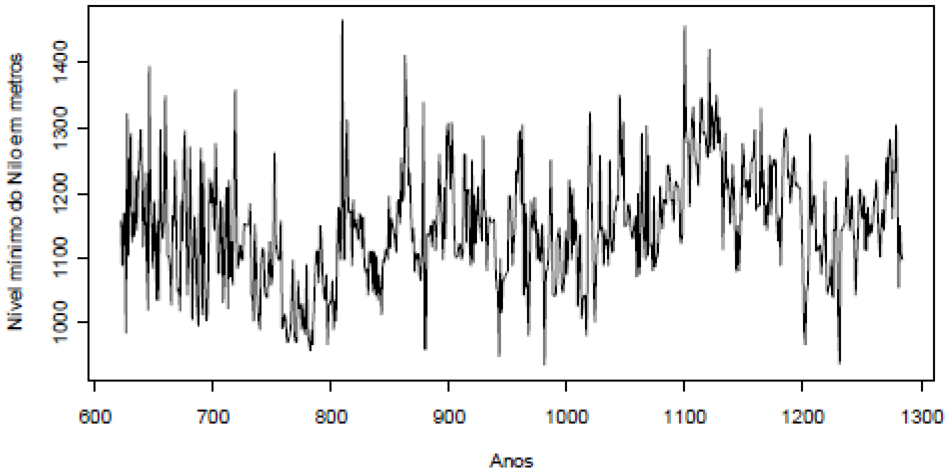
Definição 3.2 Estacionariedade fraca ou de segunda ordem [Willinger e Park 2000]. *A função de auto-covariância $\gamma(r, s) = E\left[(X_r - \mu)(X_s - \mu)\right]$ deve satisfazer à relação de invariância $\gamma(r, s) = \gamma(r + k, s + k) \quad \forall r, s, k \in \mathbb{Z}$.*

Estacionariedade

Não estacionário



Estacionário



Função de Auto-Correlação

- Correlação entre X, Y:

$$\rho_{X,Y} = \frac{\mathrm{cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{E((X-\mu_X)(Y-\mu_Y))}{\sigma_X \sigma_Y}$$

- A análise da função de auto-correlação de uma série é muito importante.
  - Auto-Covariância: $E\left[(X_t-\mu)(X_{t-k}-\mu)\right] = \gamma_k$
  - Auto-Correlação: $\rho_k = \gamma_k/\gamma_0.$

O índice k é o deslocamento (ou lag)

Função de auto-correlação

- Para estimar a função de auto-correlação, calculamos:

$$\overline{\gamma}_k = \frac{\sum_{i=1}^{N-k}(X_i - \overline{X})(X_{i+k} - \overline{X})}{\sum_{i=1}^{N}\left(X_i - \overline{X}\right)^2}$$

Função de auto-correlação

- A função de auto-correlação estabelece o coeficiente de correlação entre um valor X_i da série e o X_{i-k}.
- *A análise da função de auto-correlação tem um papel central na análise de desempenho de sistemas de comunicações.*

Exercícios

1- Calcule a função de auto-correlação para a série a seguir. Plote o gráfico da série temporal e da função de autocorrelação .

		10	20
1	3	13	6
2	8	13	3
3	10	32	2
4	27	18	13
5	5	37	2
6	4	16	1
7	5	9	10
8	16	1	2
9	38	9	1

Exercícios

2- Calcule a função de auto-correlação para a série a seguir. Plote o gráfico da série temporal e da função de auto-correlação (cotação petr4).

		10	20
1	20,80	20,03	20,16
2	20,65	20,37	19,65
3	20,55	21,04	19,70
4	20,52	20,80	20,30
5	20,32	20,55	20,87
6	20,34	20,49	20,76
7	20,15	19,90	20,99
8	20,66	19,86	20,30
9	20,47	20,23	20,21

Exercícios

3- Calcule a função de auto-correlação para a série a seguir. Plote o gráfico da série temporal e da função de auto-correlação

		10	20
1	2.46	1.87	1.59
2	2.05	1.55	1.12
3	2.14	2.04	1.05
4	2.56	2.43	1.36
5	2.49	1.24	1.67
6	1.12	1.11	2.45
7	1.74	0.89	2.93
8	2.11	1.03	2.36
9	2.34	0.77	2.44

Modelo Auto-Regressivo

3.4 Modelo Auto Regressivo

Seja $\tilde{X}_t$ a diferença entre X_t e μ, $\tilde{X}_t = X_t - \mu$. O modelo $AR(p)$, ou auto-regressivo, é dado por $\tilde{X}_t = \phi_1 \tilde{X}_{t-1} + \ldots + \phi_p \tilde{X}_{t-p} + a_t$, onde a_t é um ruído branco e $\phi_1, \ldots, \phi_p$ são parâmetros do modelo. Os parâmetros do modelo são fáceis de estimar e as séries temporais podem ser geradas de maneira simples. A função de auto-correlação decai exponencialmente, o que faz com que o modelo possa ser representado aproximadamente por um modelo de Markov Modulado [Adas 1997]. Este modelo não consegue caracterizar fluxos que possuam uma distribuição de cauda pesada.

Para identificação do modelo AR em uma série existente o analista deve examinar a estrutura da função de auto-correlação da série e estimar o valor de p [Box et al. 1994]. Após esta tarefa, podem ser estimados os valores dos parâmetros $\phi_1, \phi_2, \ldots, \phi_p$.

Modelo Média Móvel (MA)

3.5 Modelo Média Móvel

O modelo $MA(q)$, ou médias móveis, é dado por $\tilde{X}_t = a_t - \theta_1 a_{t-1} - \ldots - \theta_q a_{t-q}$, ou seja, o valor atual de $\tilde{X}_t$ é formado pela soma dos choques ponderados de ruídos aleatórios passados. Os valores de $\theta_1, \ldots, \theta_q$ são parâmetros do modelo.

Para estimação de parâmetros a partir de uma série, o analista deve estimar o valor do parâmetro q e depois estimar os valores de $\theta_1, \theta_2, \ldots, \theta_q$.

Modelo ARMA

3.6 Modelo ARMA

No modelo $ARMA(p, q)$ os valores de p e q indicam respectivamente o número de parâmetros no modelo AR e MA. O modelo é dado por

$$\tilde{X}_t = \phi_1 \tilde{X}_{t-1} + \ldots + \phi_p \tilde{X}_{t-p} + a_t - \theta_1 a_{t-1} - \ldots - \theta_q a_{t-q} \tag{3.1}$$

A estimação de parâmetros do modelos ARMA é mais difícil do que o modelo AR, e envolve a resolução de equações não-lineares [Box et al. 1994]. Na prática podem ser examinadas determinadas propriedades da função de auto-correlação e auto-correlação parcial da série na tentativa de determinar os valores de p e q. Neste modelo, soluções analíticas são mais difíceis de se obter. A geração da série pode ser realizada da mesma maneira mostrada nas Tabelas 3.1 e 3.2

Modelo ARIMA

3.7 Modelo ARIMA

O modelo *ARIMA*(*Auto Regressive Integrated Moving Average*) consiste de uma extensão do modelo $ARMA(p, q)$ e é dado por $ARIMA(p, d, q)$ onde ∇^d é um operador de diferenças, definido como $\nabla^d X_t = (X_t - X_{t-d})$. A série original é submetida ao operador de diferenças e sobre a série diferenciada aplica-se um modelo ARMA.

No modelo ARIMA, o operador de diferenças pode ser obtido por

$$\nabla^d X_t = \sum_{i=0}^{d} \binom{d}{i} (-1)^i X_{t-i}, d \in \mathbb{N} \tag{3.2}$$

A diferenciação da série normalmente é aplicada na tentativa de torná-la estacionária. Um exemplo de diferenciação é mostrado na seção 5.3.4. Desta maneira, este modelo pode ser utilizado em séries não estacionárias. Uma descrição completa do modelo é encontrada em [Box et al. 1994] e softwares de análise em [R Development Core Team 2005].

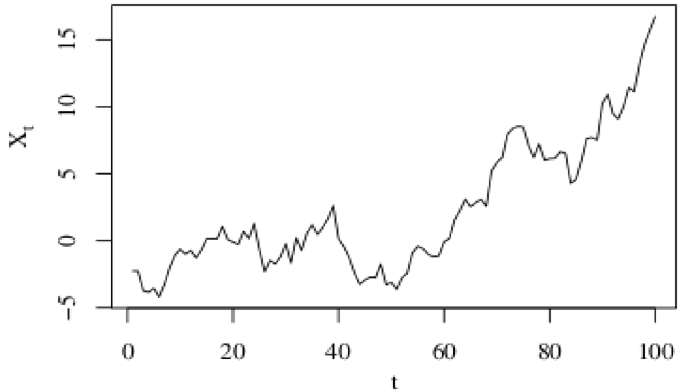
Modelo ARIMA

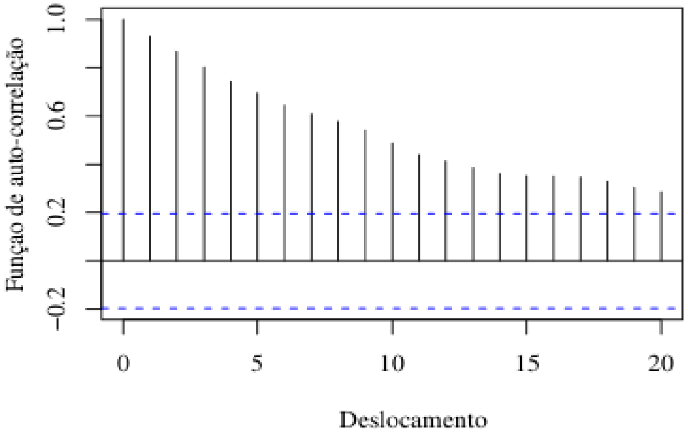
Exemplo 3.3 *Considere um modelo ARIMA(1,1,1), com parâmetros* $\phi_1 = 0,101270$ *e* $\theta_1 = 0,110472$. *O valor de* $\nabla^d X_t$ *será dado por* $\{0,101270X_{t-1} + a_t\} + \{-0,110472a_{t-1}\}$. *O valor de* X_t *pode ser obtido fazendo-se* $X_t = \nabla^d X_t + X_{t-1}$. *O resultado é mostrado na Tabela* $\boxed{3.3}$. *A série e sua função de auto-correlação são mostrados na Figura* $\boxed{3.1}$. *A função de auto-correlação mostra um decaimento lento, que pode ser confundido com a não estacionariedade da série ou mesmo com a presença de características auto-similares.*

Modelo ARIMA

Tabela 3.3: Exemplo de geração de uma série com modelo ARIMA.

t	$\tilde{X}_t$	$\nabla^d \tilde{X}_t$	a_t
1	-2,27228	-2,27228	-2,27228
2	-2,26659	0,00569	-0,01522
3	-3,71876	-1,45217	-1,45212
4	-3,84008	-0,12132	-0,13468
5	-3,55374	0,28634	0,28523
⋮	⋮	⋮	⋮





Exercícios

1. Recupere a série de valores históricos para do valor de fechamento da cotação do dolar, com 100 valores passados.
   - Esboce o gráfico da série. Analise o gráfico resultante.
   - Esboce o gráfico da função de auto-correlação (e auto-correlação parcial). Analise o gráfico resultante.
   - Estabeleça uma hipótese sobre um modelo adequado utilizando ARIMA.
   - Utilize o software R (pacote **fArima**) para estimar os parâmetros.
   - Como verificar se o modelo proposto é adequado? (utilize QQPLOT dos resíduos)

Exercícios

2. Suponha a série com o número de mortes em decorrência da Gripe A, em Curitiba, no inverno de 2009.
   - Esboce o gráfico da série. Analise o gráfico resultante.
   - Esboce o gráfico da função de auto-correlação (e auto-correlação parcial). Analise o gráfico resultante.
   - Estabeleça uma hipótese sobre um modelo adequado utilizando ARIMA.
   - Utilize o software R (pacote **fArima**) para estimar os parâmetros.
   - Como verificar se o modelo proposto é adequado? (utilize QQPLOT dos resíduos)

Série – Gripe A

Gripe A (mortes/dia – 2009)		*10*	*20*	*30*	*40*	*50*	*60*
1	2	1	1	12	3	2	2
2	0	4	2	7	7	3	0
3	1	6	6	8	5	0	1
4	1	7	8	10	4	2	1
5	0	4	13	7	4	3	1
6	3	6	9	6	6	3	
7	3	0	5	6	4	1	
8	2	4	7	2	5	0	
9	3	4	4	2	1	2	

Séries Temporais – outros modelos

- Média Móvel
  - Esse método considera como previsão para o período futuro a média das observações passadas recentes

$$x_t = \frac{x_{t-1} + x_{t-2} + \ \dots \ + x_{t-n}}{n}$$

  - O termo média móvel é utilizado porque à medida que a próxima observação se torna disponível, a média das observações é recalculada, incluindo essa observação no conjunto de observações e desprezando a observação mais antiga

Séries Temporais – outros modelos

- Alisamento Exponencial Simples (EWMA Exponentially Weighted Moving Average

$$S_t = \alpha X_t + (1\text{-}\alpha)S_{t-1}$$

  – Insere atraso na série.
  – Também conhecido como filtro passa baixa.

S_t representa a série e o parâmetro α deve estar entre 0 e 1.

Séries Temporais – outros modelos

- Alisamento Exponencial Simples
  - Quanto menor o valor de α, mais estáveis serão as previsões, visto que a utilização de baixo valor de implica na atribuição de peso maior às observações passadas e, conseqüentemente, qualquer flutuação aleatória no presente contribui com menor importância para a obtenção da previsão.
  - Não há metodologia que oriente quanto à seleção de um valor apropriado para a, sendo normalmente encontrado por tentativa e erro.

Séries Temporais – outros modelos

- Alisamento Exponencial Linear
  - Quando o método Alisamento Exponencial Simples é aplicado na previsão de séries temporais que apresentam tendência entre as observações passadas, os valores prognosticados superestimam (ou subestimam) os valores reais.
  - Para evitar esse erro sistemático, o método Alisamento Exponencial Linear foi desenvolvido procurando reconhecer a presença de tendência na série de dados

Séries Temporais – outros modelos

- Alisamento Exponencial Linear

$$F_{t+m} = S_t + mT_t$$

$$S_t = \alpha x_t + (1-\alpha)(S_{t-1} + T_{t-1})$$

$$T_t = \beta(S_t - S_{t-1}) + (1-\beta)T_{t-1}$$

Onde α é o peso atribuído à observação $0<\alpha<1$

e β é o coeficiente de alisamento, $0<\beta<1$

Séries Temporais – outros modelos

- Alisamento Exponencial Sazonal e Linear de Winter

$$S_t = \alpha \frac{x_t}{I_{t-l}} + (1-\alpha)(S_{t-1} + T_{t-1})$$

$$T_t = \beta(S_t - S_{t-1}) + (1-\beta)T_{t-1}$$

$$I_t = \gamma \frac{x_t}{S_t} + (1-\gamma) I_{t-l}$$

$$F_{t+m} = (S_t + mT_t) I_{t-l+m}$$

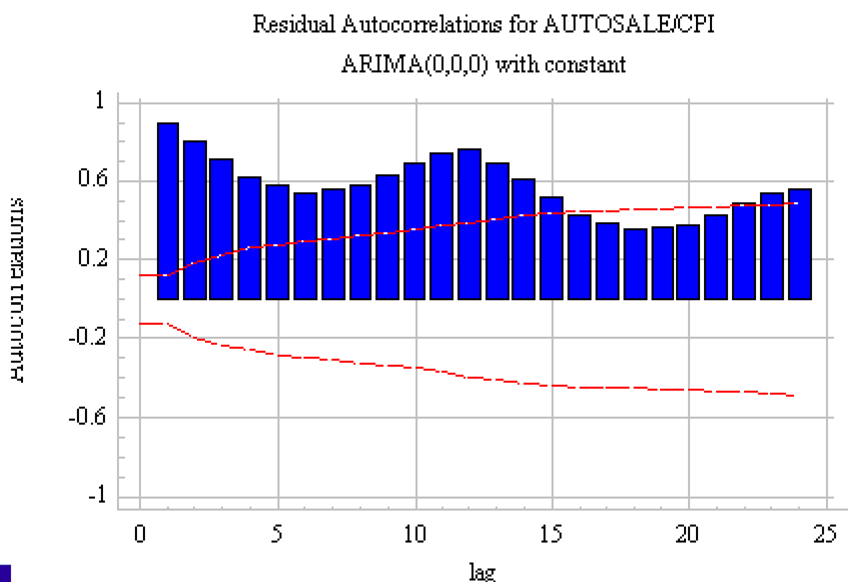
onde I_t corresponde ao alisamento do fator de sazonalidade $\frac{x_t}{S_t}$;

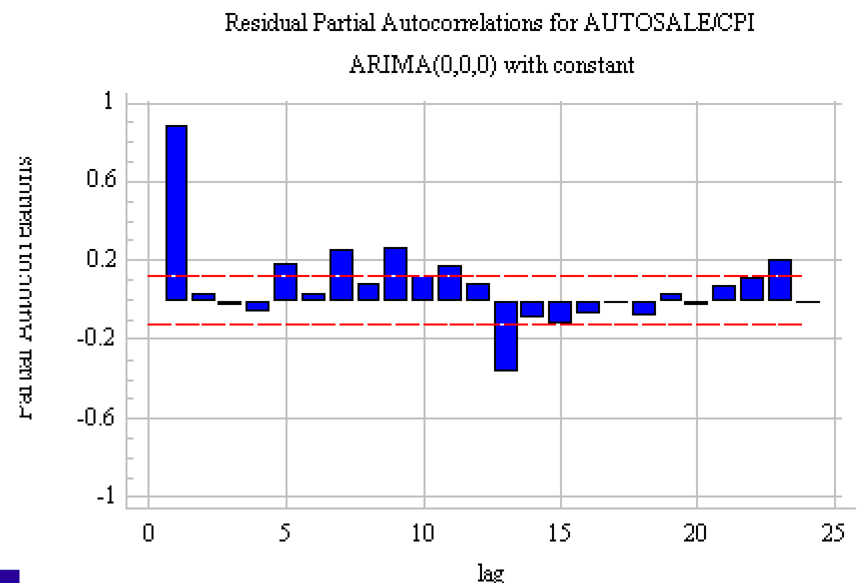
l é o intervalo da sazonalidade e

γ corresponde ao peso atribuído ao fator de sazonalidade.

Séries Temporais – outros modelos

- Modelos SARIMA:
  - Arima Sazonal.
  - ARIMA(p,d,q)x(P,D,Q)
    - P,D,Q são os componentes sazonais





Exemplo 1

- Considere a série disponível em *www.eletrica.ufpr.br/pedroso/2018/EELT7029/serie1.dat*
- *Determine o modelo usando AR(p), MA(q), ARMA(p,q)*

  1) Carregando os dados

     ```
     R> s<-scan("serie1.dat")
     ```

  2) Examine a série

     ```
     R> plot(s,type="l")
     ```

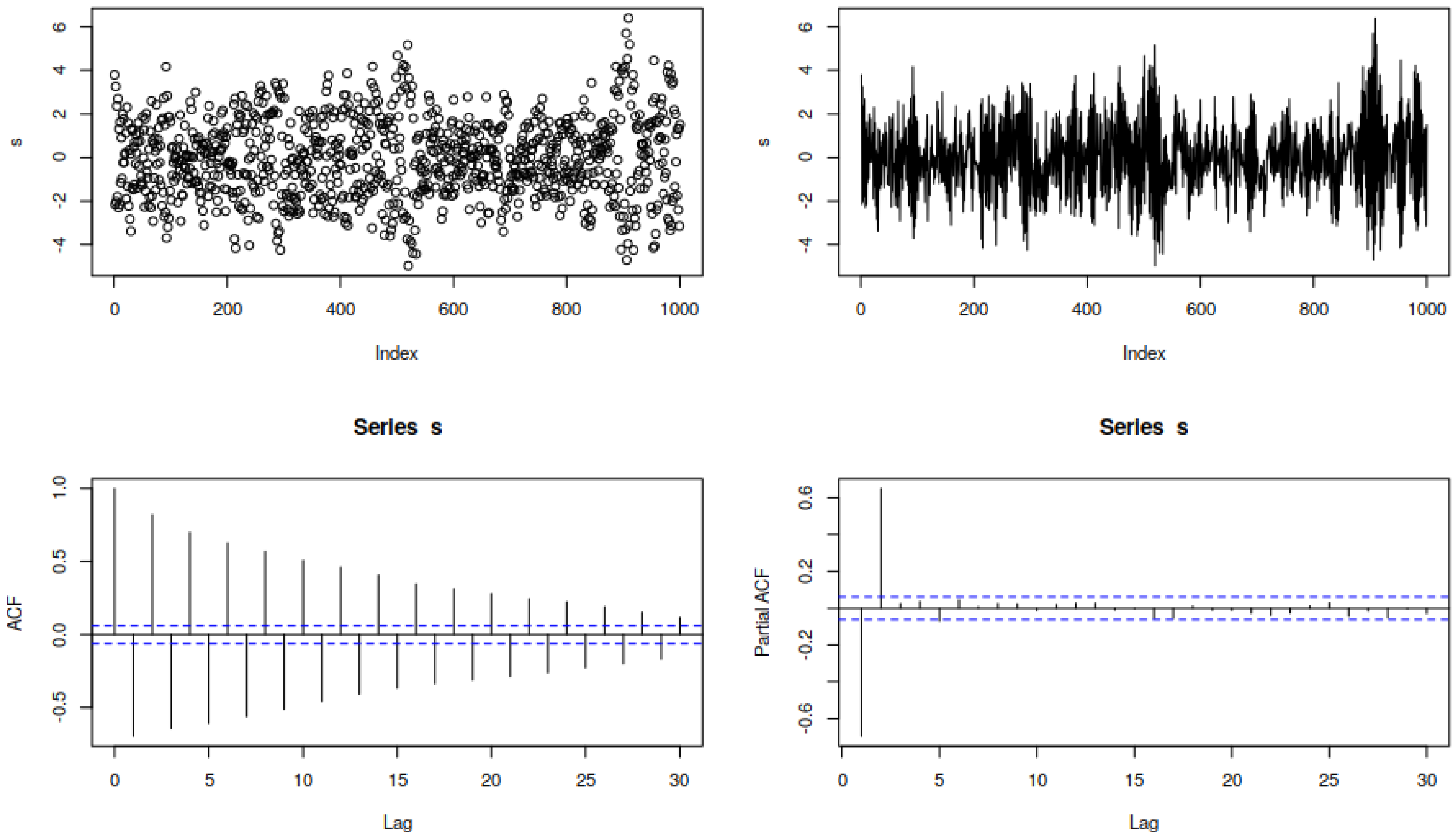
  3) Examine a função de auto correlação

     ```
     R> acf(s)
     ```

  4) Examine a função de auto correlação parcial

     ```
     R> pacf(s)
     ```

Exemplo 1



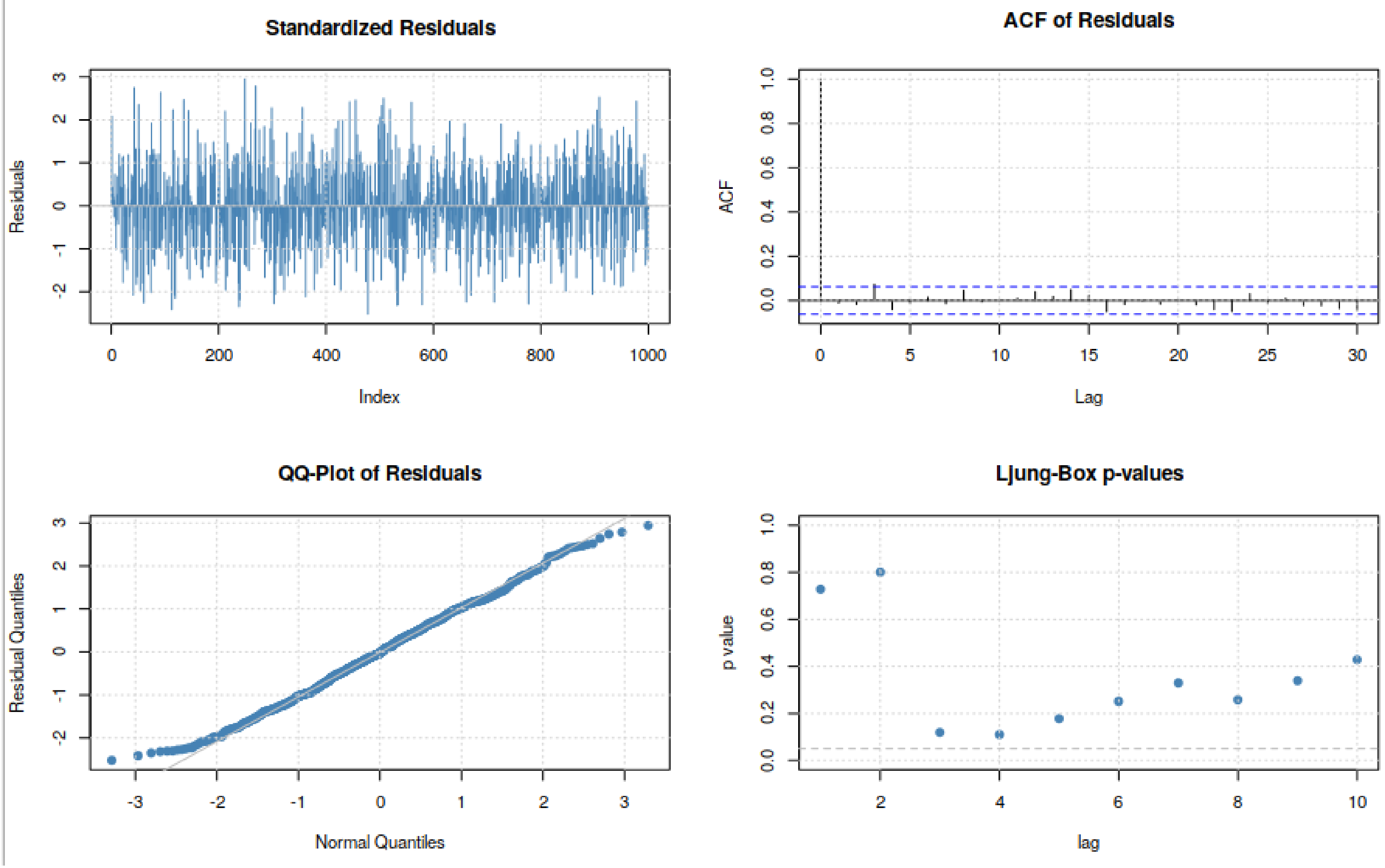
Exemplo 1

- Hipótese: Modelo AR(2)

```
library(fArma)
fit = armaFit(~ar(2), data = s)
printf(fit)
par(mfrow = c(2, 2)
summary(fit,which="all")
```

```
Coefficient(s):
     ar1       ar2      intercept
-0.24095   0.65732   -0.02361
```

Exemplo 1

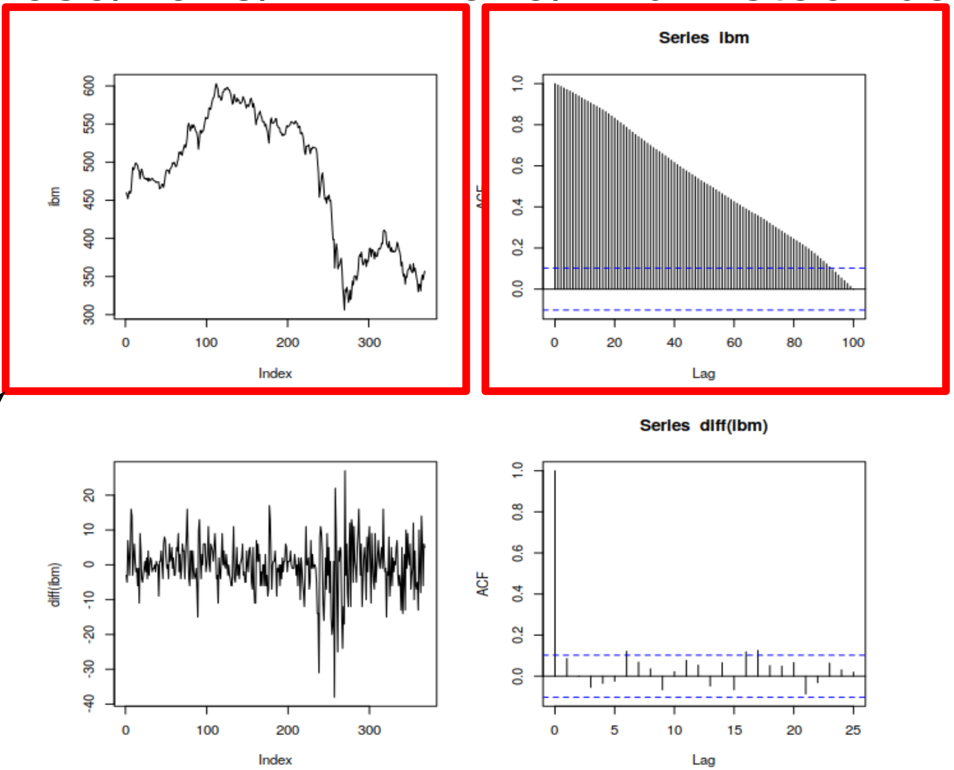


Standardized Residuals
Residuals
Index
ACF of Residuals
ACF
Lag
QQ-Plot of Residuals
Residual Quantiles
Normal Quantiles
Ljung-Box p-values
p value
lag

Exemplo 2

Considere a série do preço das ações da IBM disponível em:
www.eletrica.ufpr.br/pedroso/2018/EELT7029/B.ibm-stock.dat

```
> plot(ibm,type="l")
> acf(ibm,lag.max=100)
> plot(diff(ibm),type="l")
> acf(diff(ibm))
```



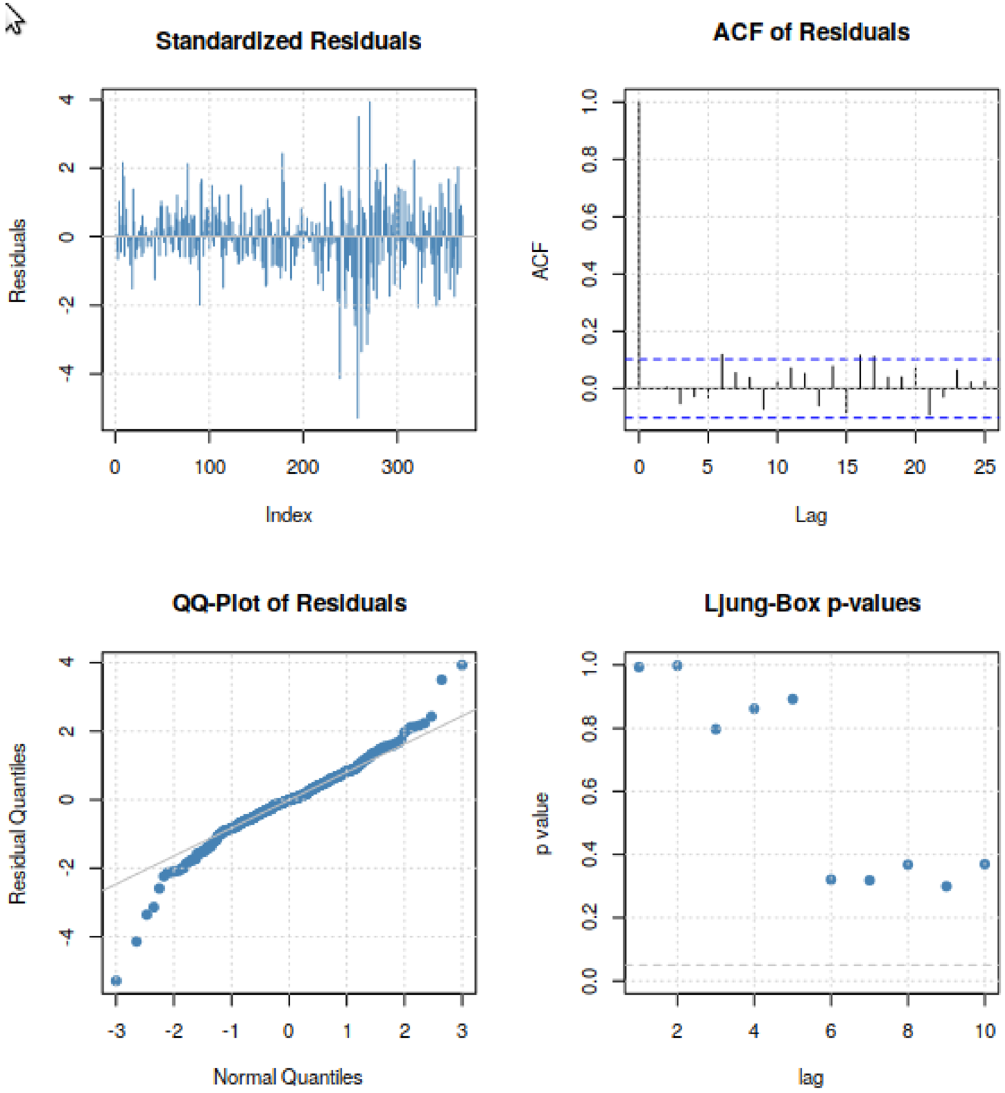
Observe:
1. *que a série não é estacionária*
2. *a série diferenciada é estacionária*
3. ***não** existe autocorrelação significativa na série diferenciada*

Exemplo2

Hipótese: ARIMA(0,1,1)

```
fit = armaFit(~ arima(0,1,1), data = ibm)
print(fit)
par(mfrow = c(2, 2), cex = 0.7)
summary(fit, which = "all")
```

```
Coefficient(s):
    ma1
0.08635
```

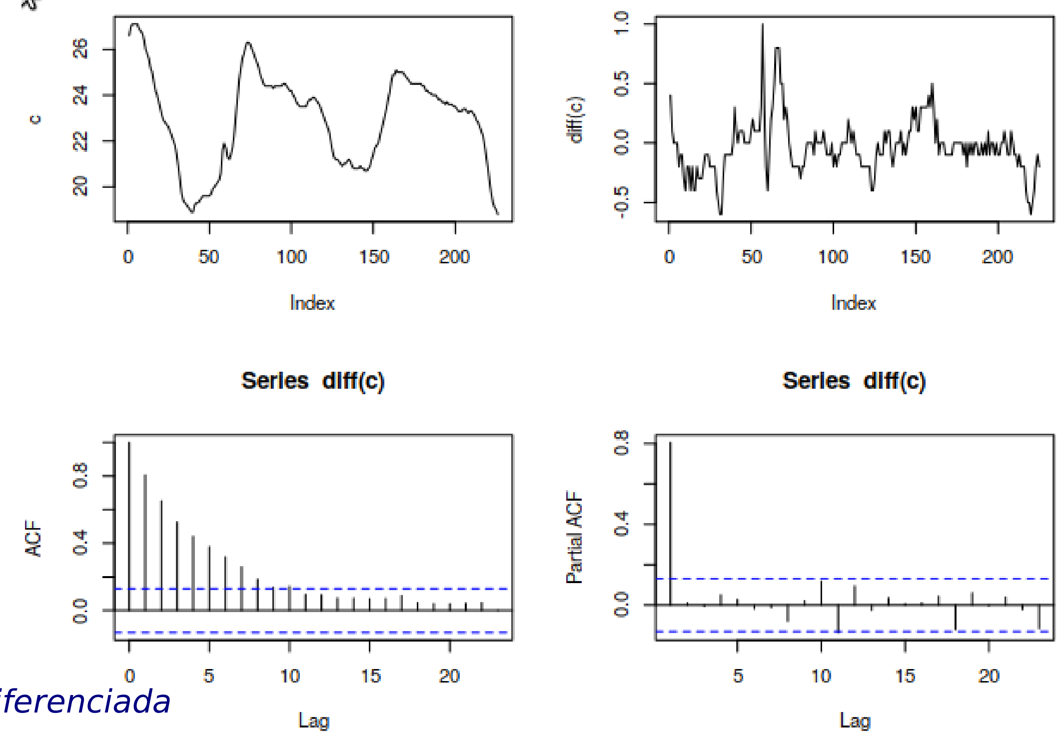


Exemplo 3

Considere a série:

www.eletrica.ufpr.br/pedroso/2018/EELT7029/C.chem-temp.dat

```
c<-scan("C.chem-temp.dat")
par(mfrow = c(2, 2))
plot(c,type="l")
plot(diff(c),type="l")
acf(diff(c))
pacf(diff(c))
```



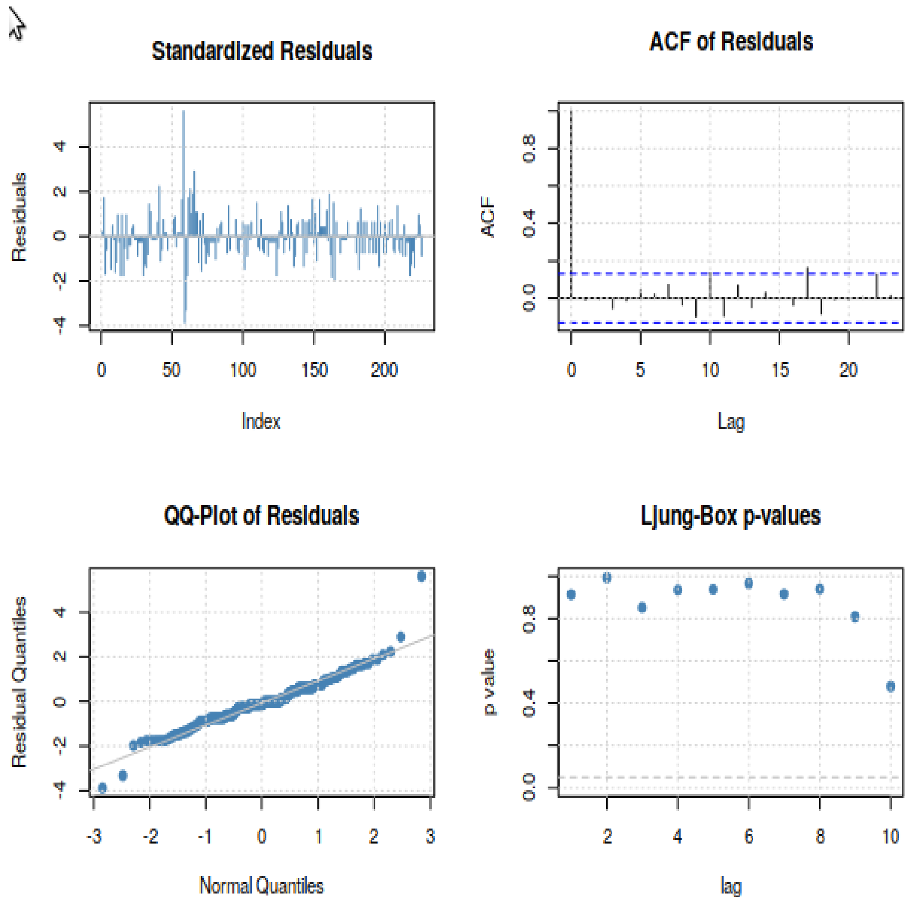
Observe:
1. que a série não é estacionária
2. a série diferenciada é estacionária
3. existe autocorrelação significativa na série diferenciada

Exemplo3

Hipótese: ARIMA(1,1,0)

```
fit = armaFit(~ arima(1,1,0), data = c)
print(fit)
par(mfrow = c(2, 2))
summary(fit, which = "all")
```

Coefficient(s):
ar1
0.8202

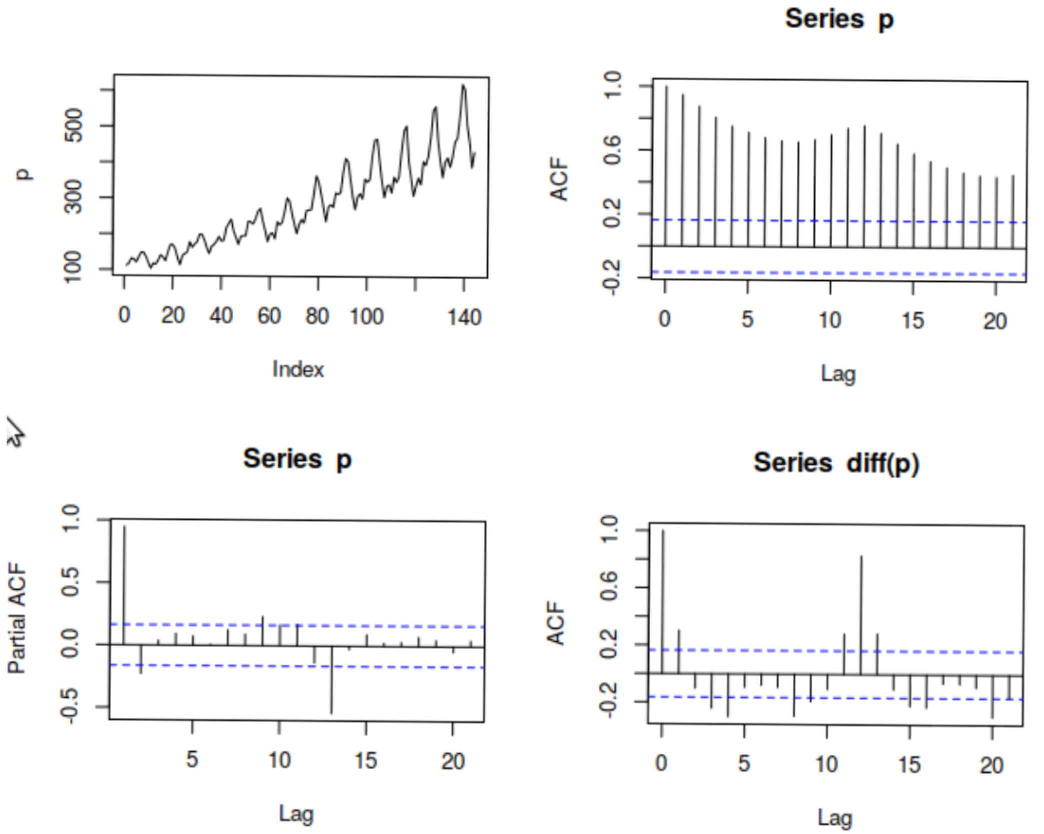


Exemplo 4

Considere a série:

www.eletrica.ufpr.br/pedroso/2018/EELT7029/G.airline-pass.dat

```
p<-scan("G.airline-pass.dat")
par(mfrow = c(2, 2))
plot(p,type="l")
acf(p)
pacf(p)
plot(diff(p),type="l")
```



Observe:
1. que a série não é estacionária
2. que a autocorrelação parcial indica uma dependência temporal nos lags 1 e 12

Exemplo4

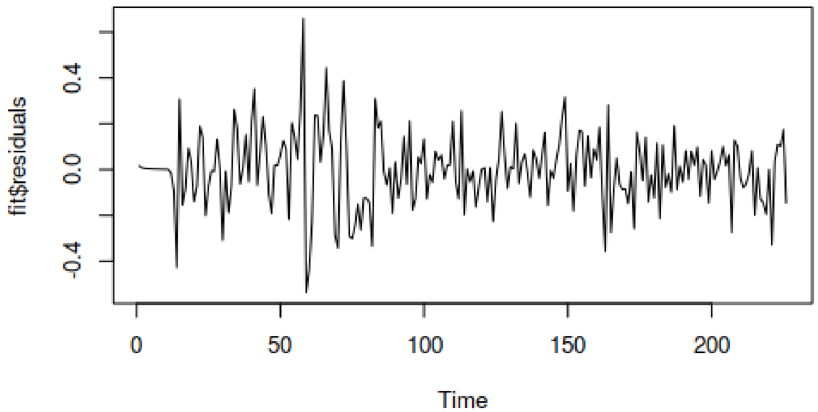
Hipótese:

$$ARIMA(1,1,0)x(1,1,0)_{12}$$

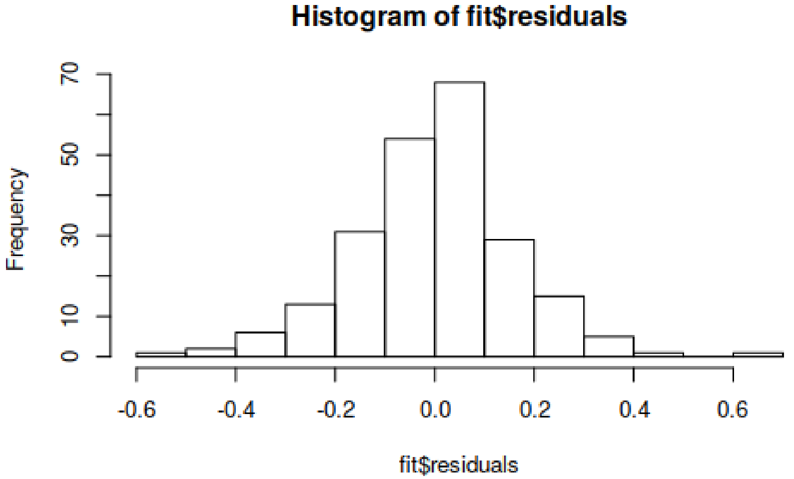
```
fit = arima(c, order=c(1,1,0), seasonal=list(order=c(1,1,0),period=12))
print(fit)
par(mfrow = c(2, 2))
plot(fit$residuals)
hist(fit$residuals)
library(car)
qqPlot(fit$residuals,distribution="norm",mean=mean(fit$residuals),sd=sd(fit$residuals))
acf(fit$residuals)
```

```
Coefficient(s):
       ar1     sar1
    0.8192  -0.4370
sigma^2 estimated as 0.02719
```

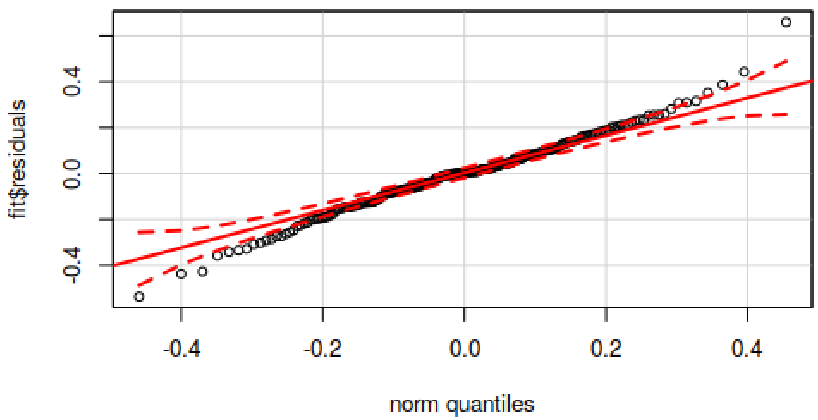
Exemplo4



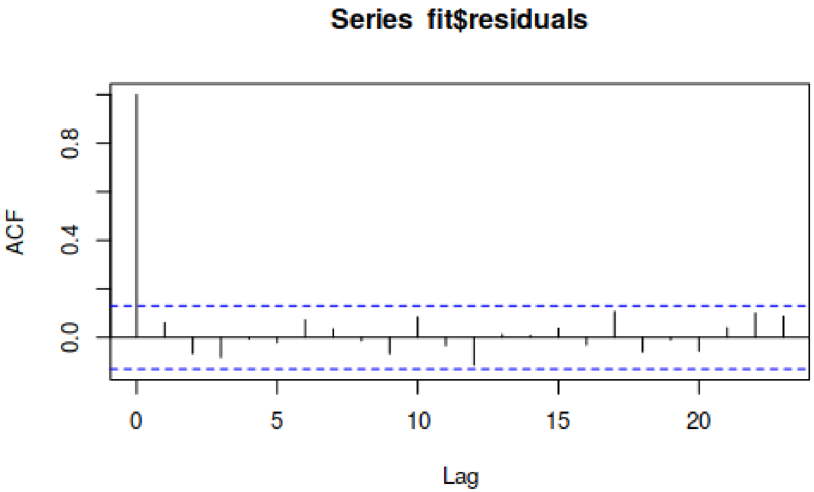
fit$residuals
Time



Histogram of fit$residuals
Frequency
fit$residuals



fit$residuals
norm quantiles



Series fit$residuals
ACF
Lag