**UNIVERSIDADE FEDERAL DE MINAS GERAIS**

Carlos Henrique Coelho Miyazawa - 2018051347

Gledes Moraleida Oliveira Gripp - 2018051118

Júlia Raquel Fonseca Freitas - 2018051096

Rafael Hilário Cruz Peixoto - 2016097951

**LISTA 1 – METODOS ESTATISTICOS DE PREVISAO**

Belo Horizonte – Minas Gerais

2020

Sumário

[Questão 1 3](#_Toc63797275)

[Questão 5 3](#_Toc63797276)

[Questão 7 4](#_Toc63797277)

[Questão 40 4](#_Toc63797278)

[Leitura dos dados 4](#_Toc63797279)

[Série V1 6](#_Toc63797280)

[Modelo 1 6](#_Toc63797281)

[Variância dos coeficientes 6](#_Toc63797282)

[P-valor dos coeficientes 7](#_Toc63797283)

[Teste “Augmented Dickey-Fuller” 7](#_Toc63797284)

[Analise de Resíduos 7](#_Toc63797285)

[Teste de Ljung 10](#_Toc63797286)

[Série V2 11](#_Toc63797287)

[Modelo 2 11](#_Toc63797288)

[Variância dos coeficientes 11](#_Toc63797289)

[Teste “Augmented Dickey-Fuller” 11](#_Toc63797290)

[Teste “Augmented Dickey-Fuller” 12](#_Toc63797291)

[Analise de Resíduos 12](#_Toc63797292)

[Teste de Ljung 14](#_Toc63797293)

[Série V3 16](#_Toc63797294)

[Modelo 3 16](#_Toc63797295)

[Variância dos coeficientes 16](#_Toc63797296)

[Teste “Augmented Dickey-Fuller” 17](#_Toc63797297)

[Analise de Resíduos 17](#_Toc63797298)

[Teste de Ljung 20](#_Toc63797299)

[Série V4 21](#_Toc63797300)

[Modelo 4 21](#_Toc63797301)

[Variância dos coeficientes 21](#_Toc63797302)

[Teste “Augmented Dickey-Fuller” 21](#_Toc63797303)

[Analise de Resíduos 22](#_Toc63797304)

[Teste de Ljung 24](#_Toc63797305)

[Série V5 26](#_Toc63797306)

[Modelo 5 26](#_Toc63797307)

[Variância dos coeficientes 26](#_Toc63797308)

[Teste “Augmented Dickey-Fuller” 26](#_Toc63797309)

[Analise de Resíduos 27](#_Toc63797310)

[Teste de Ljung 29](#_Toc63797311)

# Questão 1

|  |  |
| --- | --- |
| a | é uma série MA(1)  toda série MA(q) é estacionária  então esta série é estacionária |
| b | É uma série AR(1)  então esta série não é estacionária |
| c | então esta série é estacionária |

# Questão 5

|  |
| --- |
|  |

# Questão 7

|  |
| --- |
|  |
|  |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Questão 28 | | | | | |
| 2 | 8,8 | 18 | 158,4 | 324 |  |
| 3 | 1 | 8,8 | 8,8 | 77,44 |
| 4 | 11,2 | 1 | 11,2 | 1 |
| 5 | 30,7 | 11,2 | 343,84 | 125,44 |
| 6 | 31,9 | 30,7 | 979,33 | 942,49 |
| 7 | 22,4 | 31,9 | 714,56 | 1017,61 |  |
| 8 | 28 | 22,4 | 627,2 | 501,76 |
| 9 | 19,2 | 28 | 537,6 | 784 |
| 10 | 28,2 | 19,2 | 541,44 | 368,64 |
|  | Total | | 3922,37 | 4142,38 |

# Questão 40

## Leitura dos dados

require(forecast)

## Loading required package: forecast

## Warning: package 'forecast' was built under R version 4.0.3

## Registered S3 method overwritten by 'quantmod':  
## method from  
## as.zoo.data.frame zoo

require(urca)

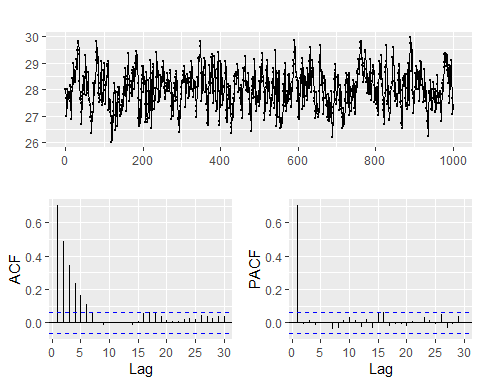
## Loading required package: urca

## Warning: package 'urca' was built under R version 4.0.3

dados <- read.csv("series.csv",header = T , sep = ",")

## Série V1

ggtsdisplay(dados$V1)



No gráfico 1 pode-se observar que o gráfico não apresenta ter nem tendência e nem sazonalidade.

No gráfico ACF pode-se observar que a covariância entre os valores da série reduz significativamente até o lag 7 e depois se mantem dentro do intervalo de confiança.

No gráfico PACF pode-se observar que não existem covariâncias muito significativas nos lags após o 0.

ndiffs(dados$V1)

## [1] 0

### **Modelo 1**

modelo1 <- auto.arima(dados$V1)  
  
modelo1

## Series: dados$V1   
## ARIMA(1,0,0) with non-zero mean   
##   
## Coefficients:  
## ar1 mean  
## 0.7023 28.0323  
## s.e. 0.0225 0.0534  
##   
## sigma^2 estimated as 0.2541: log likelihood=-733.27  
## AIC=1472.54 AICc=1472.56 BIC=1487.26

### **Variância dos coeficientes**

modelo1$var.coef

## ar1 intercept  
## ar1 5.043344e-04 -2.083487e-06  
## intercept -2.083487e-06 2.848438e-03

### **P-valor dos coeficientes**

(1-pnorm(abs(modelo1$coef)/sqrt(diag(modelo1$var.coef))))^2

## ar1 intercept   
## 0 0

Os valores do p-valor mostram que ambos os coeficientes são significativos, assim o modelo pode ser realmente considerado um modelo ARIMA.

### **Teste “**Augmented Dickey-Fuller**”**

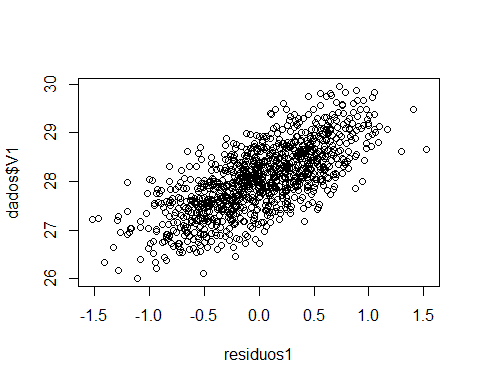
summary(ur.df(dados$V1, type='none', lags=0))

##   
## ###############################################   
## # Augmented Dickey-Fuller Test Unit Root Test #   
## ###############################################   
##   
## Test regression none   
##   
##   
## Call:  
## lm(formula = z.diff ~ z.lag.1 - 1)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.82049 -0.36262 0.00269 0.41221 1.91064   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)  
## z.lag.1 -0.0002103 0.0006162 -0.341 0.733  
##   
## Residual standard error: 0.5462 on 998 degrees of freedom  
## Multiple R-squared: 0.0001167, Adjusted R-squared: -0.0008852   
## F-statistic: 0.1165 on 1 and 998 DF, p-value: 0.7329  
##   
##   
## Value of test-statistic is: -0.3413   
##   
## Critical values for test statistics:   
## 1pct 5pct 10pct  
## tau1 -2.58 -1.95 -1.62

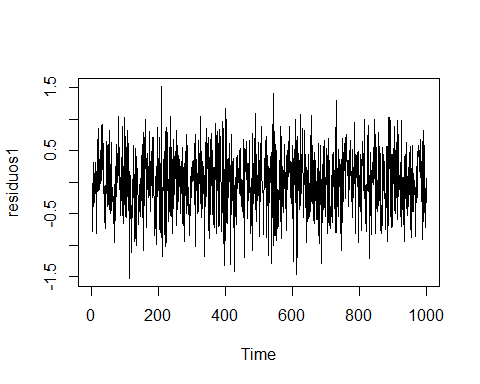
Os p-valor mostra que não existem evidencias para rejeitar que as raízes são unitárias então pode-se afirmar que a série não é estacionária a um nível alfa de 0,05.

### Analise de Resíduos

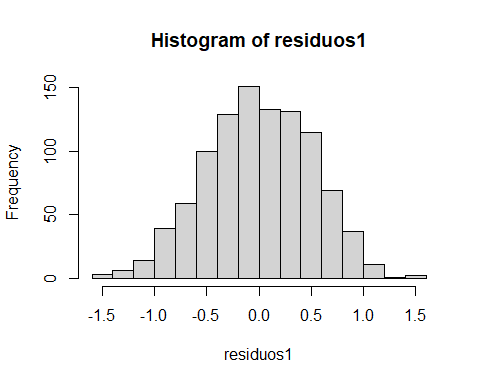
residuos1 <- modelo1$residuals  
plot(residuos1,dados$V1)



plot(residuos1)



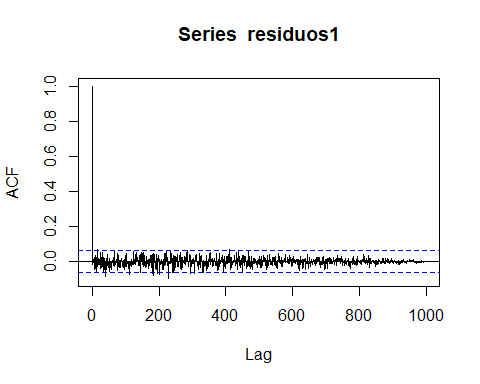
hist(residuos1)



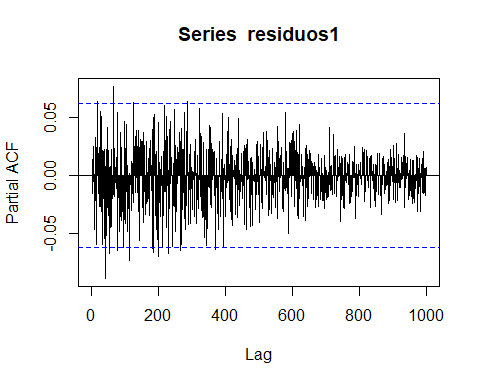
shapiro.test(residuos1)

##   
## Shapiro-Wilk normality test  
##   
## data: residuos1  
## W = 0.9966, p-value = 0.02923

acf(residuos1, lag.max = length(residuos1))



pacf(residuos1, lag.max = length(residuos1))



### Teste de Ljung

Box.test(residuos1,lag=5,type = c("Box-Pierce","Ljung-Box"),fitdf=0)

##   
## Box-Pierce test  
##   
## data: residuos1  
## X-squared = 0.53778, df = 5, p-value = 0.9907

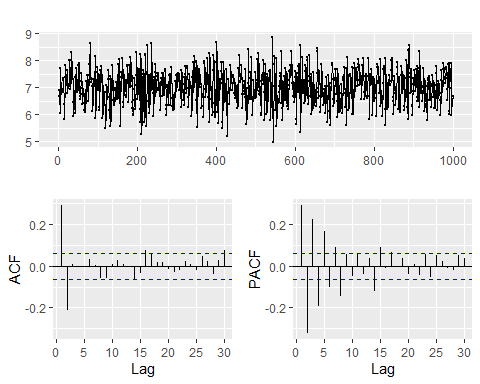
Box.test(residuos1,lag=10,type = c("Box-Pierce","Ljung-Box"),fitdf=0)

##   
## Box-Pierce test  
##   
## data: residuos1  
## X-squared = 4.9728, df = 10, p-value = 0.893

Em ambos os testes a um nível alfa de 0,05 não existem evidencias para rejeitar que as correlações são diferentes de 0

## Série V2

ggtsdisplay(dados$V2)



No gráfico 1 pode-se observar que o gráfico não apresenta ter nem tendência e nem sazonalidade.

No gráfico ACF pode-se observar que a covariância entre os valores da série reduz significativamente até o lag 3 e depois se mantem dentro do intervalo de confiança.

No gráfico PACF pode-se observar que não existem covariâncias muito significativas nos lags após o 6.

ndiffs(dados$V2)

## [1] 0

### Modelo 2

Modelo2 <- auto.arima(dados$V2)  
  
modelo2

## Series: dados$V2   
## ARIMA(0,0,2) with non-zero mean   
##   
## Coefficients:  
## ma1 ma2 mean  
## 0.6084 -0.2967 7.0124  
## s.e. 0.0303 0.0299 0.0209  
##   
## sigma^2 estimated as 0.2543: log likelihood=-733.59  
## AIC=1475.18 AICc=1475.22 BIC=1494.81

### Variância dos coeficientes

Modelo2$var.coef

## ma1 ma2 intercept  
## ma1 9.156576e-04 8.024666e-04 -7.541932e-07  
## ma2 8.024666e-04 8.914425e-04 -8.883280e-07  
## intercept -7.541932e-07 -8.883280e-07 4.362082e-04

### **P-valor dos coeficientes**

(1-pnorm(abs(modelo2$coef)/sqrt(diag(modelo2$var.coef))))^2

## ma1 ma2 intercept   
## 0 0 0

Os valores do p-valor mostram que ambos os coeficientes são significativos, assim o modelo pode ser realmente considerado um modelo ARIMA.

### **Teste “**Augmented Dickey-Fuller**”**

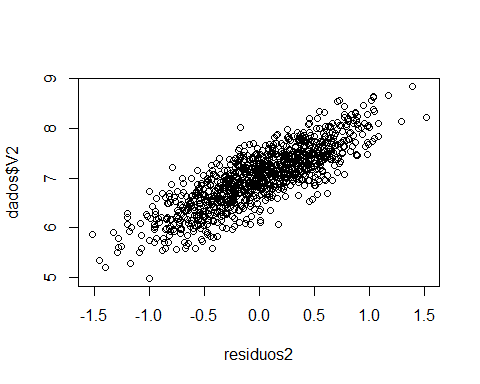
summary(ur.df(dados$V2, type='none', lags=0))

##   
## ###############################################   
## # Augmented Dickey-Fuller Test Unit Root Test #   
## ###############################################   
##   
## Test regression none   
##   
##   
## Call:  
## lm(formula = z.diff ~ z.lag.1 - 1)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -2.69727 -0.42457 0.06182 0.52496 2.57107   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)  
## z.lag.1 -0.00538 0.00327 -1.645 0.1  
##   
## Residual standard error: 0.7274 on 998 degrees of freedom  
## Multiple R-squared: 0.002705, Adjusted R-squared: 0.001706   
## F-statistic: 2.707 on 1 and 998 DF, p-value: 0.1002  
##   
##   
## Value of test-statistic is: -1.6454   
##   
## Critical values for test statistics:   
## 1pct 5pct 10pct  
## tau1 -2.58 -1.95 -1.62

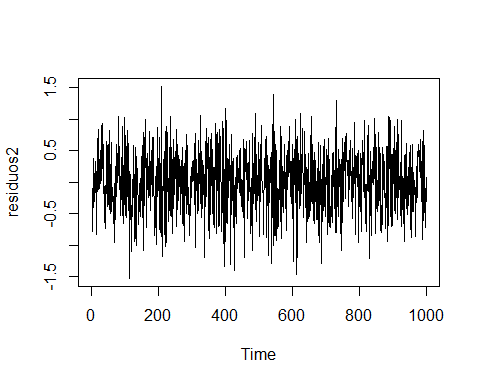
Os p-valor mostra que não existem evidencias para rejeitar que as raízes são unitárias então pode-se afirmar que a série não é estacionária a um nível alfa de 0,05.

### Analise de Resíduos

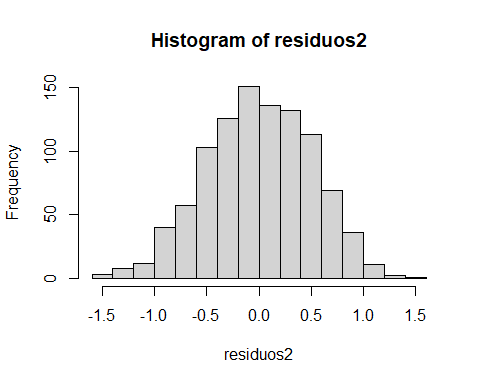
residuos2 <- modelo1$residuals  
plot(residuos2,dados$V2)



plot(residuos2)



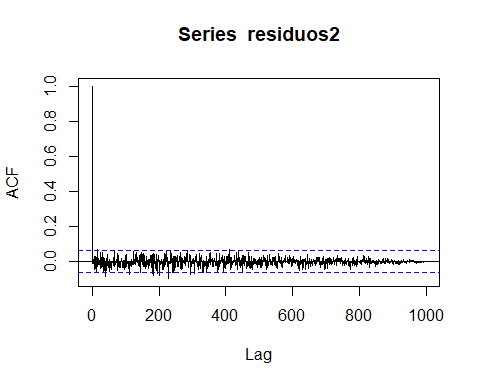
hist(residuos2)



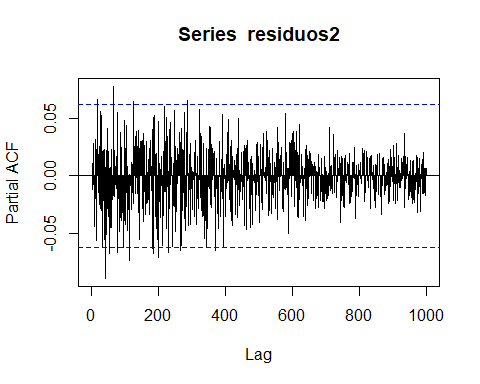
shapiro.test(residuos2)

##   
## Shapiro-Wilk normality test  
##   
## data: residuos2  
## W = 0.99659, p-value = 0.02875

acf(residuos2, lag.max = length(residuos2))



pacf(residuos2, lag.max = length(residuos2))



### Teste de Ljung

Box.test(residuos2,lag=5,type = c("Box-Pierce","Ljung-Box"),fitdf=0)

##   
## Box-Pierce test  
##   
## data: residuos2  
## X-squared = 0.33355, df = 5, p-value = 0.997

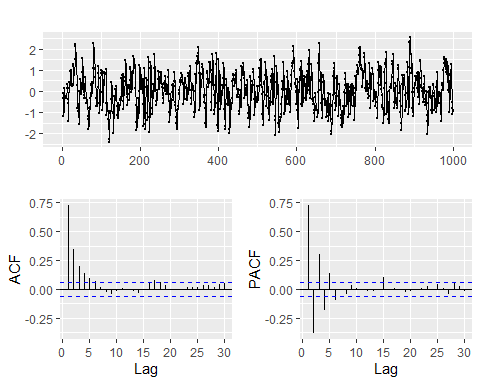
Box.test(residuos2,lag=10,type = c("Box-Pierce","Ljung-Box"),fitdf=0)

##   
## Box-Pierce test  
##   
## data: residuos2  
## X-squared = 4.7823, df = 10, p-value = 0.9052

Em ambos os testes a um nível alfa de 0,05 não existem evidencias para rejeitar que as correlações são diferentes de 0

## Série V3

ggtsdisplay(dados$V3)



No gráfico 1 pode-se observar que o gráfico não apresenta ter nem tendência e nem sazonalidade.

No gráfico ACF pode-se observar que a covariância entre os valores da série reduz significativamente até o lag 6 e depois se mantem dentro do intervalo de confiança.

No gráfico PACF pode-se observar que não existem covariâncias muito significativas nos lags após o 6

ndiffs(dados$V3)

## [1] 0

### Modelo 3

modelo3 <- auto.arima(dados$V3)  
  
modelo3

## Series: dados$V3   
## ARIMA(1,0,1) with zero mean   
##   
## Coefficients:  
## ar1 ma1  
## 0.4745 0.7305  
## s.e. 0.0308 0.0229  
##   
## sigma^2 estimated as 0.2555: log likelihood=-736.44  
## AIC=1478.87 AICc=1478.9 BIC=1493.59

### **Variância dos coeficientes**

modelo3$var.coef

## ar1 ma1  
## ar1 0.000946005 -0.0003022610  
## ma1 -0.000302261 0.0005266432

### **P-valor dos coeficientes**

(1-pnorm(abs(modelo3$coef)/sqrt(diag(modelo3$var.coef))))^2

## ar1 ma1   
## 0 0

Os valores do p-valor mostram que ambos os coeficientes são significativos, assim o modelo pode ser realmente considerado um modelo ARIMA.

### Teste “Augmented Dickey-Fuller”

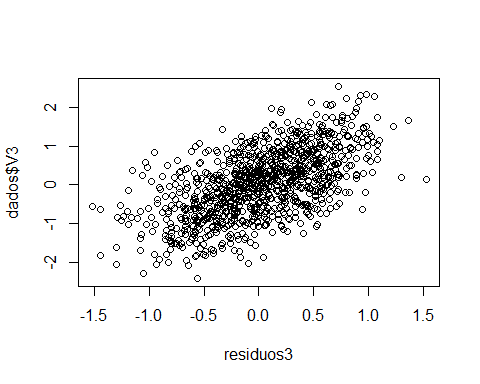
summary(ur.df(dados$V3, type='none', lags=0))

##   
## ###############################################   
## # Augmented Dickey-Fuller Test Unit Root Test #   
## ###############################################   
##   
## Test regression none   
##   
##   
## Call:  
## lm(formula = z.diff ~ z.lag.1 - 1)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.98626 -0.38050 0.03792 0.42152 1.78932   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## z.lag.1 -0.27319 0.02177 -12.55 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.5893 on 998 degrees of freedom  
## Multiple R-squared: 0.1363, Adjusted R-squared: 0.1354   
## F-statistic: 157.5 on 1 and 998 DF, p-value: < 2.2e-16  
##   
##   
## Value of test-statistic is: -12.5502   
##   
## Critical values for test statistics:   
## 1pct 5pct 10pct  
## tau1 -2.58 -1.95 -1.62

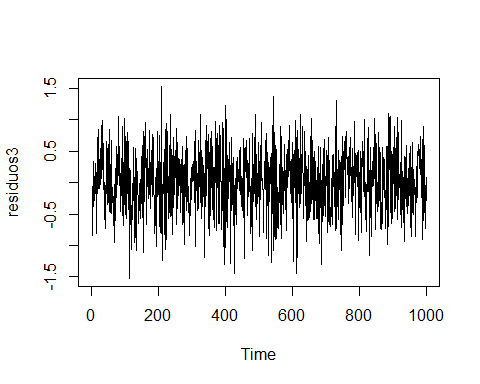
Os p-valor mostra que não existem evidencias para rejeitar que as raízes são unitárias então pode-se afirmar que a série é estacionária a um nível alfa de 0,05.

### Analise de Resíduos

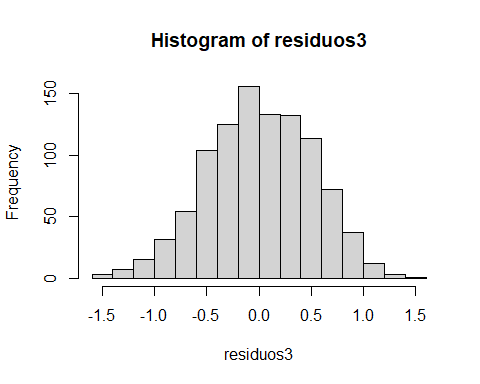
residuos3 <- modelo3$residuals  
plot(residuos3,dados$V3)



plot(residuos3)



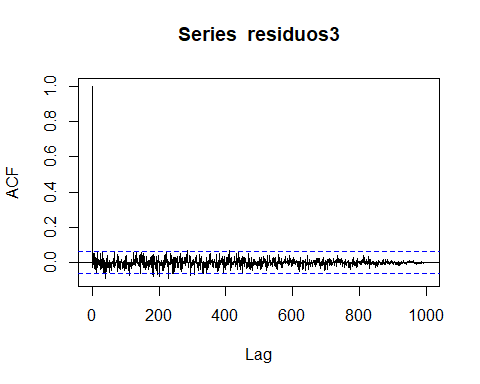
hist(residuos3)



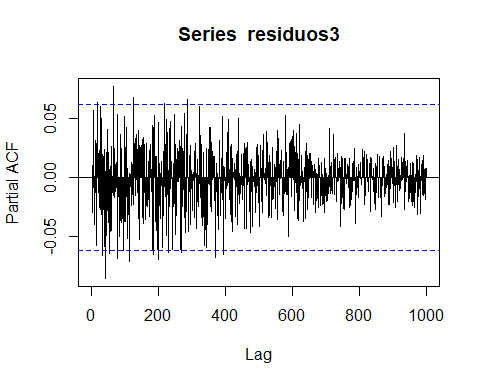
shapiro.test(residuos3)

##   
## Shapiro-Wilk normality test  
##   
## data: residuos3  
## W = 0.99688, p-value = 0.04651

acf(residuos3, lag.max = length(residuos3))



pacf(residuos3, lag.max = length(residuos3))



### Teste de Ljung

Box.test(residuos3,lag=5,type = c("Box-Pierce","Ljung-Box"),fitdf=0)

##   
## Box-Pierce test  
##   
## data: residuos3  
## X-squared = 3.9653, df = 5, p-value = 0.5544

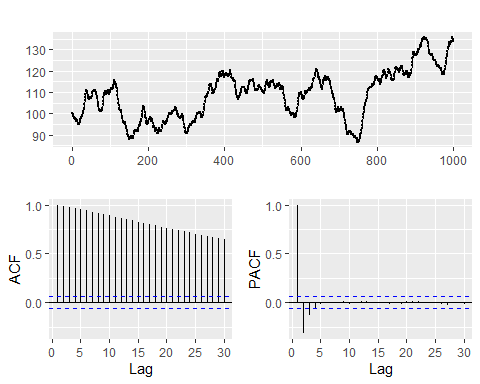
Box.test(residuos3,lag=10,type = c("Box-Pierce","Ljung-Box"),fitdf=0)

##   
## Box-Pierce test  
##   
## data: residuos3  
## X-squared = 9.6287, df = 10, p-value = 0.4737

Em ambos os testes a um nível alfa de 0,05 não existem evidencias para rejeitar que as correlações são diferentes de 0

## Série V4

ggtsdisplay(dados$V4)



No gráfico 1 pode-se observar que o gráfico não apresenta ter nem tendência e nem sazonalidade.

No gráfico ACF pode-se observar que a covariância entre os valores da série não se reduz significativamente até o lag.

No gráfico PACF pode-se observar que não existem covariâncias muito significativas nos lags após o 3.

ndiffs(dados$V4)

## [1] 1

### Modelo 4

modelo4 <- auto.arima(dados$V4)  
  
modelo4

## Series: dados$V4   
## ARIMA(1,1,0)   
##   
## Coefficients:  
## ar1  
## 0.7031  
## s.e. 0.0225  
##   
## sigma^2 estimated as 0.2542: log likelihood=-733.2  
## AIC=1470.41 AICc=1470.42 BIC=1480.22

### **Variância dos coeficientes**

modelo4$var.coef

## ar1  
## ar1 0.0005047389

### **P-valor dos coeficientes**

(1-pnorm(abs(modelo4$coef)/sqrt(diag(modelo4$var.coef))))^2

## ar1   
## 0

Os valores do p-valor mostram que ambos os coeficientes são significativos, assim o modelo pode ser realmente considerado um modelo ARIMA.

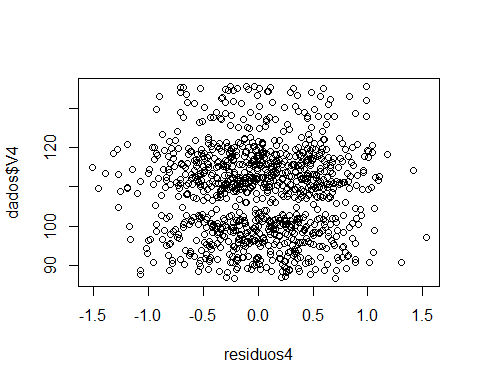
### Teste “Augmented Dickey-Fuller”

summary(ur.df(dados$V4, type='none', lags=0))

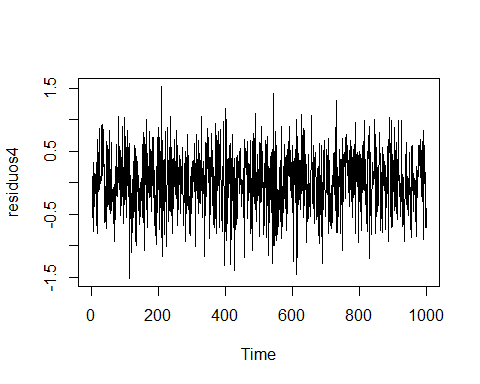
##   
## ###############################################   
## # Augmented Dickey-Fuller Test Unit Root Test #   
## ###############################################   
##   
## Test regression none   
##   
##   
## Call:  
## lm(formula = z.diff ~ z.lag.1 - 1)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -2.02330 -0.48204 -0.00207 0.48374 1.91514   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)  
## z.lag.1 0.0003124 0.0002052 1.522 0.128  
##   
## Residual standard error: 0.7083 on 998 degrees of freedom  
## Multiple R-squared: 0.002316, Adjusted R-squared: 0.001316   
## F-statistic: 2.317 on 1 and 998 DF, p-value: 0.1283  
##   
##   
## Value of test-statistic is: 1.5221   
##   
## Critical values for test statistics:   
## 1pct 5pct 10pct  
## tau1 -2.58 -1.95 -1.62

### Analise de Resíduos

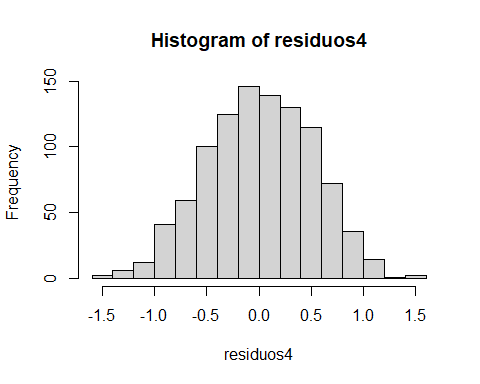
residuos4 <- modelo4$residuals  
plot(residuos4,dados$V4)



plot(residuos4)



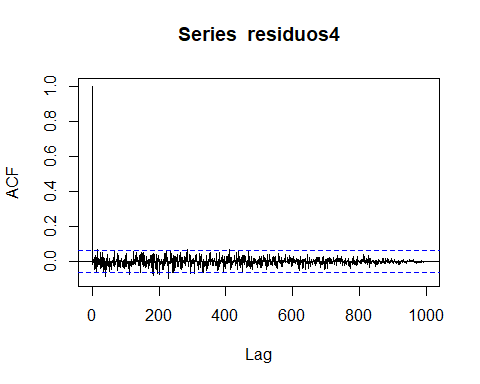
hist(residuos4)



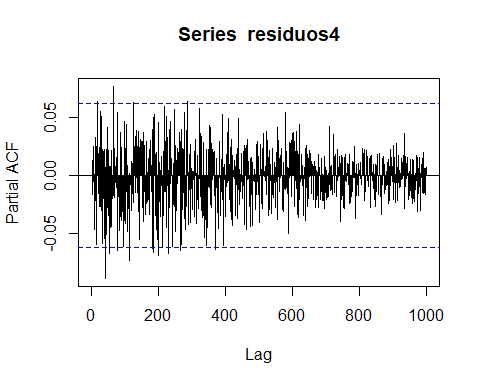
shapiro.test(residuos4)

##   
## Shapiro-Wilk normality test  
##   
## data: residuos4  
## W = 0.99659, p-value = 0.02918

acf(residuos4, lag.max = length(residuos4))



pacf(residuos4, lag.max = length(residuos4))



### Teste de Ljung

Box.test(residuos4,lag=5,type = c("Box-Pierce","Ljung-Box"),fitdf=0)

##   
## Box-Pierce test  
##   
## data: residuos4  
## X-squared = 0.54032, df = 5, p-value = 0.9906

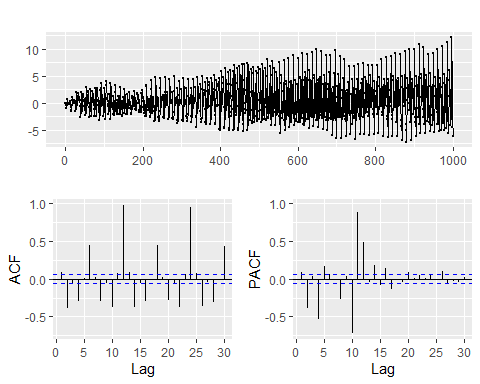
Box.test(residuos4,lag=10,type = c("Box-Pierce","Ljung-Box"),fitdf=0)

##   
## Box-Pierce test  
##   
## data: residuos4  
## X-squared = 4.9865, df = 10, p-value = 0.8921

Em ambos os testes a um nível alfa de 0,05 não existem evidencias para rejeitar que as correlações são diferentes de 0

## Série V5

ggtsdisplay(dados$V5)



No gráfico 1 pode-se observar que o gráfico não apresenta ter nem tendência e nem sazonalidade, as vai aumentando a amplitude com o tempo.

No gráfico ACF pode-se observar que a covariância entre os valores da série não se reduz significativamente e apresenta alguns picos.

No gráfico PACF pode-se observar que não existem covariâncias muito significativas e apresenta alguns picos.

ndiffs(dados$V5)

## [1] 1

### Modelo 5

modelo5 <- auto.arima(dados$V5)  
  
modelo5

## Series: dados$V5   
## ARIMA(2,1,0)   
##   
## Coefficients:  
## ar1 ar2  
## -0.3674 -0.5263  
## s.e. 0.0269 0.0269  
##   
## sigma^2 estimated as 13.28: log likelihood=-2708.84  
## AIC=5423.67 AICc=5423.7 BIC=5438.39

### Variância dos coeficientes

modelo5$var.coef

## ar1 ar2  
## ar1 0.0007230824 0.0001736694  
## ar2 0.0001736694 0.0007213632

### **P-valor dos coeficientes**

(1-pnorm(abs(modelo5$coef)/sqrt(diag(modelo5$var.coef))))^2

## ar1 ar2   
## 0 0

Os valores do p-valor mostram que ambos os coeficientes são significativos, assim o modelo pode ser realmente considerado um modelo ARIMA.

### **Teste “**Augmented Dickey-Fuller**”**

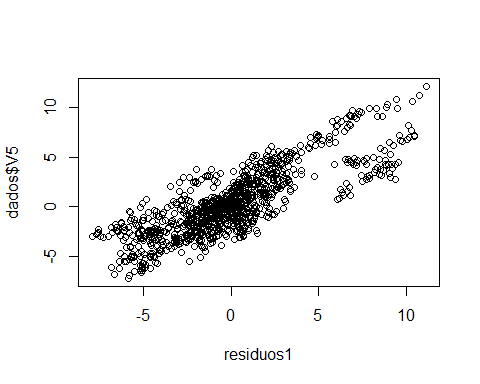
summary(ur.df(dados$V5, type='none', lags=0))

##   
## ###############################################   
## # Augmented Dickey-Fuller Test Unit Root Test #   
## ###############################################   
##   
## Test regression none   
##   
##   
## Call:  
## lm(formula = z.diff ~ z.lag.1 - 1)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -6.8023 -2.0745 -0.1263 2.1154 12.2362   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## z.lag.1 -0.90374 0.03156 -28.63 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 3.273 on 998 degrees of freedom  
## Multiple R-squared: 0.451, Adjusted R-squared: 0.4505   
## F-statistic: 819.9 on 1 and 998 DF, p-value: < 2.2e-16  
##   
##   
## Value of test-statistic is: -28.6339   
##   
## Critical values for test statistics:   
## 1pct 5pct 10pct  
## tau1 -2.58 -1.95 -1.62

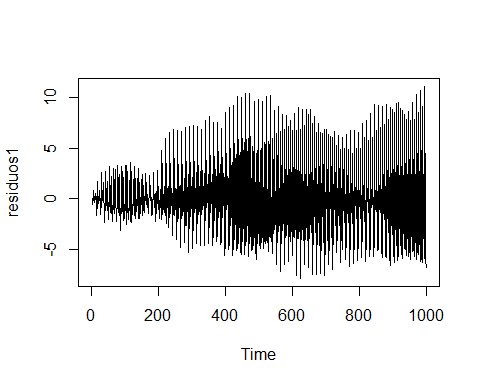
Os p-valor mostra que não existem evidencias para rejeitar que as raízes são unitárias então pode-se afirmar que a série é estacionária a um nível alfa de 0,05.

### Analise de Resíduos

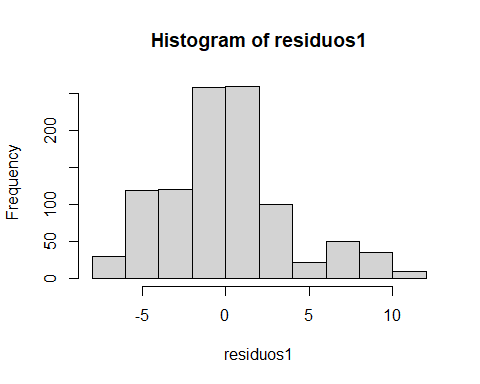
residuos5 <- modelo5$residuals  
plot(residuos5,dados$V5)



plot(residuos5)



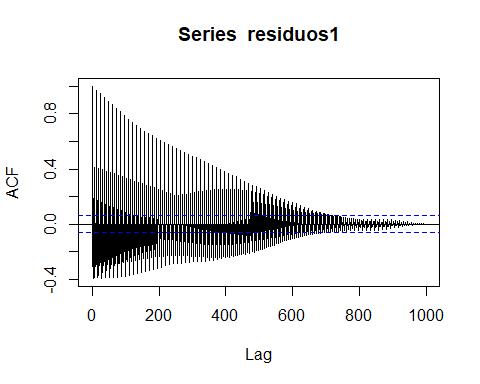
hist(residuos5)



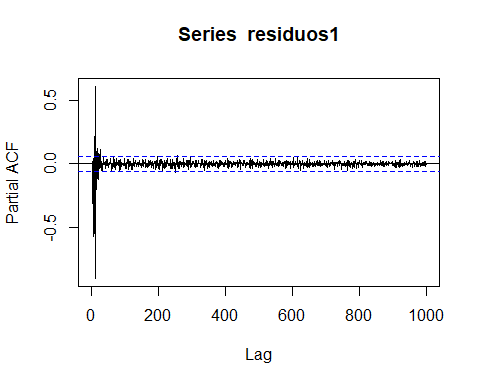
shapiro.test(residuos5)

##   
## Shapiro-Wilk normality test  
##   
## data: residuos1  
## W = 0.96021, p-value = 7.269e-16

acf(residuos1, lag.max = length(residuos1))



pacf(residuos5, lag.max = length(residuos1))



### Teste de Ljung

Box.test(residuos5,lag=5,type = c("Box-Pierce","Ljung-Box"),fitdf=0)

##   
## Box-Pierce test  
##   
## data: residuos1  
## X-squared = 325.11, df = 5, p-value < 2.2e-16

Box.test(residuos5,lag=10,type = c("Box-Pierce","Ljung-Box"),fitdf=0)

##   
## Box-Pierce test  
##   
## data: residuos1  
## X-squared = 814.31, df = 10, p-value < 2.2e-16

Em ambos os testes a um nível alfa de 0,05 existem evidencias para rejeitar que as correlações são diferentes de 0