

# Capítulo 3

## Exercício 03

### Vítimas de rodovias na Grã-Bretanha 1969–1984.

UKDriverDeaths é uma série temporal que fornece os totais mensais de motoristas de automóveis na Grã-Bretanha mortos ou gravemente feridos de janeiro de 1969 a dezembro de 1984. O uso obrigatório de cintos de segurança foi introduzido em 31 de janeiro de 1983.

### 3.1 Análise Descritiva

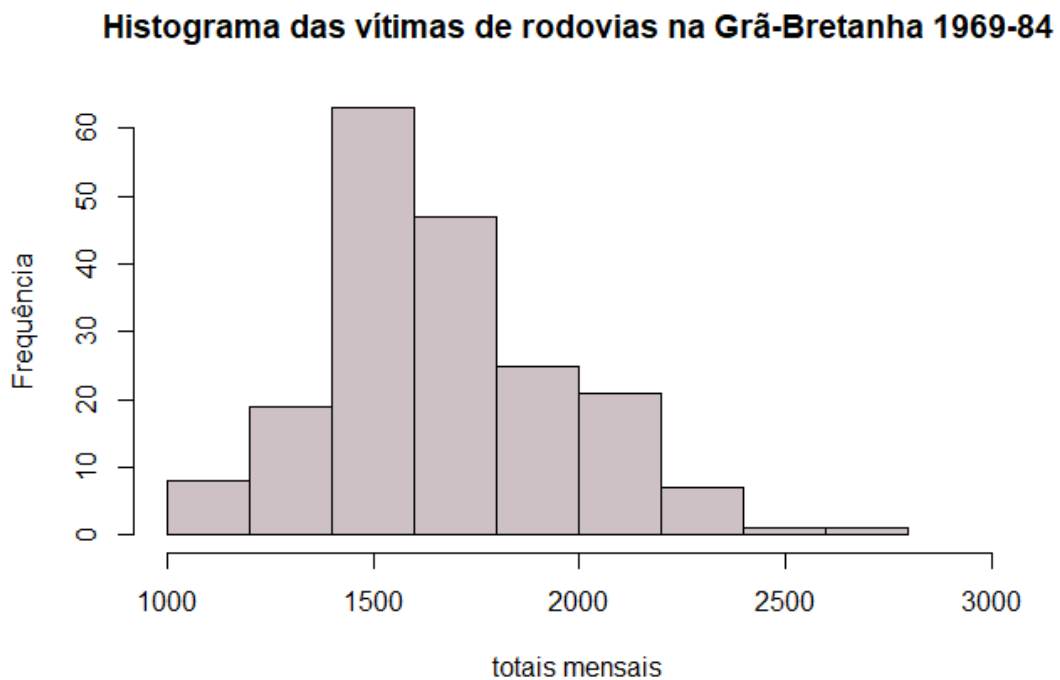


Figura 3.1: Gráfico Histograma das vítimas de rodovias na Grã-Bretanha 1969-84

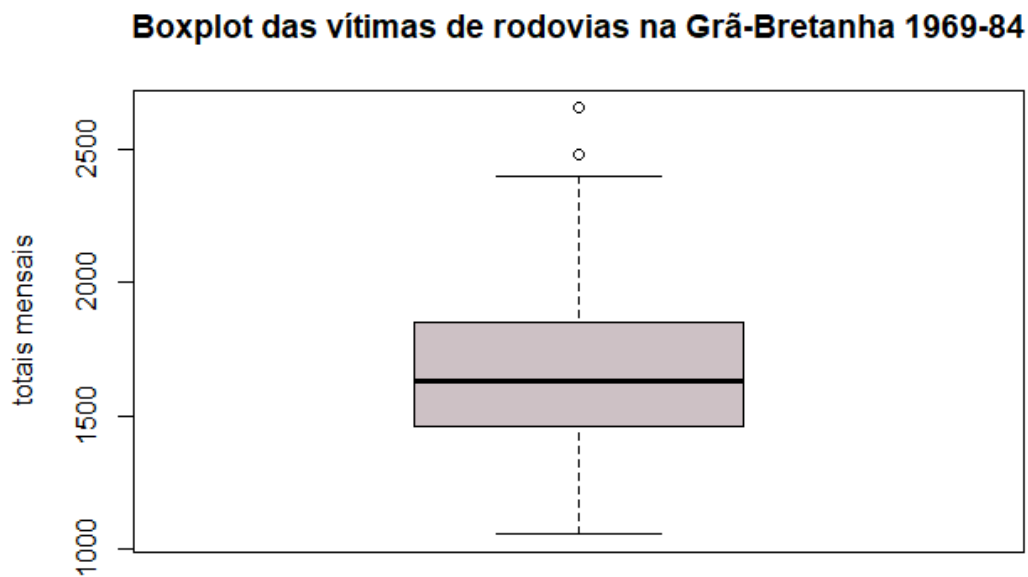


Figura 3.2: Gráfico Boxplot das vítimas de rodovias na Grã-Bretanha 1969-84

Tabela 3.1: Summary da Serie

Min	1º Quartil	Mediana	Média	3º Quartil	Max
1057	1462	1631	1670	1851	2654

Analisando os gráficos e a tabela de resumo da análise foram obtidas as seguintes conclusões:

- Podemos observar que o histograma tem assimetria à esquerda: a frequência é mais alta entre 1500 e cerca de até 1750, depois ocorre um decréscimo gradual para a calda a esquerda. A média e a mediana são aproximadamente iguais, como podemos observar na nossa tabela.
- Podemos observar que existe presença de dois outliers no boxplot.

## 3.2 Análise da Série Temporal

Agora iremos mostrar um plot de uma série temporal para a nossa análise correspondente.

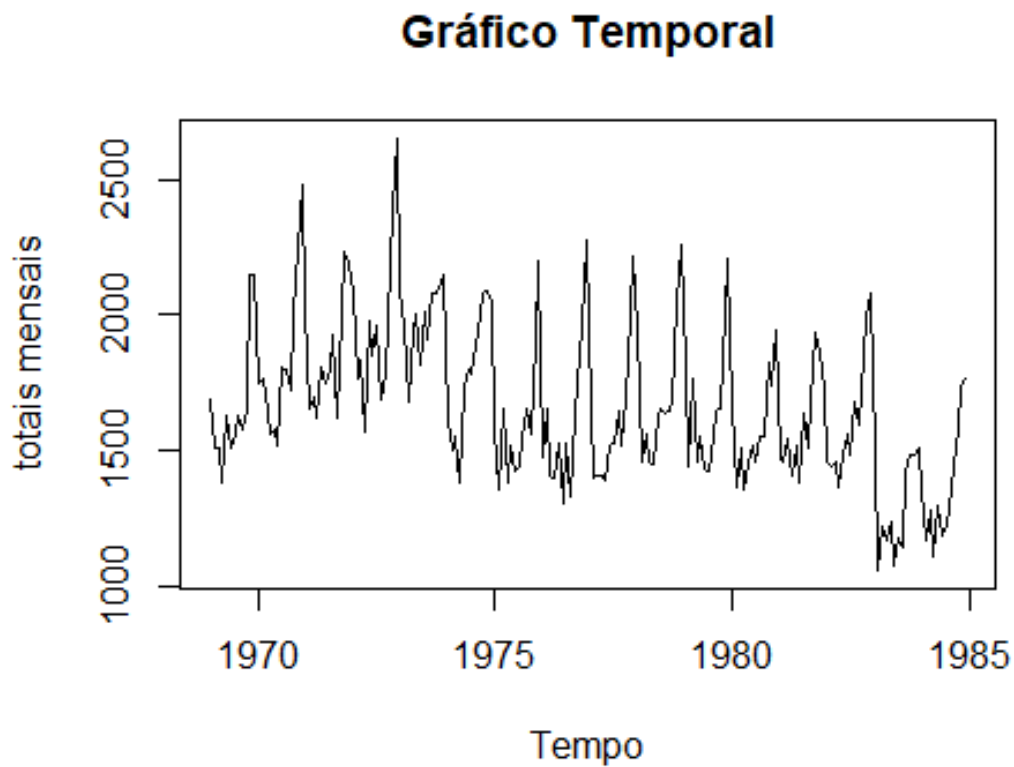


Figura 3.3: Gráfico Temporal

A princípio, a série temporal aparenta, ter a existência de possíveis tendência. Além disso, também podemos perceber haver um padrão no comportamento da série, o que indica haver presença de sazonalidade. E na possível ausência de heterodasticidade apresentada pela Série. Adiante, analisaremos a fundo para confirmar a presença ou não dessas características.

Iremos primeiramente analisar a tendência da série.

### 3.2.1 Análise de Tendência

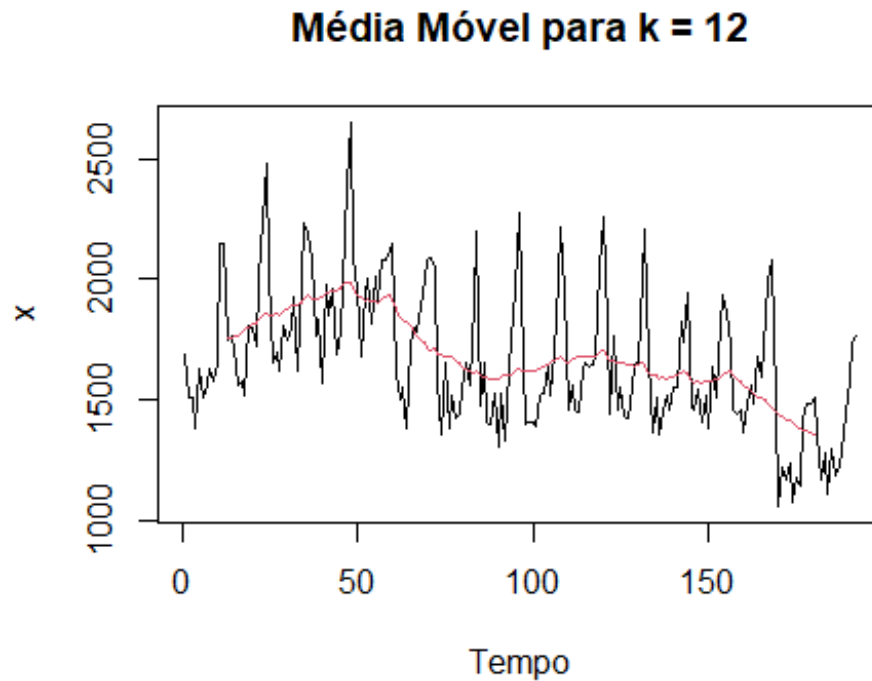


Figura 3.4: Gráfico de Médias Móveis

Pelo gráfico acima, a linha vermelha referente à média móvel de 12 meses já nos remete a uma tendência dos dados, uma vez que esta média não se mantém aproximadamente constante.

Para nos auxiliar na análise, realizamos os testes Wald-Wolfowitz e o Cox Stuart, onde:

$$\begin{cases} H_0 : \text{A série temporal não tem tendência.} \\ H_1 : \text{A série temporal tem tendência.} \end{cases}$$

E obtemos os seguintes resultados:

Tabela 3.2: Teste Wald-Wolfowitz e Cox Stuart

Teste	$n_1$	$n_2$	$n$	p-valor
Wald-Wolfowitz	96	96	192	4.626e-13
Cox Stuart	-	-	96	2.547e-15

Assim, como nos dois testes, temos o *p-valor* menor do que  $\alpha = 0.05$ , rejeitamos  $H_0$ , ou seja, temos evidências de que há tendência nos dados.

### 3.2.2 Análise de Heterocedasticidade

Para verificar se a série apresenta heterocedasticidade, realizamos o gráfico para analisarmos a relação entre a média e a amplitude dos dados tomando um intervalo de 12 meses.

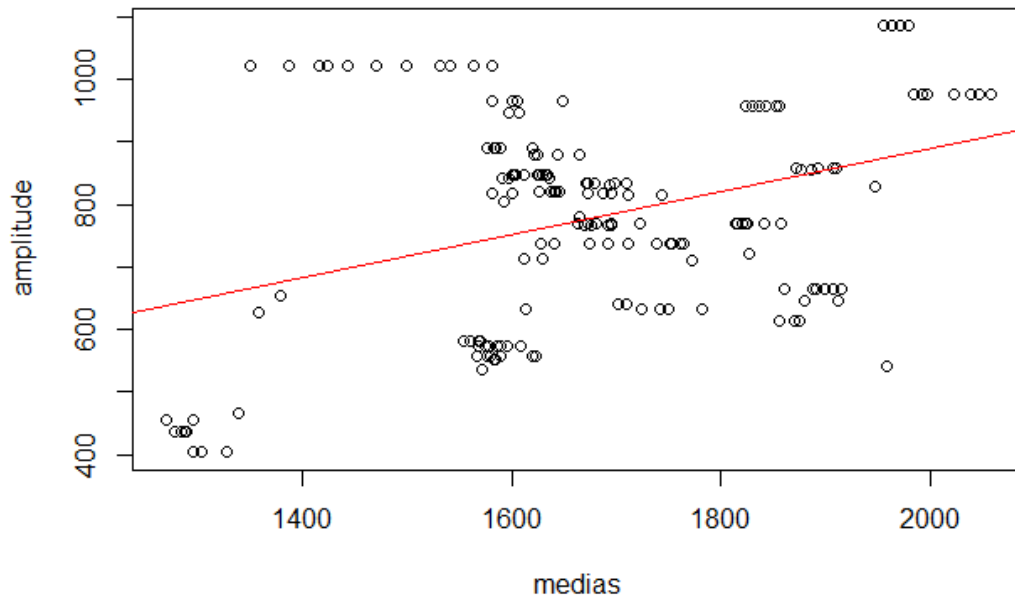


Figura 3.5: Gráfico de amplitude média da Série Temporal UKDriverDeaths

De acordo com o gráfico acima, podemos perceber que os pontos, num geral, tem um comportamento indicando que a amplitude não aumenta conforme a média aumenta, ou seja, pode-se observar que a variância não se mantém constante.

### 3.2.3 Análise de Sazonalidade

Para verificar se a série é sazonal, realizamos a análise gráfica por mês do ano 1969 até 1984.

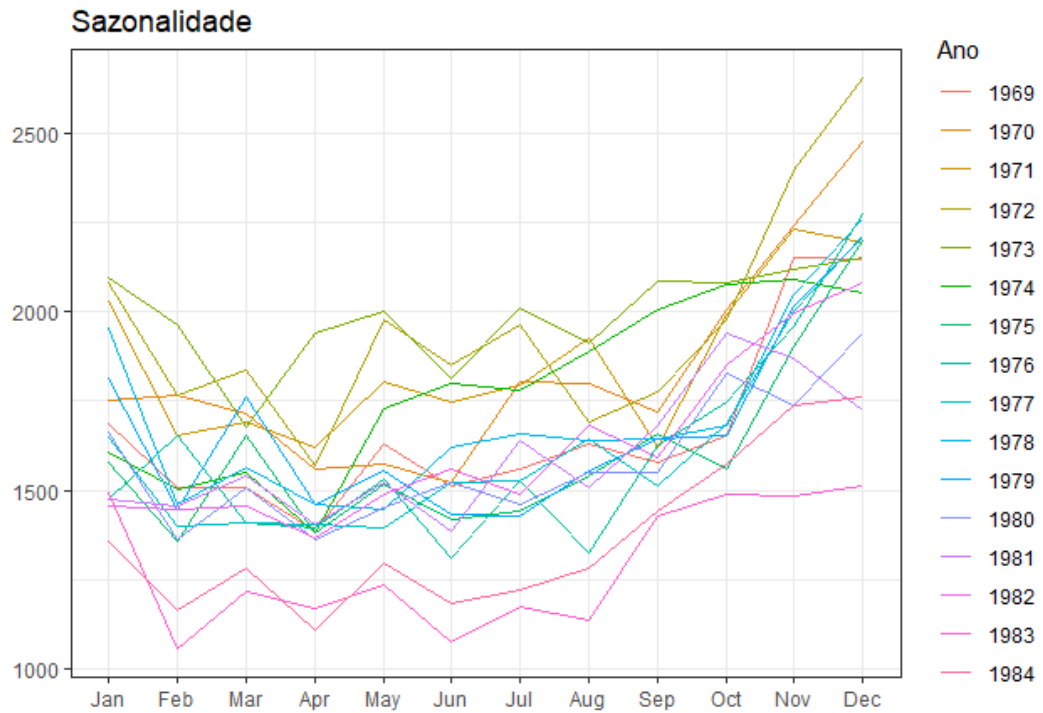
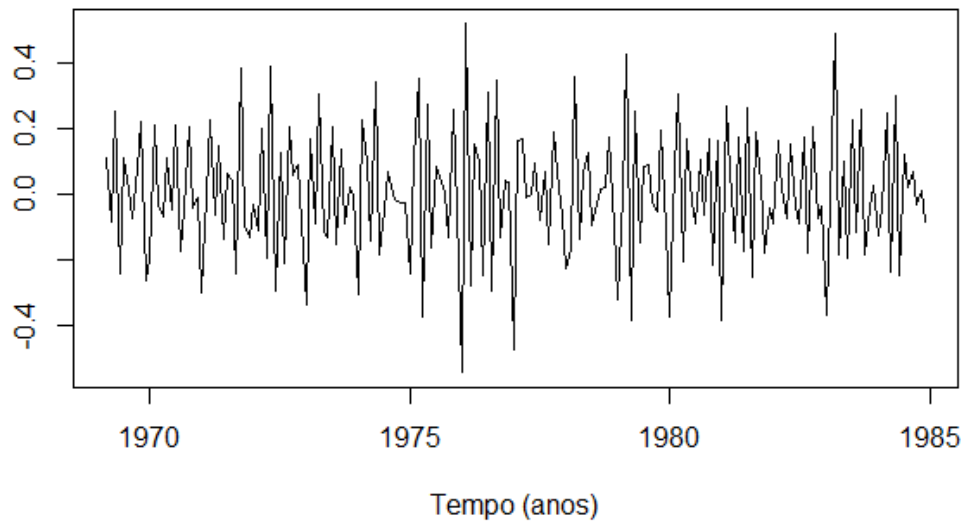


Figura 3.6: Gráfico de Sazonalidade da Série Temporal UKDriverDeaths

Podemos notar que as observações anuais possuem trajetórias similares, o que indica a presença de sazonalidade. Com isso observa-se um padrão de comportamento que se repete a cada ano. Embora o comportamento não seja o mesmo, podemos verificar algumas semelhanças, como, por exemplo, temos que em todos os anos ha um aumento no número de mortes de acidentes entre outubro e dezembro.

Fazendo a diferença em um período de 12 meses para remover a tendência e a sazonalidade. E para retirar a heterocedasticidade iremos estar fazendo o log da série.

### Gráfico temporal sem tendência e sem sazonalidade



Agora, portanto, temos uma série temporal sem tendência e sem comportamento sazonal. Podemos, de outra forma, decompor nossa série inicial utilizando a função `decompose()`, do R.

### Decomposition of additive time series

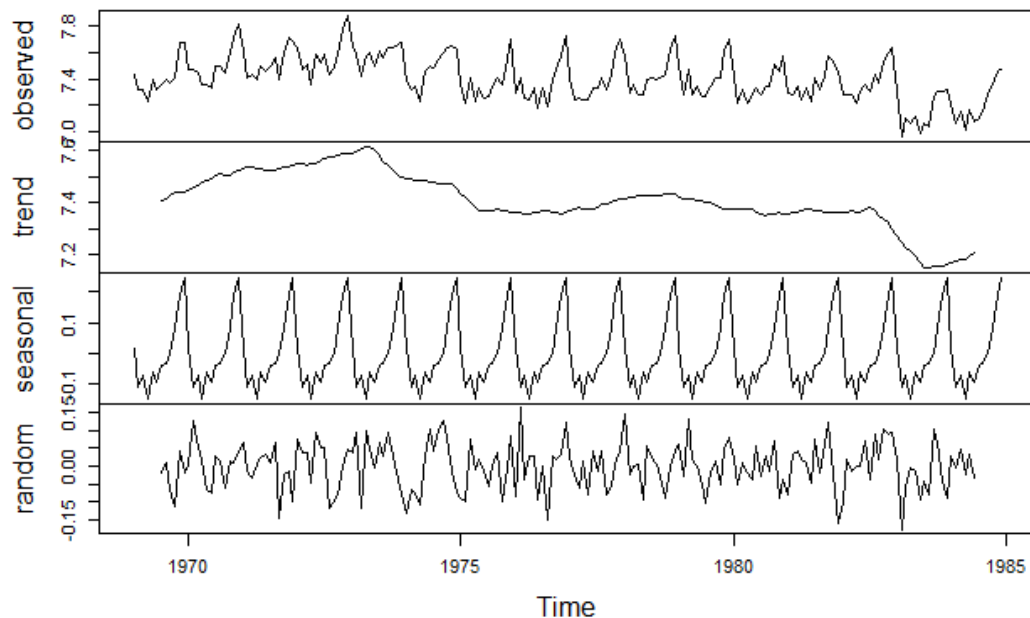


Figura 3.7: Gráfico Decomposição da Série Temporal Aditiva - Série Temporal UKDriverDeaths

A partir do gráfico de Decomposição da Série Temporal Aditiva, realizamos o cálculo de amplitude dos mesmos, para podermos ver quanto cada componente esta contribuindo.

- Assim, a partir do gráfico da série temos que a amplitude da série é de 0.9206332.
- O gráfico da Tendência temos que a amplitude da tendência é de 0.4688247.
- O gráfico da Sazonalidade temos que a amplitude da sazonalidade é de 0.3926792.
- E para o gráfico Aleatório, temos que a amplitude do componente aleatório é de 0.3369406.

Portanto, podemos notar que o componente Aleatório e o componente sazonalidade são os com menor contribuição para a série, já a tendência, é o componente que mais contribui para a série.

### 3.3 Modelo Ajustado

Ajustando um modelo para a série UKDriverDeaths

#### 3.3.1 Função de autocorrelação (ACF)

Primeiramente, vamos avaliar a função de autocorrelação da série.

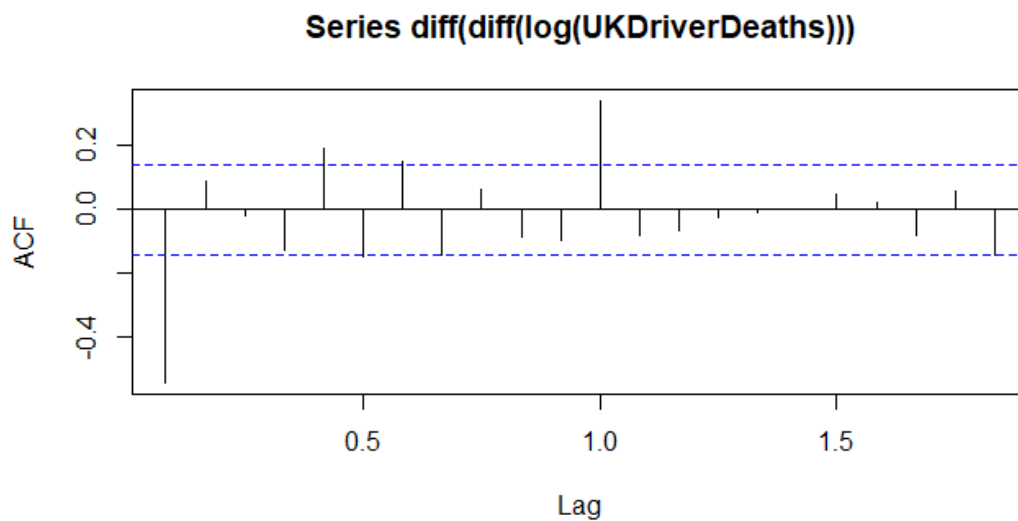


Figura 3.8: Gráfico da Função de autocorrelação da serie

Neste gráfico, podemos observar correlação significativas no primeiro lag. O número de correlações significativas indica a ordem do termo da média móvel. Notamos também que temos  $m = 12$  períodos sazonais.



### 3.3.2 Função de autocorrelação Parcial (PACF)

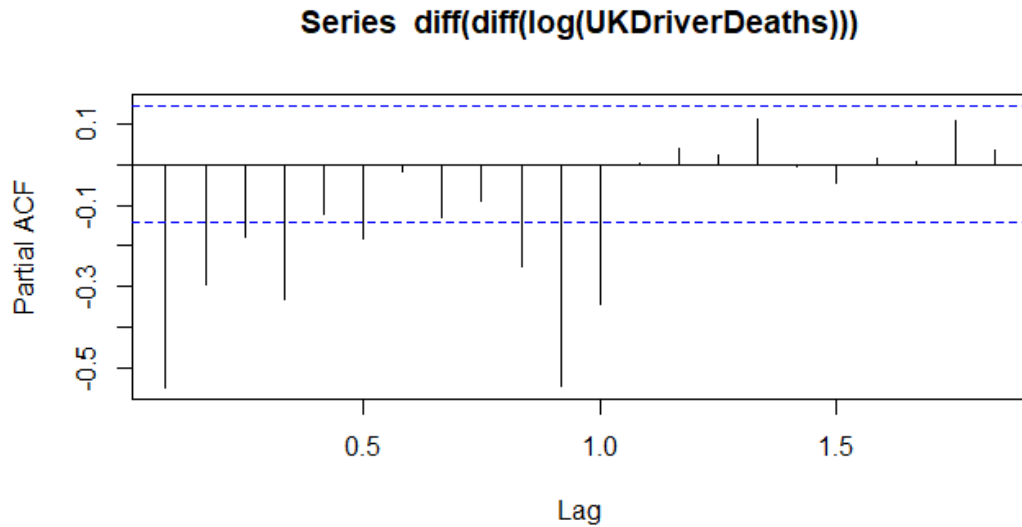


Figura 3.9: Gráfico da Função de autocorrelação Parcial da serie

Podemos observar um grande pico no lag 1 que diminui depois de alguns lags. Um termo de média móvel nos dados.

### 3.3.3 Escolhendo o modelo: SARIMA

Modelos ARIMA são capazes também de modelar séries que apresentam um componente sazonal, sendo descrito como:

$$\text{ARIMA}(p, d, q)(P, D, Q)_m.$$

Onde o primeiro parênteses se refere à parte não-sazonal do modelo e o segundo à parte sazonal.  $m$  corresponde ao número de períodos sazonais.

Ajustar um modelo SARIMA é semelhante ao processo de ajustar um ARIMA.

Olhando os gráficos temos que a PACF sugere um modelo AR(1) enquanto a ACF sugere um modelo MA(1), foram testados usando diferenciação ou não, chegando a conclusão de que os melhores modelos são os sem diferenciação, para a parte não sazonal. Já para a parte sazonal tem um AR(0), MA(1) e usando diferenciação. Assim, vamos estimar todas as combinações para SARIMA  $(1, 0, 1)(0, 1, 1)_{12}$ , e obtemos:

Tabela 3.3: Ajuste modelo SARIMA  $(1, 0, 1)(0, 1, 1)_{12}$

	ar1	ma1	sma1
	0.9557	-0.5499	-0.8718
s.e.	0.0318	0.0838	0.0764

Além disso, através do ajuste temos  $\sigma^2 = 0.006284$ , Log verossimilhança = 192.43, aic = -378.86

Verificando se o modelo é adequado

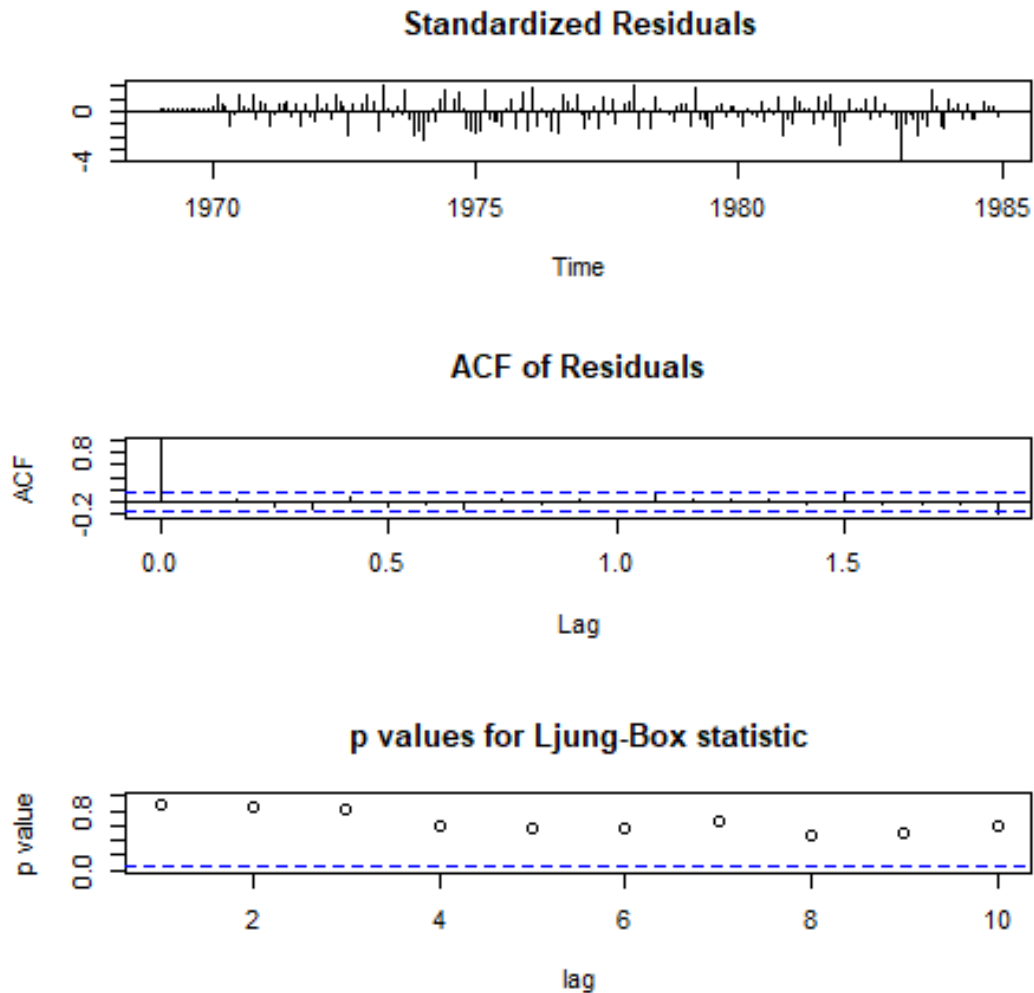


Figura 3.10: Gráfico do teste Ljung-Box para o ajuste SARIMA (1,0,1)(0,1,1)

No primeiro gráfico temos os resíduos padronizados, note que, a princípio não vemos nenhum padrão neles, indicando aleatoriedade nos resíduos.

No segundo gráfico temos a Função de Auto Correlação dos resíduos, onde no lag 0 o valor é 1 e depois as correlações são muito pequenas e estão todas no intervalo, com isso, meus resíduos são ruídos branco.

No terceiro gráfico, temos o teste Ljung-Box, note que todos os p-valores são maiores que o nível de significância  $\alpha = 0.05$ , ou seja, temos evidências de que os resíduos são ruído branco.

Portanto, o ajuste SARIMA(1,0,1)(0,1,1) apresenta fortes resultados que, de fato os resíduos são ruído branco e possuem aleatoriedade, então podemos assim concluir que para a nossa série temporal o ajuste com o modelo SARIMA (1,0,1)(0,1,1) é adequado.

### 3.4 Análise de Intervenção

Agora será incorporado uma intervenção sabendo que em 31/01/1983 teve início a obrigatoriedade de uso de cinto de segurança.

Os coeficientes ajustados, o diagnóstico do modelo e o gráfico do ajuste são dados a seguir:

Tabela 3.4: Coeficientes do modelo ajustado com a intervenção

	<b>ar1</b>	<b>ma1</b>	<b>sma1</b>	<b>I3101-MA0</b>
Estimativas	0.9589	-0.5480	-0.8919	-0.1989
s.e.	0.0311	0.0812	0.0884	0.0694

Além disso, através do ajuste com a intervenção obtemos uma variância estimada menor  $\sigma^2 = 0.005953$ , Log verossimilhança = 196.45 maior e aic = -384.9 menor, que os estimados com o modelo sem a intervenção.

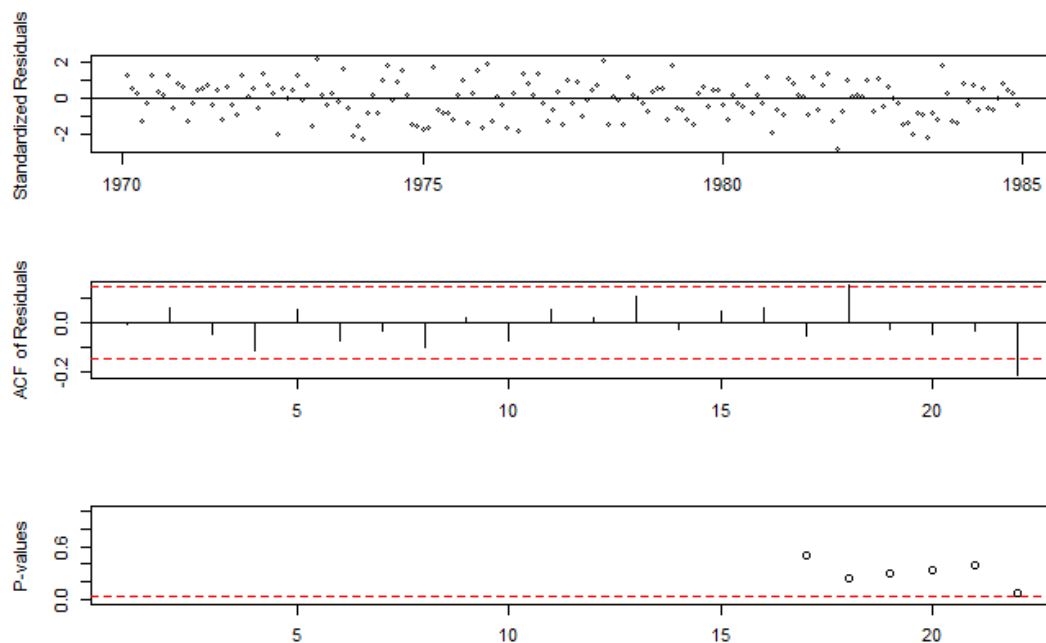


Figura 3.11: Diagnóstico do modelo

No primeiro gráfico temos os resíduos padronizados, note que, a princípio não vemos nenhum padrão neles, indicando aleatoriedade nos resíduos.

No segundo gráfico temos a Função de Auto Correlação dos resíduos, onde notamos indícios de que os resíduos são ruídos branco.

No terceiro gráfico, temos o teste Ljung-Box, note que todos os p-valores são maiores que o nível de significância  $\alpha = 0.05$ , ou seja, temos evidências de que os resíduos são ruído branco.

Portanto, o ajuste com a intervenção apresenta fortes resultados que, de fato os resíduos são ruído branco e possuem aleatoriedade, então podemos assim concluir que para a nossa série o ajuste com a intervenção é adequado.

Assim, com o ajuste incorporado à intervenção de que em 31/01/1983 teve início a obrigatoriedade de uso de cinto de segurança, plotamos o Gráfico valores ajustados, onde as linhas representam a série e os pontos os valores estimados.

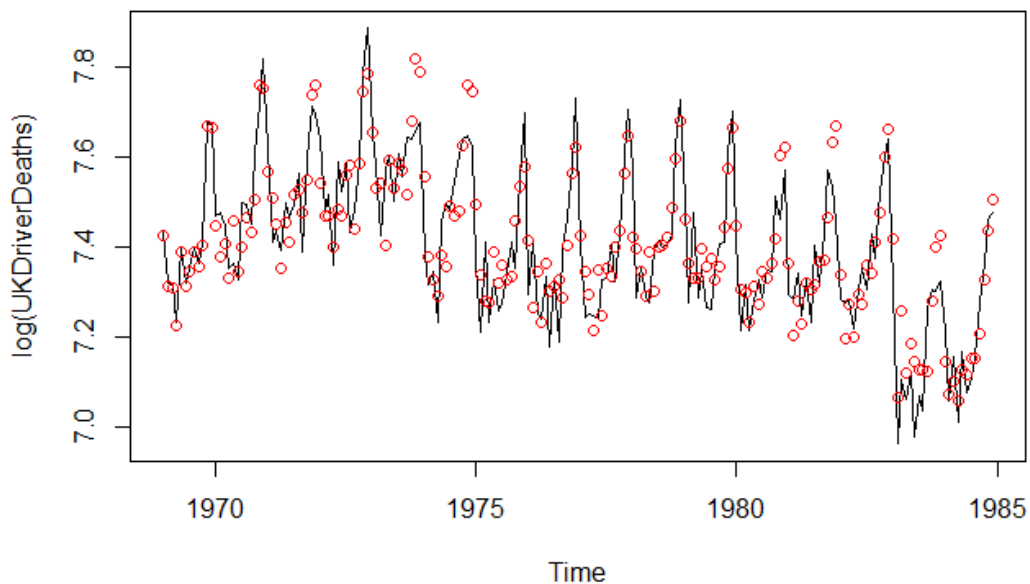


Figura 3.12: Gráfico valores ajustados

Como podemos observar nossos pontos em alguns instantes de tempo não estão tão precisos, como no ano de 1974 e 1975, pois, a partir do ano 1973 tivemos uma queda diferente do que foi estimado, isso é pela crise petrolífera de 1973 que teve início em outubro, onde o preço do barril de petróleo subiu de três dólares para 12 dólares no

mundo inteiro. Consequentemente menos pessoas passaram a usar seus carros. Com isso, como essa crise afetou a economia dos países, teve uma queda no uso de transportes, como foi visto no ano de 1974 e 1975, por conta desse acontecimento podemos observar que o nosso modelo fez uma previsão menor do que o esperado para o ano de 1976 até 1980, além disso, as previsões de 1976 até 1980 tem uma certa estabilidade um padrão de repetição entre os dados. E assim a partir de 1981 a previsão ficou maior do que o esperado e assim sucessivamente. O modelo não está com a precisão que gostaríamos, mas no futuro será visto outras técnicas para melhorar. Mas é importante ressaltar que olhando graficamente, tivemos um ótimo ajuste para os nossos dados.

**Referência Bibliográfica usado no item 3** <https://www.stoodi.com.br/blog/historia/crise-do-petroleo-o-que-foi/>