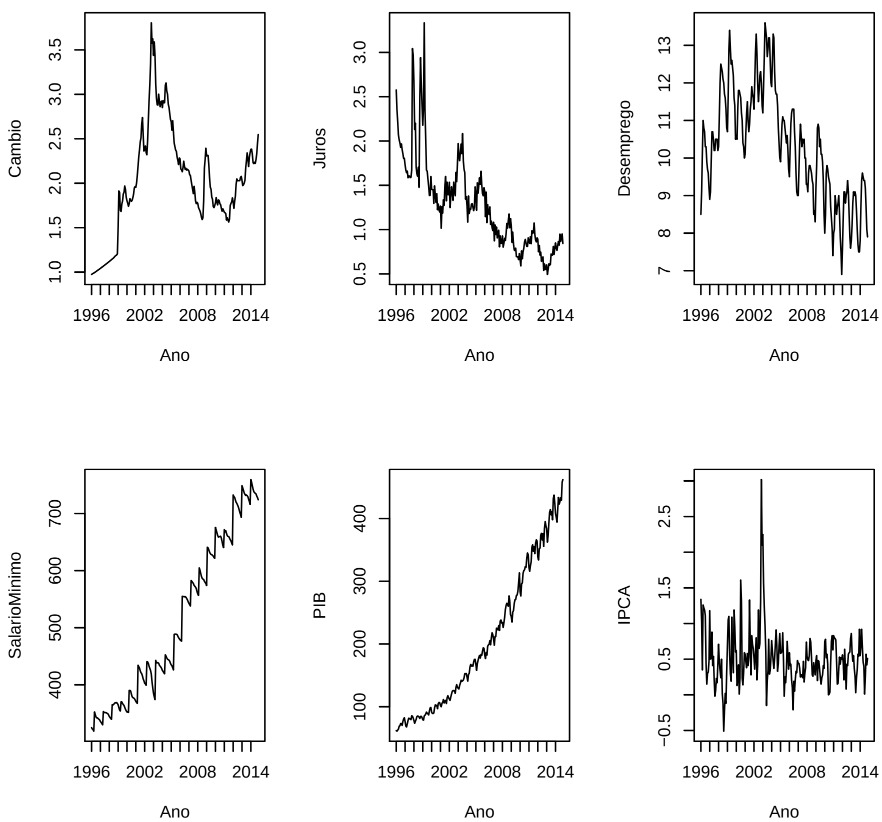
Monitoria Econometria Avançada
Lista 2

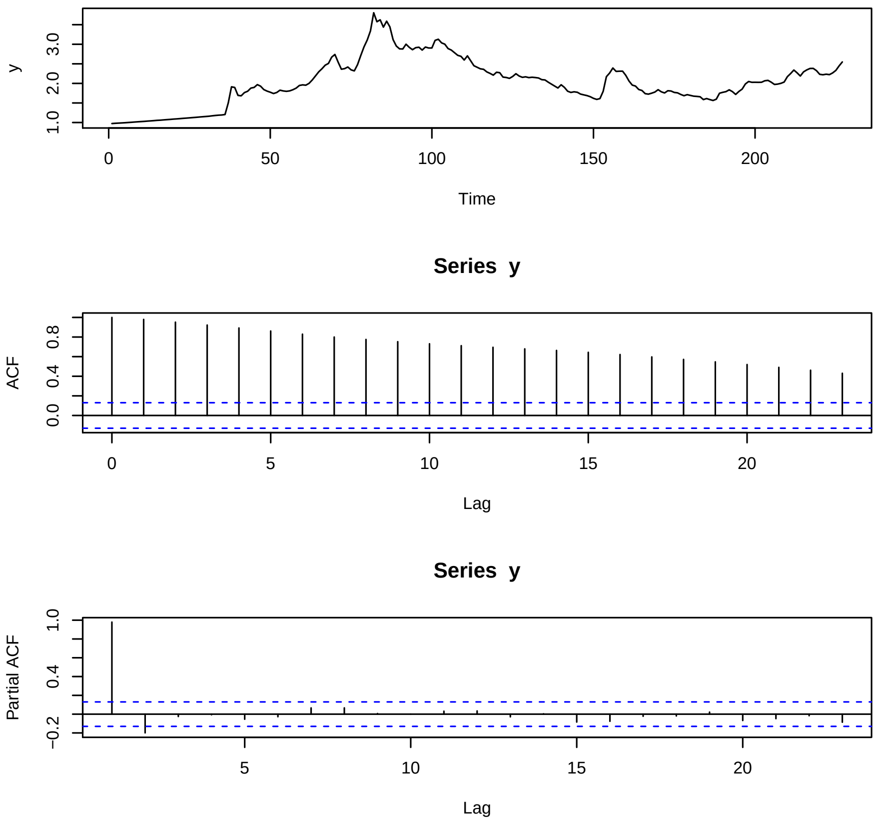
Professor: Hedibert Lopes

Primeiramente, vamos plotar os gráficos das séries, rodando a primeira parte do programa macro-arima.R disponível no site do professor, observando o caminho de rede onde está salvo o arquivo macro.csv.



Agora vamos analisar a série do Câmbio, rodando a seguinte programação, de forma a identificar um modelo ARIMA adequado:

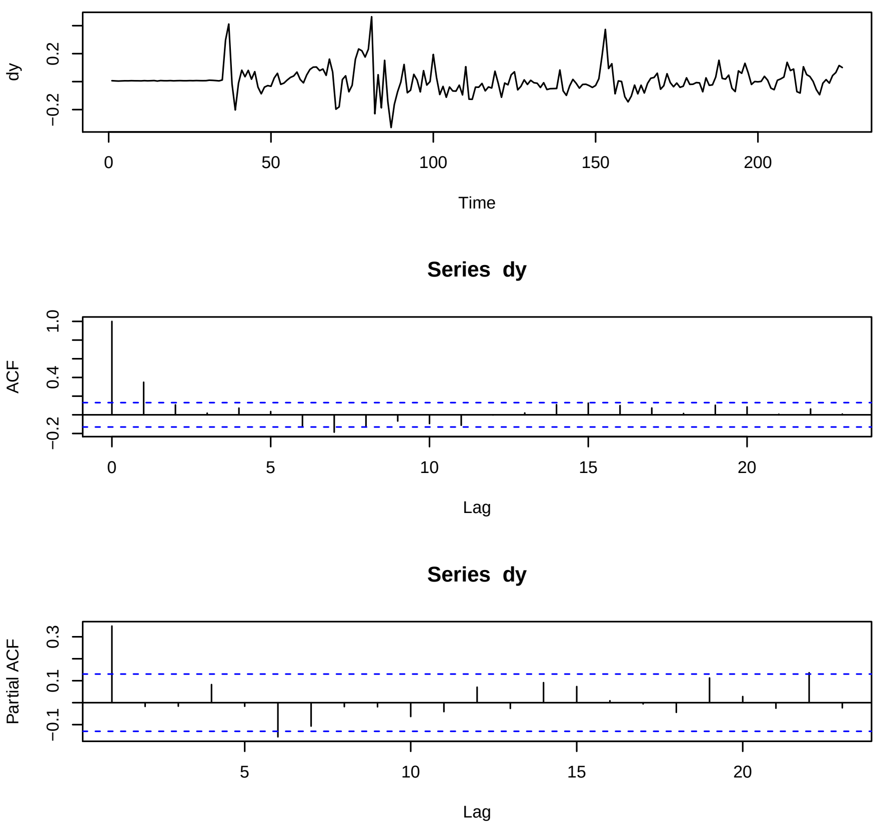
```
> y  = Cambio
> n  = length(y)
> # Grafico da serie temporal, sua funcao de aucorrelacao (FAC) e
> # sua funcao de autocorrelacao parcial (FACP)
> par(mfrow=c(3,1))
> ts.plot(y)
> acf(y)
> pacf(y)
```



Pelo gráfico da série, vimos que esta possui tendência e analisando o gráfico da ACF, nota-se que a série não é estacionária. Portanto, vamos proceder analisando a primeira diferença da série de Câmbio:

```
> dy  = diff(y,1)
> # Grafico da serie temporal, sua funcao de aucorrelacao (FAC) e
> # sua funcao de autocorrelacao parcial (FACP)
```

```
> par(mfrow=c(3,1))
> ts.plot(dy)
> acf(dy)
> pacf(dy)
```

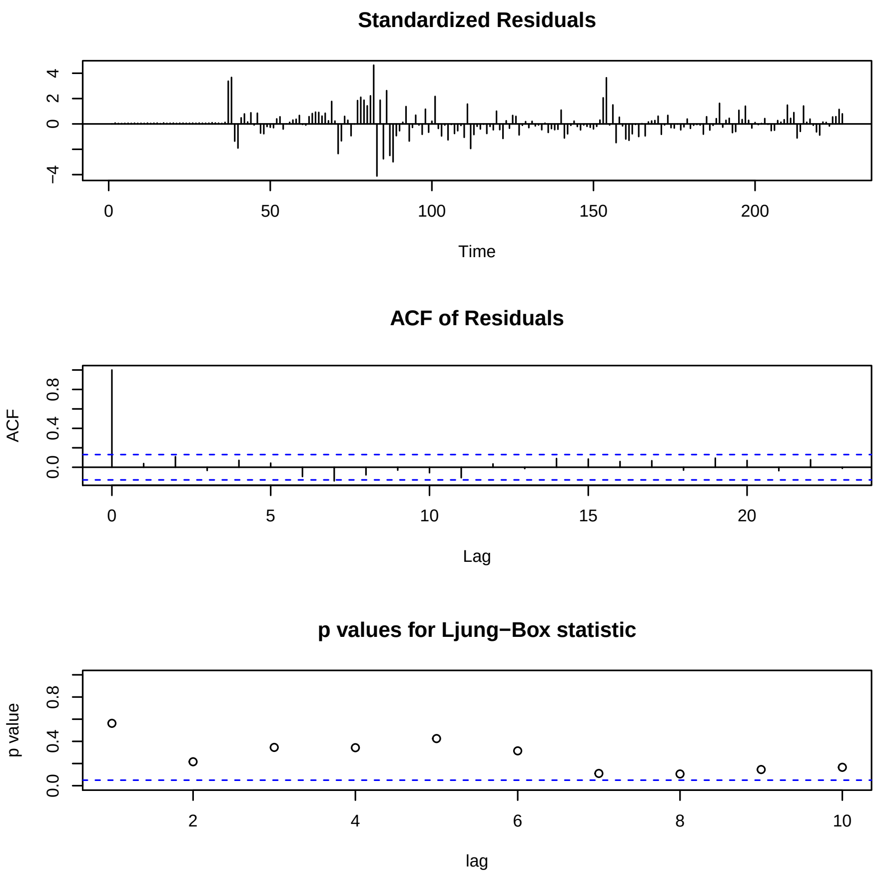


Nesse caso, vemos que a série torna-se estacionária e que a ACF encontra-se fora do intervalo de confiança (ou seja, é estatisticamente diferente de zero) somente até o lag 1. Já a PACF está praticamente dentro do intervalo de confiança, exceto talvez pelo lag 6. Sendo assim, há indícios de que trata-se de um processo ARIMA(0,1,1) ou de um processo ARIMA(1,1,0). Vamos ajustar o modelo IMA(1,1), primeiramente:

```
> p=0
> d=1
> q=1
> k=1+p+q+1
> fit = arima(y,order=c(p,d,q),method="ML")
```

```
> tsdiag(fit)
> fit$aic
```

```
[1] -458.7622
```



O modelo parece bem ajustado, uma vez que a ACF dos resíduos está dentro do intervalo de confiança (ou seja, os resíduos são ruído branco). O p-valor do teste de Ljung-Box (hipótese nula é de não autocorrelação dos resíduos) são relativamente altos para todos os lags, logo não rejeitamos a hipótese.

Contudo, é interessante rodarmos uma combinação de modelos ARIMA e analisar os critérios de informação (quanto menores, melhor):

```
> model = NULL
> aic=NULL
> for (p in 0:2)
+ for (q in 0:2)
```

```
+ for (d in 0:2){
+   model = rbind(model,c(p,d,q))
+   aic=c(aic,arima(y,order=c(p,d,q),method="ML")$aic)
+  }
> # Modelos e respectivos AIC
> cbind(model,aic)
```

```
                   aic
 [1,] 0 0 0  415.55299
 [2,] 0 1 0 -433.92228
 [3,] 0 2 0 -373.68561
 [4,] 0 0 1  141.98935
 [5,] 0 1 1 -458.76218
 [6,] 0 2 1 -426.38694
 [7,] 0 0 2  -58.47124
 [8,] 0 1 2 -460.49335
 [9,] 0 2 2 -449.70331
[10,] 1 0 0 -428.71866
[11,] 1 1 0 -461.91111
[12,] 1 2 0 -395.18974
[13,] 1 0 1 -454.20986
[14,] 1 1 1 -459.94767
[15,] 1 2 1 -452.95598
[16,] 1 0 2 -453.76406
[17,] 1 1 2 -458.69006
[18,] 1 2 2 -450.98250
[19,] 2 0 0 -458.05160
[20,] 2 1 0 -459.95110
[21,] 2 2 0 -406.40305
[22,] 2 0 1 -456.05244
[23,] 2 1 1 -457.94876
[24,] 2 2 1 -450.98463
[25,] 2 0 2 -452.81532
[26,] 2 1 2 -456.74621
[27,] 2 2 2 -449.06744
```

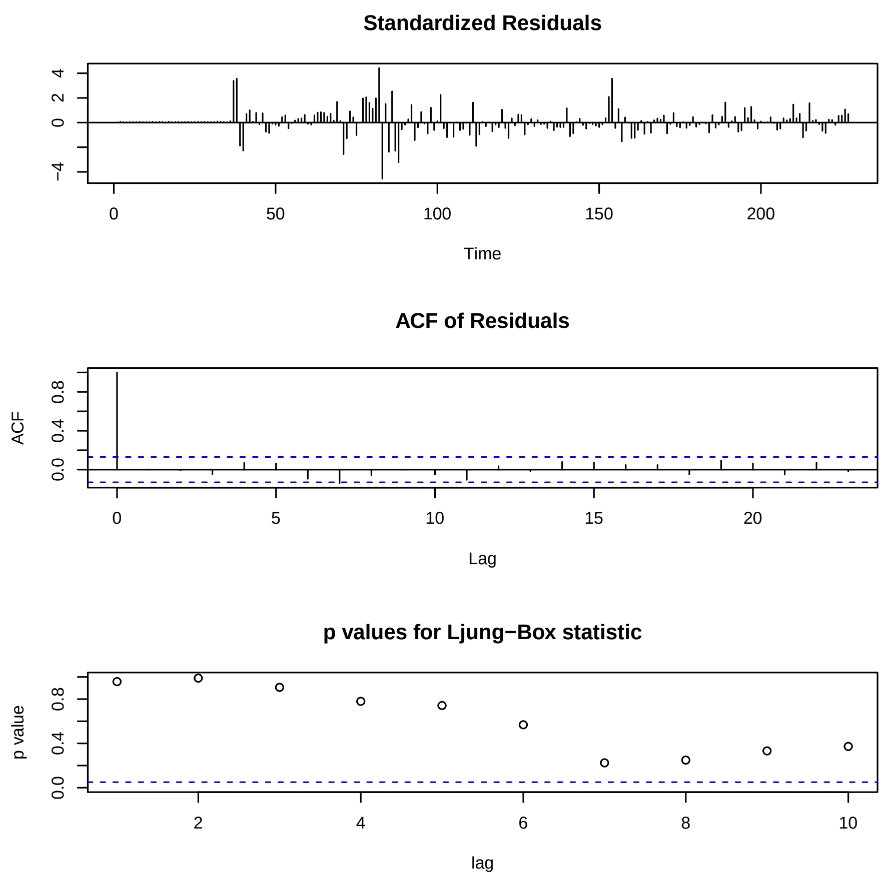
```
> # Melhor modelo (menor AIC)
> model[aic==min(aic)]
```

```
[1] 1 1 0
```

```
> p=1
> d=1
```

```
> q=0
> k=1+p+q+1
> fit = arima(y,order=c(p,d,q),method="ML")
> tsdiag(fit)
> fit$aic
```

```
[1] -461.9111
```

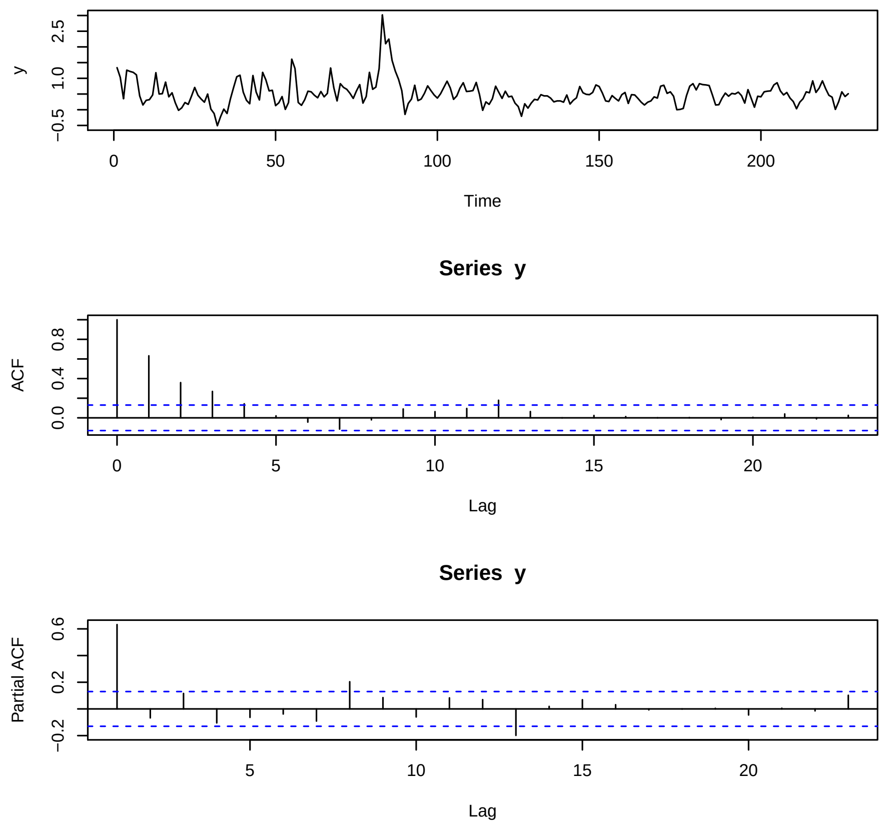


Ou seja, o modelo com menor AIC é o ARIMA(1,1,0). Analisando os resíduos deste modelo (exibido acima), verificamos que trata-se de um ruído branco e que os p-valores são altos para todos os lags. Logo, ficamos com este modelo.

Agora vamos repetir o procedimento para a série do IPCA:

```
> y  = IPCA
> n  = length(y)
> # Grafico da serie temporal, sua funcao de aucorrelacao (FAC) e
```

```
> # sua funcao de autocorrelacao parcial (FACP)
> par(mfrow=c(3,1))
> ts.plot(y)
> acf(y)
> pacf(y)
```



Pelo gráfico da série, esta aparenta ser estacionária, o que é confirmado pela ACF, que decai rapidamente para zero. Sendo assim, não é necessário tomar uma diferença. Nesse sentido, vamos ajustar diretamente um modelo ARMA. Nota-se que a ACF é diferente de zero até o lag 3, portanto, vamos ajustar modelos ARIMA(p,0,q) considerando p,q = 1,2,3. Note que a PACF é diferente de zero para os lags 8 e 13. Porém, por questões de parcimônia, vamos manter somente ordens até 3 e verificar o ajuste dos modelos:

```
> model = NULL
> aic=NULL
> for (p in 0:3)
```

```
+ for (q in 0:3)
+ for (d in 0:0){
+   model = rbind(model,c(p,d,q))
+   aic=c(aic,arima(y,order=c(p,d,q),method="ML")$aic)
+  }
> # Modelos e respectivos AIC
> cbind(model,aic)
```

```
                 aic
 [1,] 0 0 0 230.1609
 [2,] 0 0 1 135.0623
 [3,] 0 0 2 122.9111
 [4,] 0 0 3 119.8780
 [5,] 1 0 0 113.8847
 [6,] 1 0 1 114.2338
 [7,] 1 0 2 114.3138
 [8,] 1 0 3 115.5455
 [9,] 2 0 0 114.8089
[10,] 2 0 1 114.8879
[11,] 2 0 2 116.0515
[12,] 2 0 3 111.9929
[13,] 3 0 0 114.1762
[14,] 3 0 1 115.4370
[15,] 3 0 2 102.4268
[16,] 3 0 3 102.7421
```

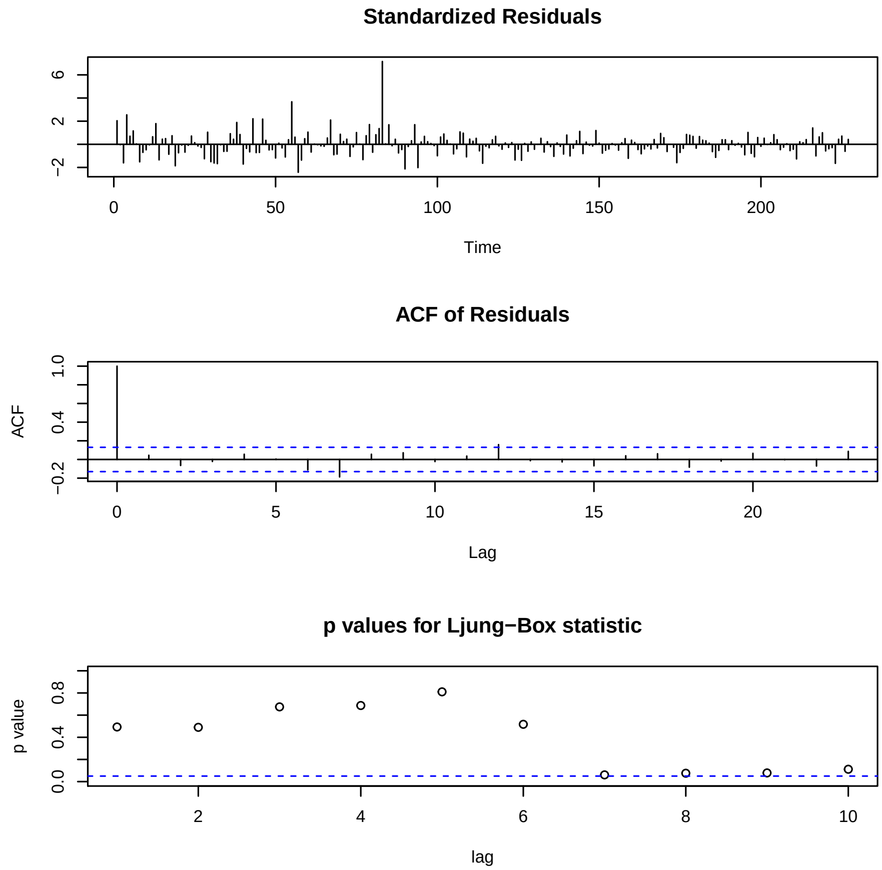
```
> # Melhor modelo (menor AIC)
> model[aic==min(aic)]
```

```
[1] 3 0 2
```

Ou seja, o modelo com menor AIC é o ARMA(3,2). Vamos analisar os resíduos deste modelo:

```
> p=3
> d=0
> q=2
> k=1+p+q+1
> fit = arima(y,order=c(p,d,q),method="ML")
> tsdiag(fit)
> fit$aic
```

```
[1] 102.4268
```

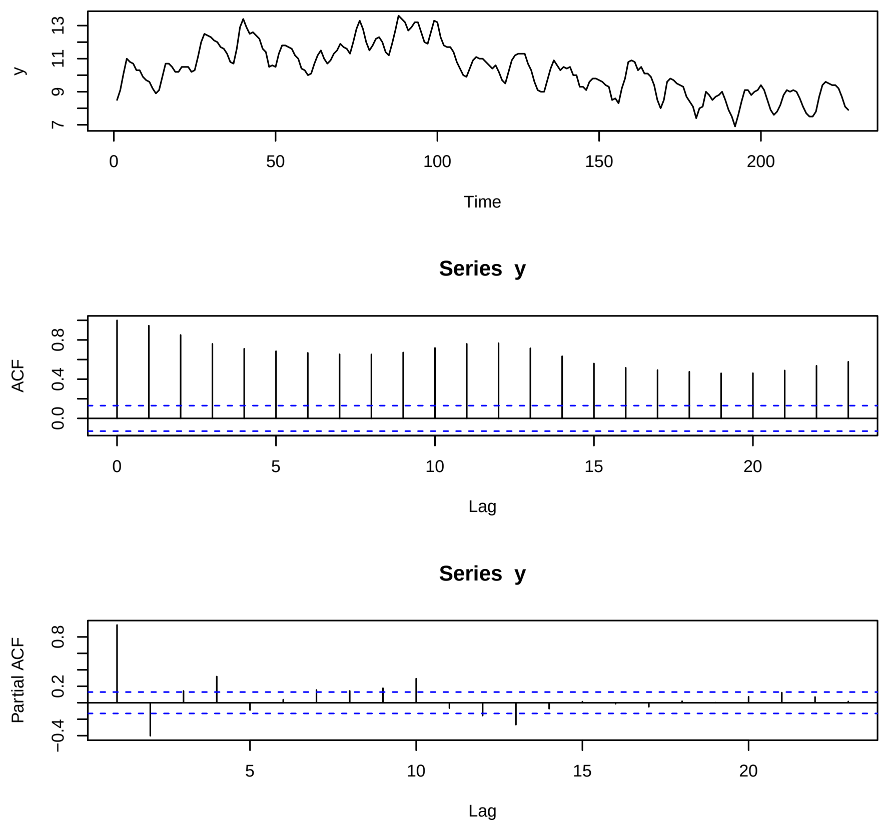


Nota-se que a ACF dos resíduos deste modelo é diferente de zero para os lags 7 e 12. Ou seja, os resíduos são autocorrelacionadas considerando altas defasagens. Note também que aparentam ser heterocedásticos. Nesse sentido, um modelo ARMA talvez não seja adequado para o propósito.

Agora vamos analisar a série de Desemprego, rodando a seguinte programação, de forma a identificar um modelo ARIMA adequado:

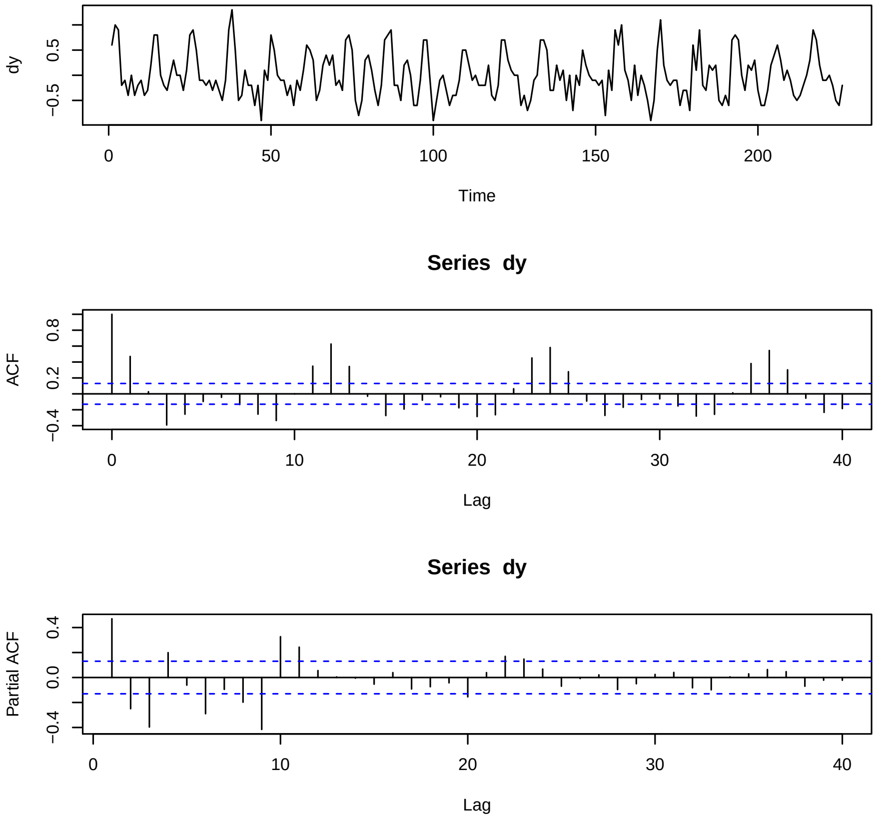
```
> y  = Desemprego
> n  = length(y)
> # Grafico da serie temporal, sua funcao de aucorrelacao (FAC) e
> # sua funcao de autocorrelacao parcial (FACP)
> par(mfrow=c(3,1))
> ts.plot(y)
```

```
> acf(y)
> pacf(y)
```



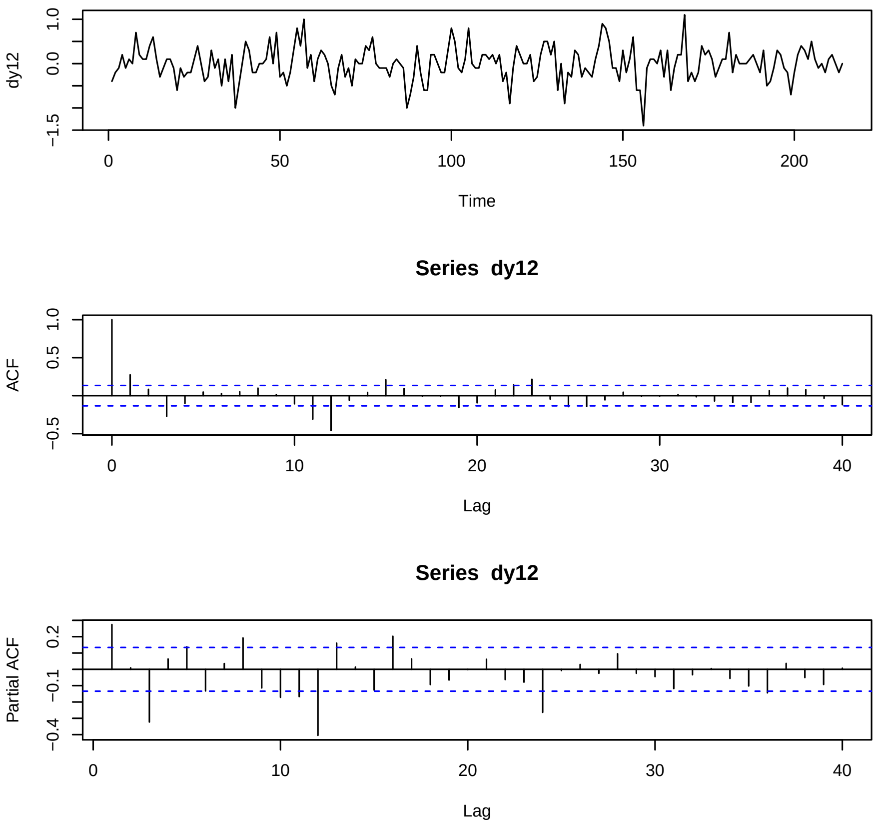
A série exibe persistência no decaimento da ACF. Portanto, vamos aplicar a primeira diferença e analisar novamente:

```
> dy=diff(y,1)
> # Grafico da serie temporal, sua funcao de aucorrelacao (FAC) e
> # sua funcao de autocorrelacao parcial (FACP)
> par(mfrow=c(3,1))
> ts.plot(dy)
> acf(dy,40)
> pacf(dy,40)
```



Nota-se que a ACF nos lags 1,3,11,12,13,24 é estatisticamente diferente de zero. Tanto pelo gráfico da série, quanto pelo comportamento da ACF (diferente de zero nos lags múltiplos de 12), podemos ver que a série exibe sazonalidade. Deveríamos, portanto, estimar um modelo ARIMA sazonal ou SARIMA. Para determinar as ordens P e Q do SARIMA(p,d,q)$(P,D,Q)_{12}$, podemos tomar a diferença de ordem 12 da primeira diferença da série original:

```
> dy12=diff(dy,12)
> par(mfrow=c(3,1))
> ts.plot(dy12)
> acf(dy12,40)
> pacf(dy12,40)
```



Para obter os valores de P e Q basta olhar para os picos fora do intervalo de confiança nos lags sazonais da PACF e da ACF, respectivamente. Assim tomaremos inicialmente, Q=2 e P=2.

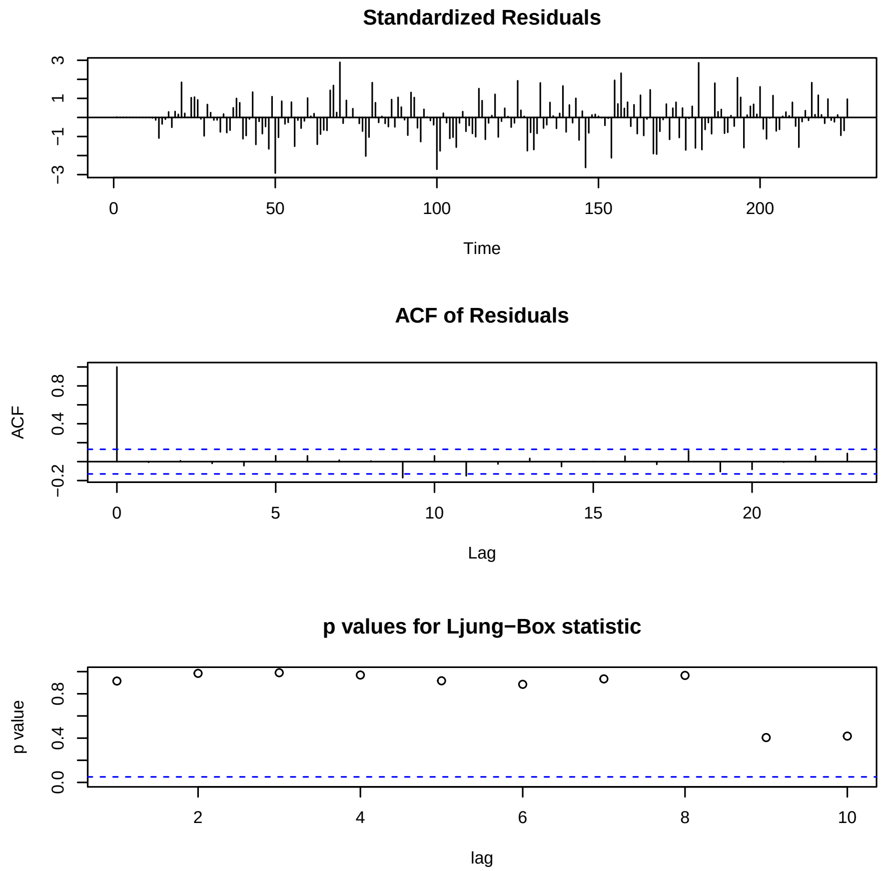
Já para obter os valores de p e q para a parte não sazonal do modelo, observamos quantos picos fora do intervalo de confiança ocorrem entre os lags sazonais na PACF e na ACF. Assim, podemos tomar p,q=1,2,3 e verificar qual modelo tem menor AIC.

```
> model = NULL
> aic=NULL
> for (p in 1:3)
+ for (q in 1:3)
+ for (d in 1:1){
+   model = rbind(model,c(p,d,q))
+   aic=c(aic,arima(y,order=c(p,d,q),
+            seasonal=list(order=c(2,1,2),period=12),method="ML")$aic)
+  }
```

```
> # Modelos e respectivos AIC
> cbind(model,aic)

                aic
 [1,] 1 1 1 91.67725
 [2,] 1 1 2 48.88383
 [3,] 1 1 3 39.49949
 [4,] 2 1 1 79.91845
 [5,] 2 1 2 42.63201
 [6,] 2 1 3 40.62021
 [7,] 3 1 1 56.69558
 [8,] 3 1 2 43.82829
 [9,] 3 1 3 42.56398

> # Melhor modelo (menor AIC)
> model[aic==min(aic)]

[1] 1 1 3
```

Ou seja, o modelo com menor AIC é o SARIMA(1,1,3)(2,1,2)$_{12}$. Vamos verificar o ajuste do modelo:

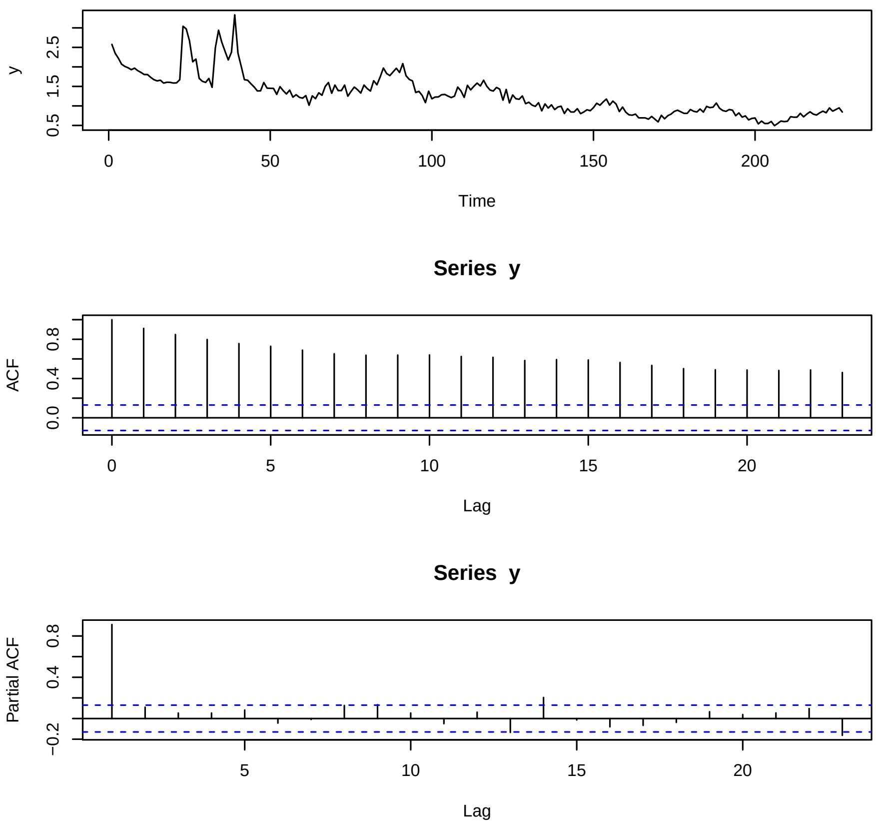
```
> fit=arima(y,order=c(1,1,3),seasonal=list(order=c(2,1,2),period=12),method="ML")
> tsdiag(fit)
```



Verificamos que o modelo está corretamente ajustado, uma vez que os resíduos são não autocorrelacionados.

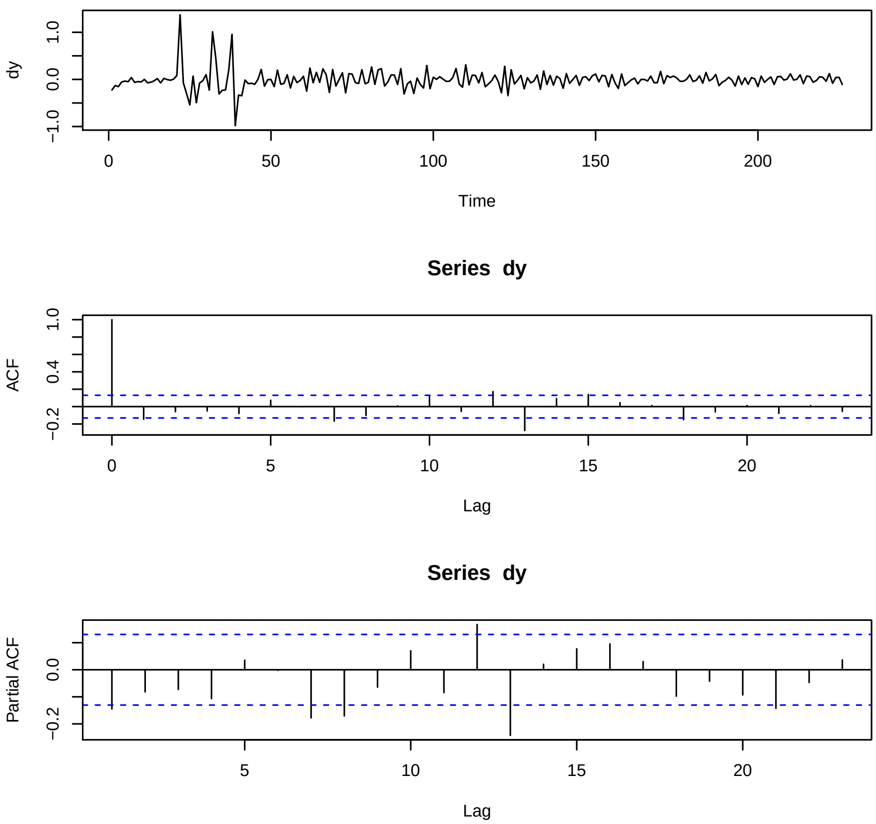
Agora vamos analisar a série de Juros:

```
> y  = Juros
> n  = length(y)
> # Grafico da serie temporal, sua funcao de aucorrelacao (FAC) e
> # sua funcao de autocorrelacao parcial (FACP)
> par(mfrow=c(3,1))
> ts.plot(y)
> acf(y)
> pacf(y)
```



Podemos ver que a série não é estacionária, vamos verificar, portanto, a primeira diferença:

```
> y=Juros
> dy=diff(y,1)
> # Grafico da serie temporal, sua funcao de aucorrelacao (FAC) e
> # sua funcao de autocorrelacao parcial (FACP)
> par(mfrow=c(3,1))
> ts.plot(dy)
> acf(dy)
> pacf(dy)
```

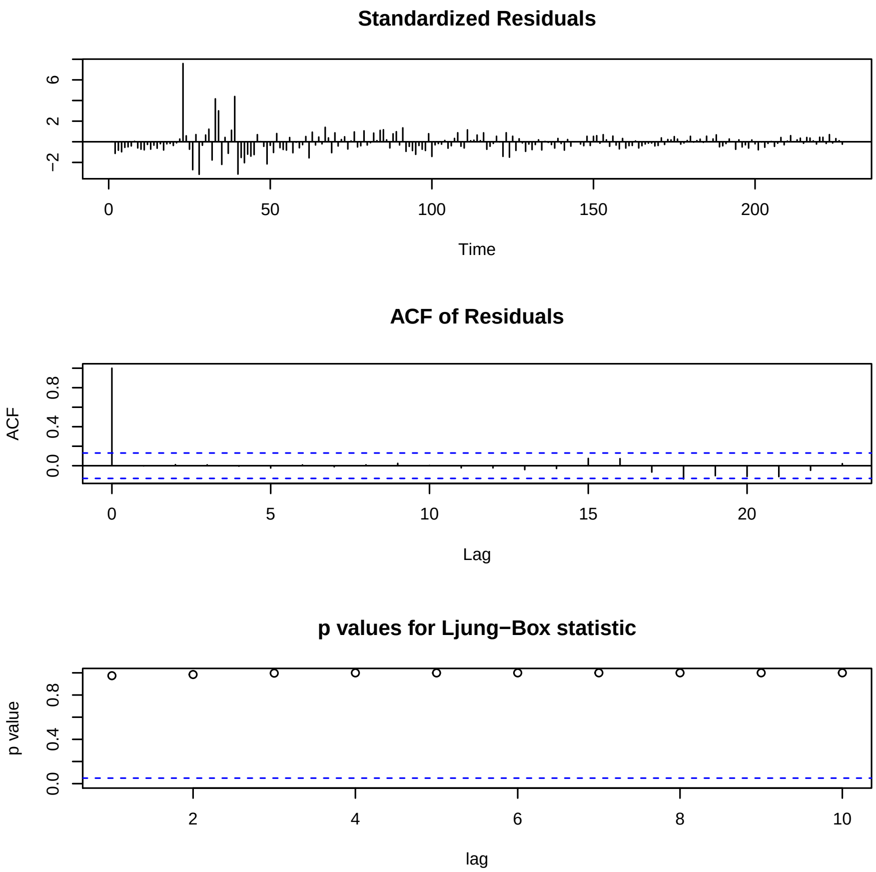


Vemos praticamente toda a ACF está dentro do intervalo de confiança, mas não a PACF, em que o lag 13 é estatisticamente significante, por exemplo. Vamos rodar um modelo ARIMA(1,1,1) e um modelo ARIMA(13,1,0) e verificar qual o menor AIC.

```
> fit=arima(y,order=c(13,1,0))
> fit2=arima(y,order=c(1,1,1))
> fit$aic

[1] -103.1677

> fit2$aic

[1] -92.96227
```

Verificamos que neste caso, o modelo ARIMA(13,1,0) é mais adequado. Vamos verificar o ajuste do mesmo:

```
> tsdiag(fit)
```



O modelo está corretamente ajustado.