  
Universidade Federal do Pará

Instituto de Ciências Exatas e Naturais

3ª Avaliação de Análise de Séries Temporais

**3ª Avaliação de séries temporais**

Aluna: Michelle da Costa Santana

Belém

2015

1. **Introdução**

Neste trabalho encontra-se a resolução das questões referentes à terceira avaliação da disciplina análise de séries temporais, da faculdade de estatística da Universidade Federal do Pará.

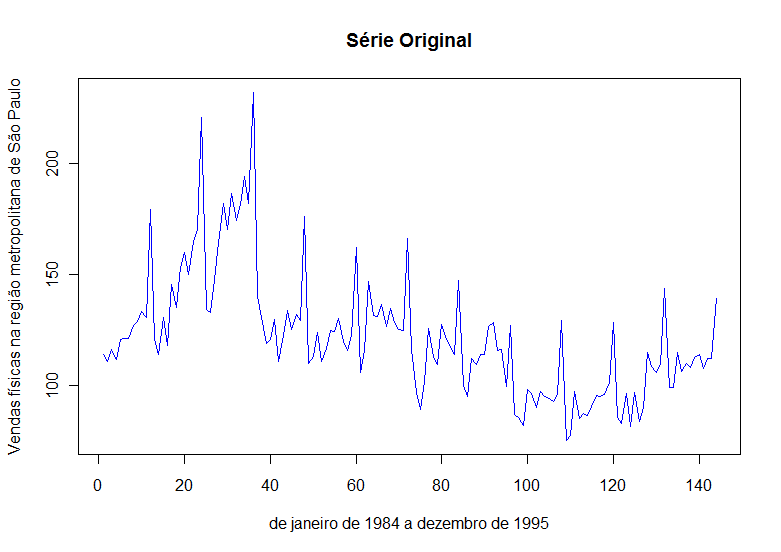
**2 – Metodologia**

Para realização da simulação foi utilizado o software R, versão 2.12.2, foi utilizado também o Microsoft Excel para construção de algumas tabelas.

**Série 1: A11 Consumo**

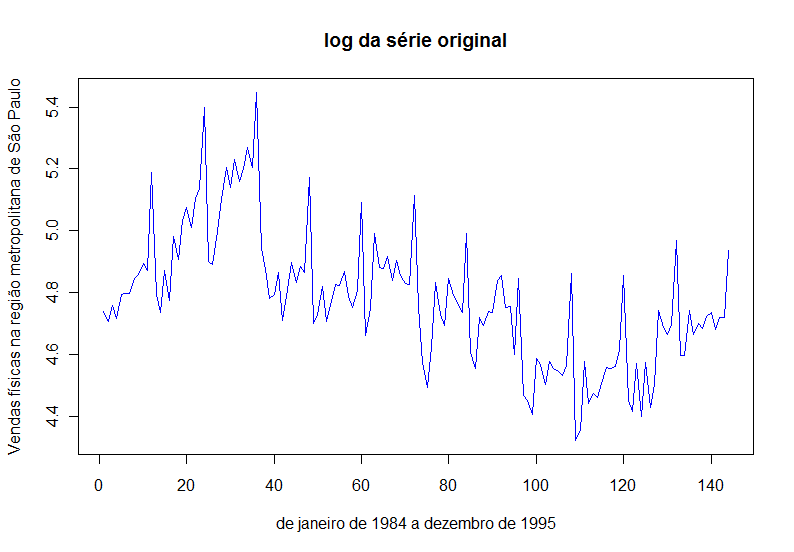
**Identificação, Estimação, Diagnóstico de Modelos e Previsão**

**Série 1: A11 Consumo** – Livro texto: Vendas físicas na região metropolitana de São Paulo, observações mensais de janeiro de 1984 a outubro de 1996. Foram utilizadas 144 das 154 observações.



A série ilustrada no gráfico acima aparentemente apresenta sazonalidade, o que será verificado com a função de auto correlação. Já em relação a tendência, pode-se se perceber a presença de uma tendência razoavelmente decrescente na série.

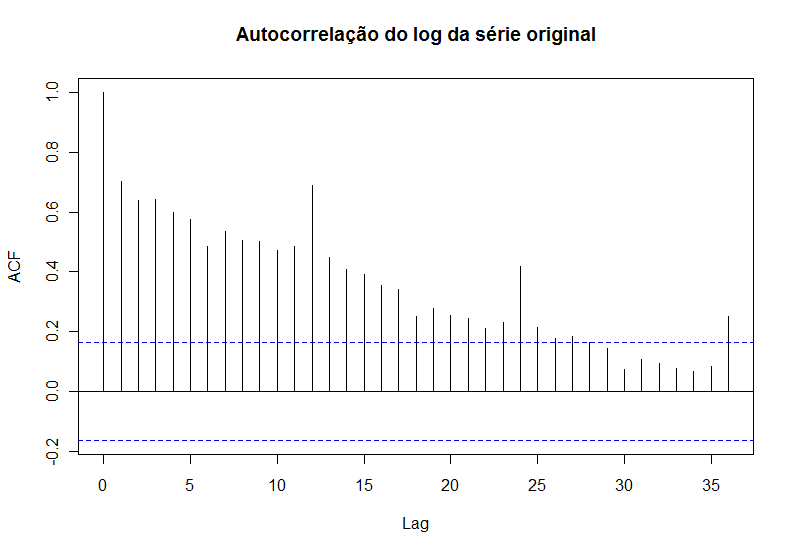
Será feito transformação logarítmica nos dados (esta transformação é sugerida na literatura). A adição da função logarítmica em uma série temporal faz com a variância dos dados se torne mais estável, o que é suficiente para obter a homogeneidade.



**Identificação**

A série é estacionária? Para resposta desta pergunta fez-se a Função de Autocorrelação (FAC) de y, onde y = log (série original).

**acf(y,lag.max=36, plot=T,main="Autocorrelação do log da série original")**

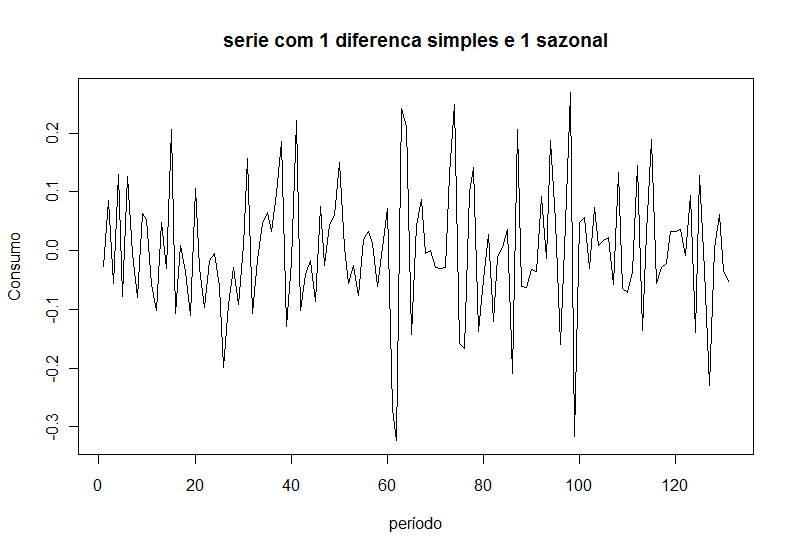


Pode-se observar no gráfico acima que a série é não estacionária e que ainda há sazonalidade, onde o período é igual a 12, pois há uma alta correlação nos lags 12, 24 e 36. Logo será feito uma diferença simples e uma diferença sazonal e verificar como se comporta a FAC da série com essas duas diferenças.

**z =(diff(diff(y,lag=12)))**

**plot(z,main='serie com 1 diferenca simples e 1 sazonal',xlab='período',**

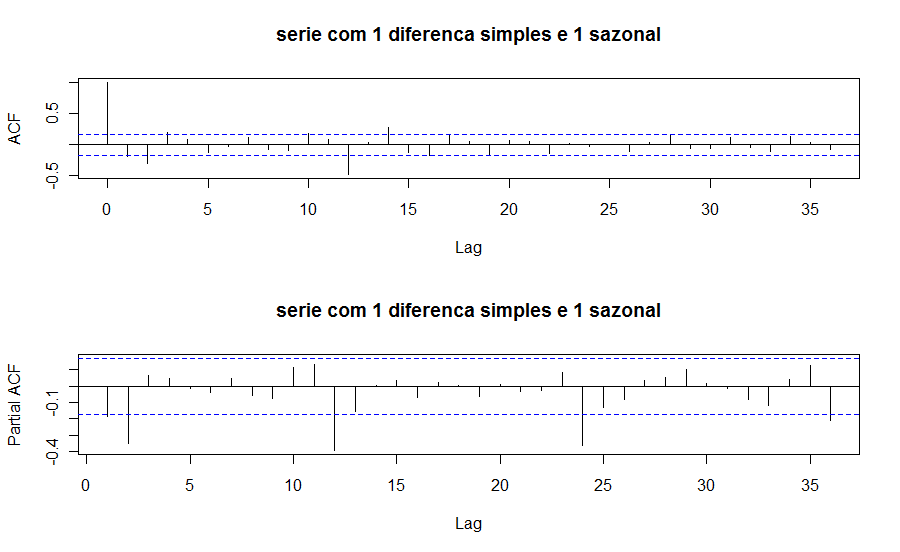
**ylab='Consumo',type="l")**



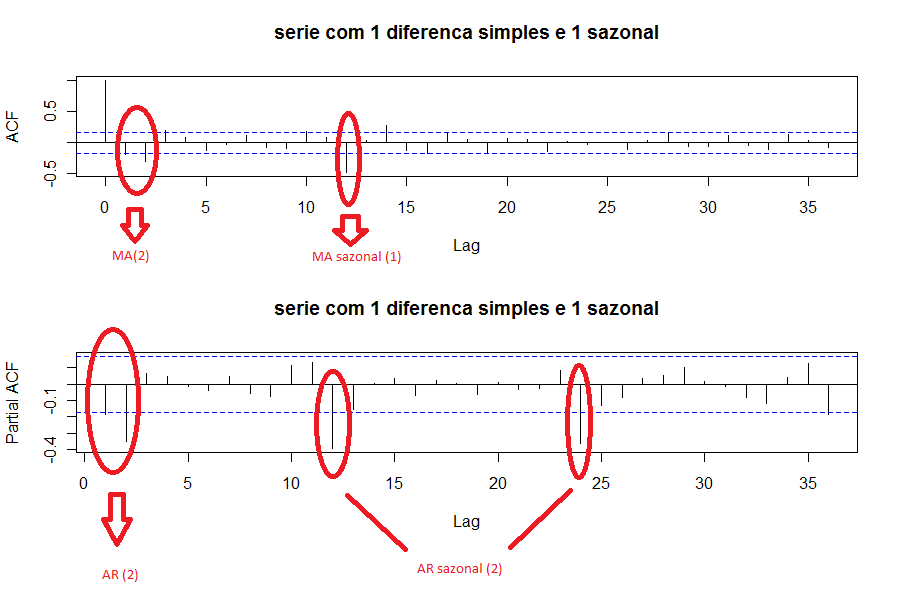
**par(mfrow=c(2,1))**

**ACFZ = acf(z, lag.max=36,main='serie com 1 diferenca simples e 1 sazonal')**

**ACPZ<-pacf(z, lag.max =36,main='serie com 1 diferenca simples e 1 sazonal')**



Ao fazer a autocorrelação e a autocorrelação parcial, identificou-se um **SARIMA(2,1,2)(2,1,1)** com período igual a 12. Os componentes de médias móveis afetam a autocorrelação e os componentes autorregressivos afetam a autocorrelação parcial. Veja a imagem abaixo e entenda.



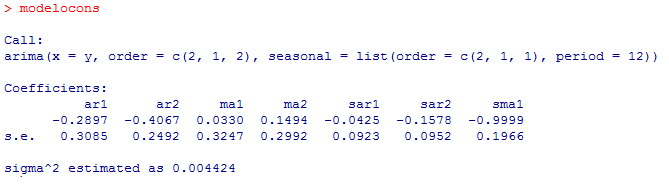
**Estimação**

Para estimação dos parâmetros utilizou-se o seguinte comando:

**modelocons=arima(y,order=c(2,1,2),seasonal=list(order=c(2,1,1),period=12))**

**modelocons**

**Parâmetros Estimados:**



Modelo obtido é:

(Y(t) - phi1Y(t-1) – phi2Y(t-2)) (Y(t) - PHI1Y(t-12) – PHI2Y(t-24)) Y(t) =

(a(t) - teta1a(t-1) – teta2a(t-2)) (a(t) - TETA1a(t-12))

Onde Y(t) = log(Z(t)).

Substituindo os valores:

(Y(t) – (-0,2897)Y(t-1) – (-0,4067)Y(t-2)) (Y(t) – (-0,0425)Y(t-12) – (-0,1578)Y(t-24)) Y(t) = (a(t) – 0,033a(t-1) – 0,1494a(t-2)) (a(t) – (-0,9999)a(t-12))

**Diagnóstico**

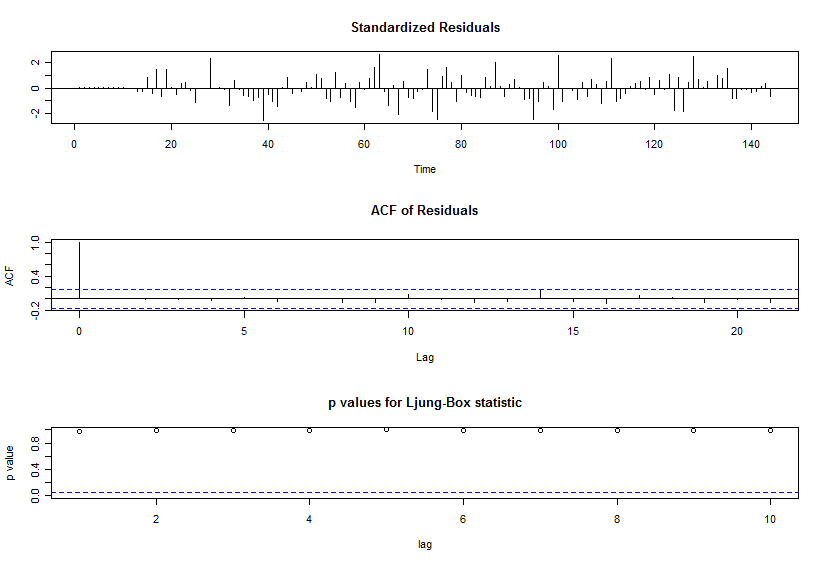
**#Diagnóstico**

**tsdiag(modelocons)**

**res=modelocons$residuals #retorna o valor dos resíduos**

Pode-se observar no gráfico dos resíduos padronizados que eles se distribuem em torno de média 0, no gráfico da autorrelação nota-se que não há correlação significativa, o que sugere que o modelo é adequado.

O teste de Ljung-Box, verifica se os resíduos podem ser considerados ruído branco, a hipótese nula é de que os erros gerados pelo modelo é ruído branco, isto é, rejeita-se a hipótese nula quando o p-valor é menor ou igual a 5%. Como podemos verificar no gráfico deste teste os valores estão acima de 5%, logo os resíduos são independentes e identicamente distribuídos, ou seja, o resíduo é ruído branco. Isso nos mostra que o modelo é adequado.



**Previsão**

Para previsão será utilizada a função predict do software R, este código é baseado no modelo obtido.

**#Previsão**

**pv=predict(modelocons,n.ahead=10) # n.head= passos a frente que se deseja**

**pv**

**prev=matrix(pv$pred) #Valores da Previsão**

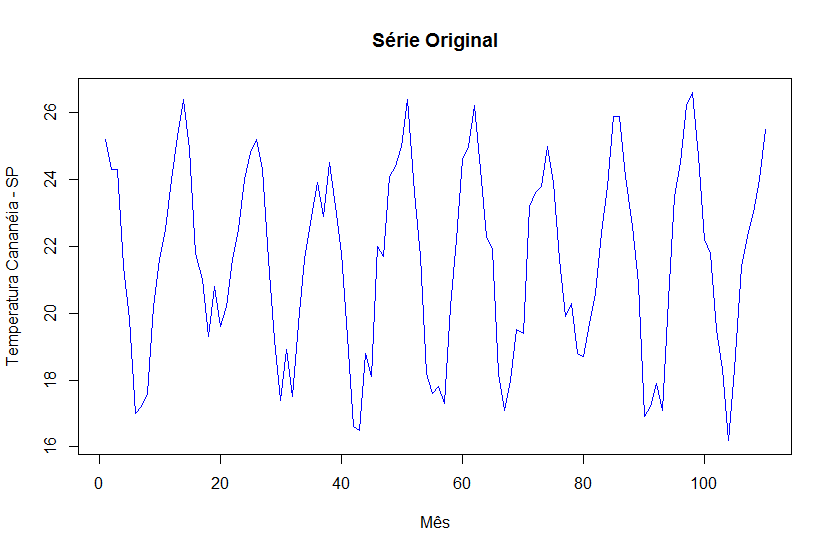
**prev**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Período | Série | Log (série) | Previsão para o log | Erro % |
| jan/96 | 92.24 | 4,524394 | 4,5113 | 0,3 |
| fev/96 | 93.56 | 4,538603 | 4,5009 | 0,8 |
| mar/96 | 107.37 | 4,676281 | 4,5693 | 2,3 |
| abr/96 | 102.89 | 4,633660 | 4,5660 | 1,5 |
| mai/96 | 114.78 | 4,743017 | 4,6376 | 2,2 |
| jun/96 | 102.88 | 4,633563 | 4,6087 | 0,5 |
| jul/96 | 118.41 | 4,774153 | 4,6469 | 2,7 |
| ago/96 | 119.23 | 4,781054 | 4,6742 | 2,2 |
| set/96 | 117.36 | 4,765246 | 4,6370 | 2,7 |
| out/96 | 122.06 | 4,804513 | 4,6553 | 3,1 |

**Série 2: A1 Temperatura Cananéia - SP**

**Identificação, Estimação, Diagnóstico de Modelos e Previsão**

**Série 2: A1 temperatura** – Livro texto: Temperaturas médias mensais, em graus centígrados, de janeiro de 1976 a dezembro de 1985, em Cananéia, SP. As observações utilizadas para simulação é de janeiro/76 a fevereiro/85, ou seja, 110 observações.

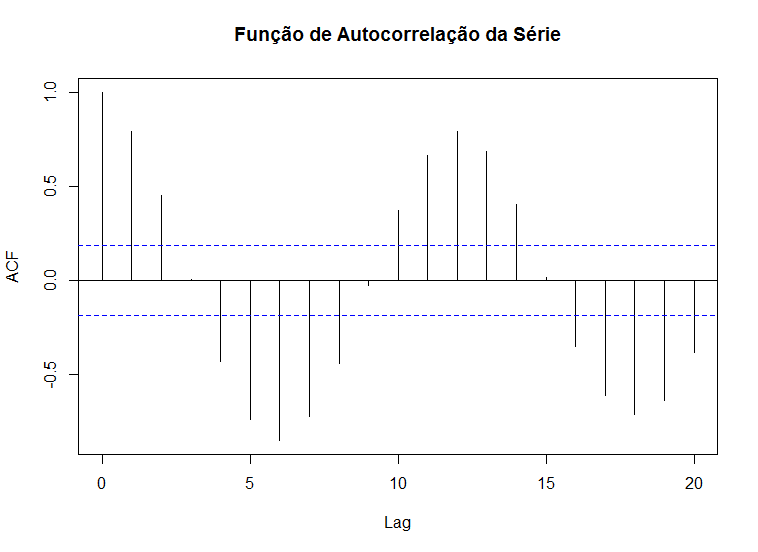


Observa-se na série ilustrada acima, a presença de componentes sazonais.

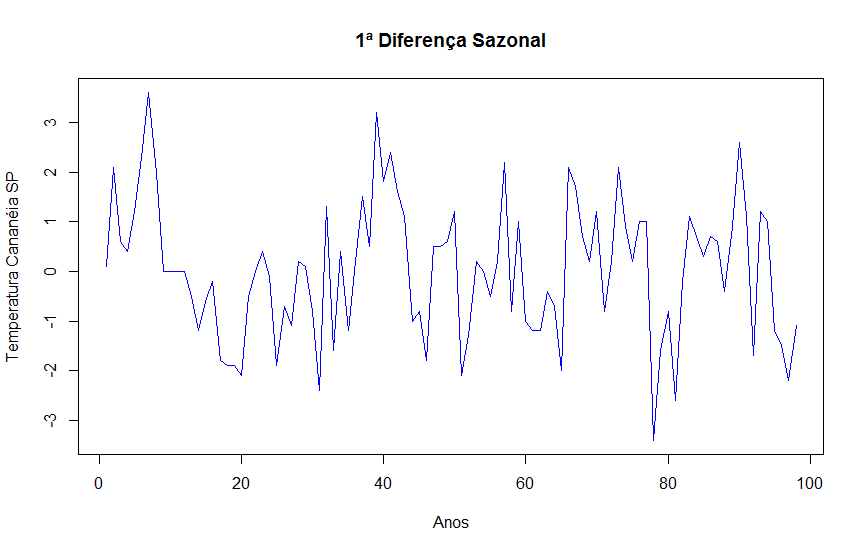
**Identificação**

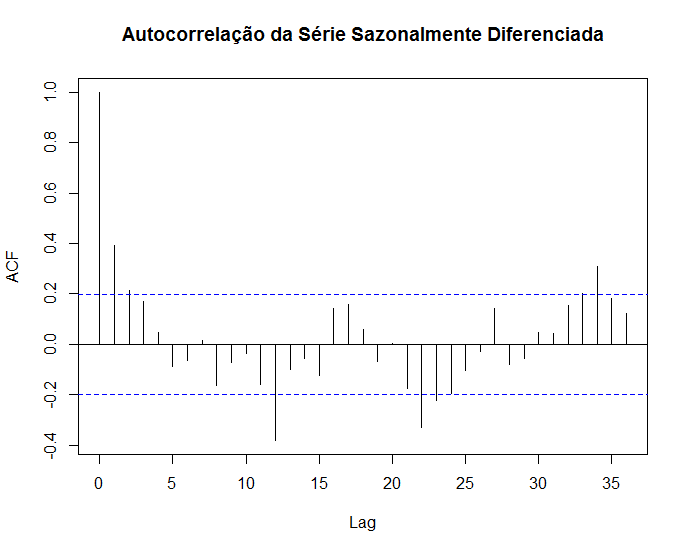
Função de Autocorrelação:

**acf(x,lag.max=20, plot=T, main="Função de Autocorrelação da Série") # x=série**



A partir do gráfico acima pode-se dizer que a série não é estacionária, que há sazonalidade com período igual a 12. Para tentar tornar a série estacionária faz-se primeiramente uma diferença sazonal. Abaixo tem-se a série com uma diferença sazonal.





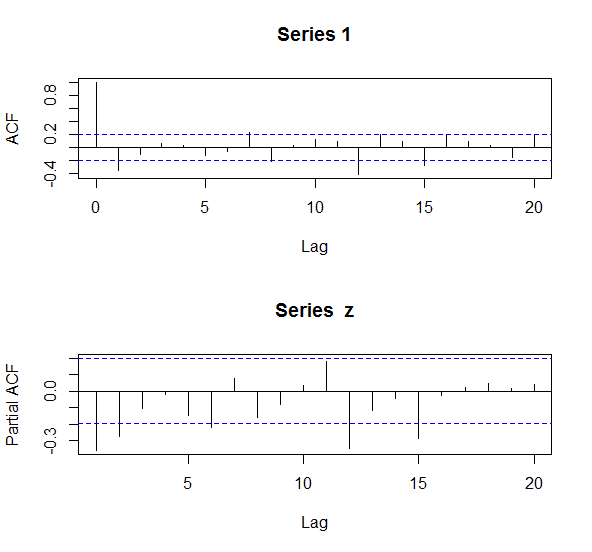
Fez-se uma diferença simples e uma sazonal, obteve-se as seguintes autocorrelações:

**z =(diff(diff(x,lag=12)))**

**par(mfrow=c(2,1))**

**ACFZ = acf(z, lag.max=20)**

**ACPZ<-pacf(z, lag.max =20)**

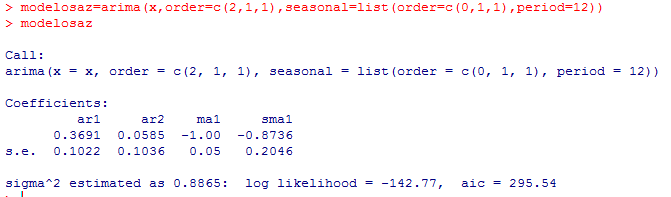


Ao analisar os gráfico da autocorrelação acima observa-se que há uma correlação significativa no lag 1 e 12, logo tem-se MA(1) e MA sazonal (1). Analisando a autocorrelação parcial nota-se autocorrelação significativa no dois primeiros lags e nos lags 12 e 15, isto indica um AR(2) e AR sazonal(1). Logo o modelo identificado foi: **SARIMA(2,1,1)(0,1,1).** Caso o lag 15 continue significativo após a modelagem, acrescentaremos um parâmetro o phi15.

**Estimação**

**modelosaz=arima(x,order=c(2,1,1),seasonal=list(order=c(0,1,1),period=12))**

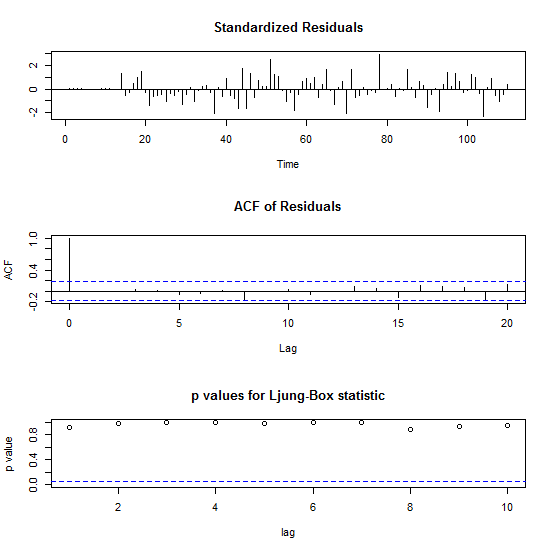
**Parâmetros Estimados:**



**Diagnóstico**

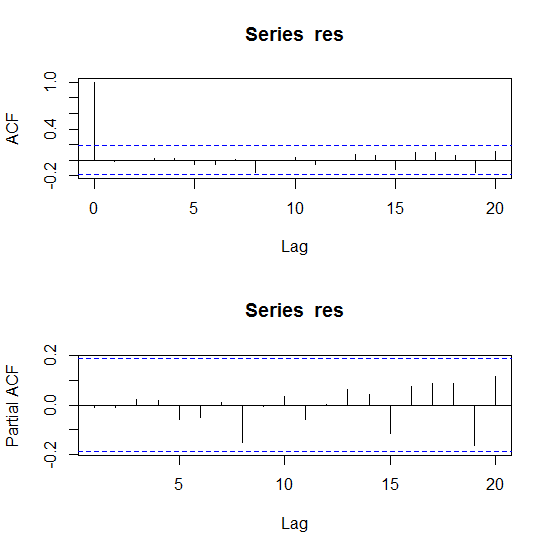
**tsdiag(modelosaz)**

**res=modelosaz$residuals**



Pode-se observar no gráfico dos resíduos padronizados que eles se distribuem em torno de média 0, no gráfico da autorrelação nota-se que não há correlação significativa, o que sugere que o modelo é adequado.

O teste de Ljung-Box, verifica se os resíduos podem ser considerados ruído branco, a hipótese nula é de que os erros gerados pelo modelo é ruído branco, isto é, rejeita-se a hipótese nula quando o p-valor é menor ou igual a 5%. Como podemos verificar no gráfico deste teste os valores estão acima de 5%, logo os resíduos são independentes e identicamente distribuídos, ou seja, o resíduo é ruído branco. Isso nos mostra que o modelo é adequado.



O gráfico acima mostra que não foi necessário a adição de mais um parâmetro para tornar a correlação no lag 15 não significativa.

**Previsão**

**#Previsão**

**pv=predict(modelosaz,n.ahead=10)**

**pv**

**prev=matrix(pv$pred)**

**prev**

|  |  |  |  |
| --- | --- | --- | --- |
| Período | Série | Previsão | Erro (%) |
| mar/85 | 25,30 | 24,37 | 3,67 |
| abr/85 | 23,80 | 22,09 | 7,17 |
| mai/85 | 20,10 | 20,68 | 2,87 |
| jun/85 | 18,00 | 18,22 | 1,21 |
| jul/85 | 18,00 | 17,99 | 0,07 |
| ago/85 | 19,10 | 17,90 | 6,30 |
| set/85 | 19,20 | 18,77 | 2,24 |
| out/85 | 21,60 | 20,91 | 3,21 |
| nov/85 | 22,60 | 22,59 | 0,06 |
| dez/85 | 24,00 | 23,90 | 0,43 |