

Lista 2: Séries Temporais

André Felipe Berdusco Menezes

Profa: Dra. Eniuce Menezes

Departamento de Estatística, Universidade Estadual de Maringá

Maringá, PR, Brazil

1. É importante lembrar que até o momento a definição de série estacionária é aquela na qual a média e variância são constantes ao longo do tempo. Seguindo a definição de Pérez (2017) podemos dizer que a estacionariedade significa que o caráter probabilístico da série não muda ao longo do tempo, ou seja, qualquer seção de série temporal é típica para todas as outras seções com o mesmo comprimento.

(a) Analisando os gráficos das séries temporais da Figura 1 podemos notar empiricamente que não há indícios de estacionariedade em ambas as variáveis, temperatura (à direita) e umidade (à esquerda). Essa conclusão é baseada no fato de que em ambas as séries observa-se bruscas variações em torno da média (reta horizontal). Os boxplots da cada variável conforme o mês também indicam a não estacionariedade da série. De fato, a variável temperatura difere substancialmente em relação a média (mediana), bem como a amplitude (variabilidade). Já a variável umidade, embora apresente medianas e médias próximas, nota-se que existe alta dispersão conforme o mês, outro indicativo empírico de que a série é não estacionária.

Por fim, vale ressaltar que a série referente a temperatura apresenta uma tendência de decréscimo até o mês de Junho e posteriormente fica evidente que a mesma exibe um comportamento crescente até o fim do ano. Já na série referente a umidade não foi possível observar claramente alguma tendência.

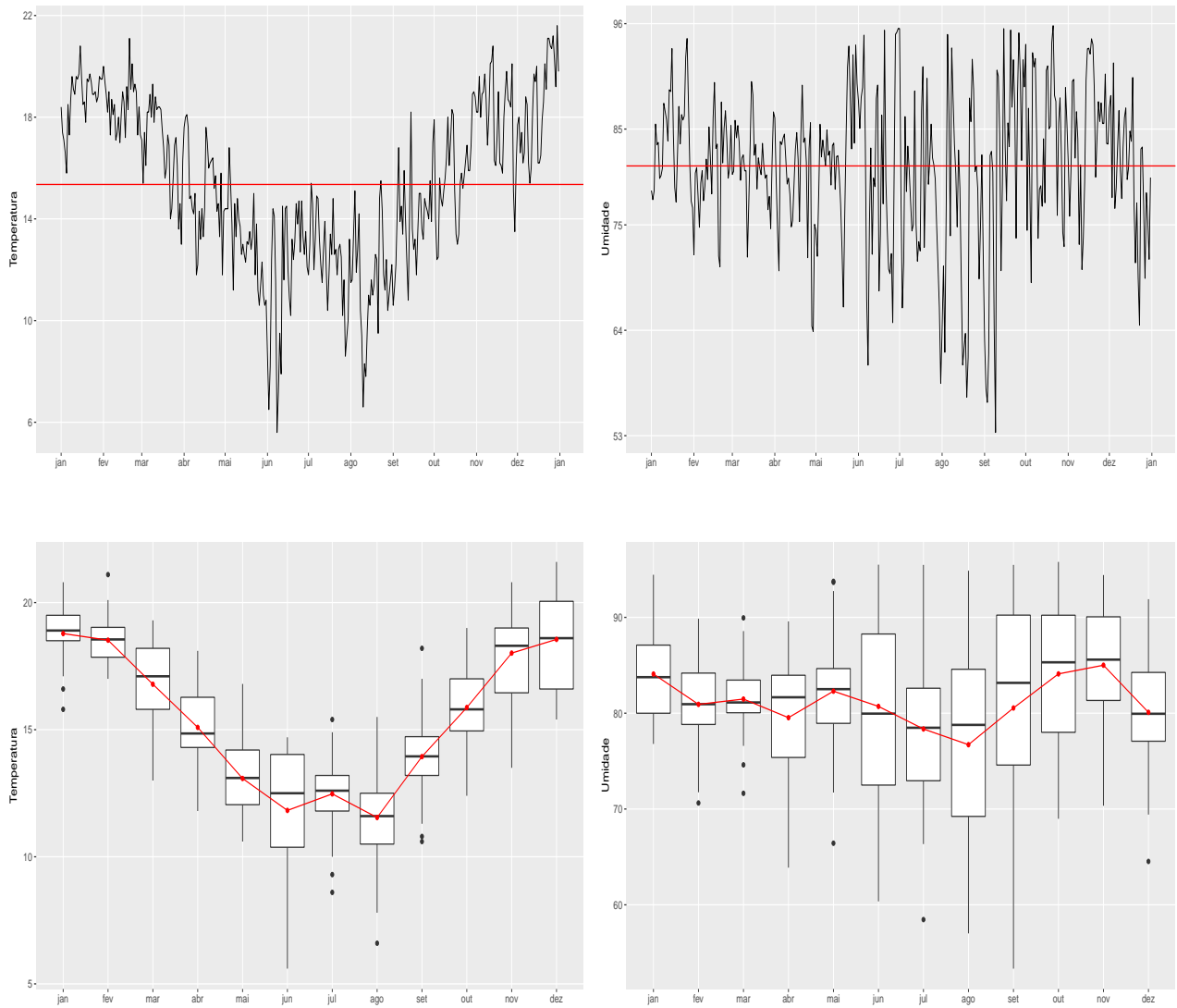


Figura 1: Temperatura e umidade diária do ano de 1997.

- (b) Na Figura 2 apresenta-se as séries obtidas pela primeira e segunda diferenciação. Pode-se notar um comportamento aleatório em torno da média para ambas as variáveis, obviamente a série de segunda diferenciação possui comportamento mais próximo da média. No entanto em todas as séries é possível observar uma forte variabilidade, indicando que elas não possuem estacionariedade.

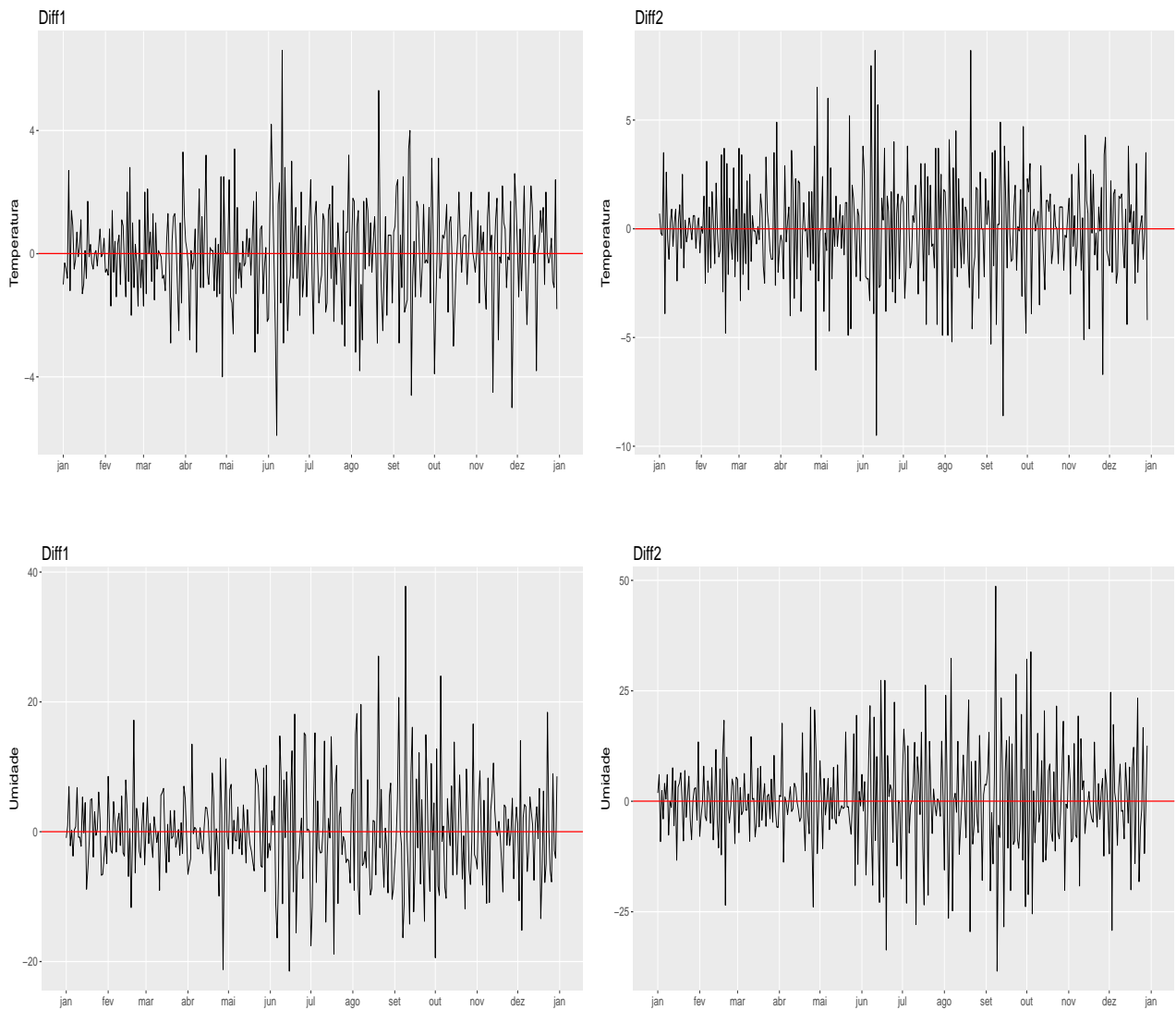


Figura 2: Séries das diferenciações da temperatura e umidade diária do ano de 1997.

Por fim, o comportamento descrito anteriormente é observado nos boxplots da Figura 3. Fica evidente que todas as séries possuem médias (medianas) muito semelhantes ao longo dos meses, entretanto o verifica-se que uma forte variabilidade, ou seja, as amplitudes dos boxplots são diferentes.

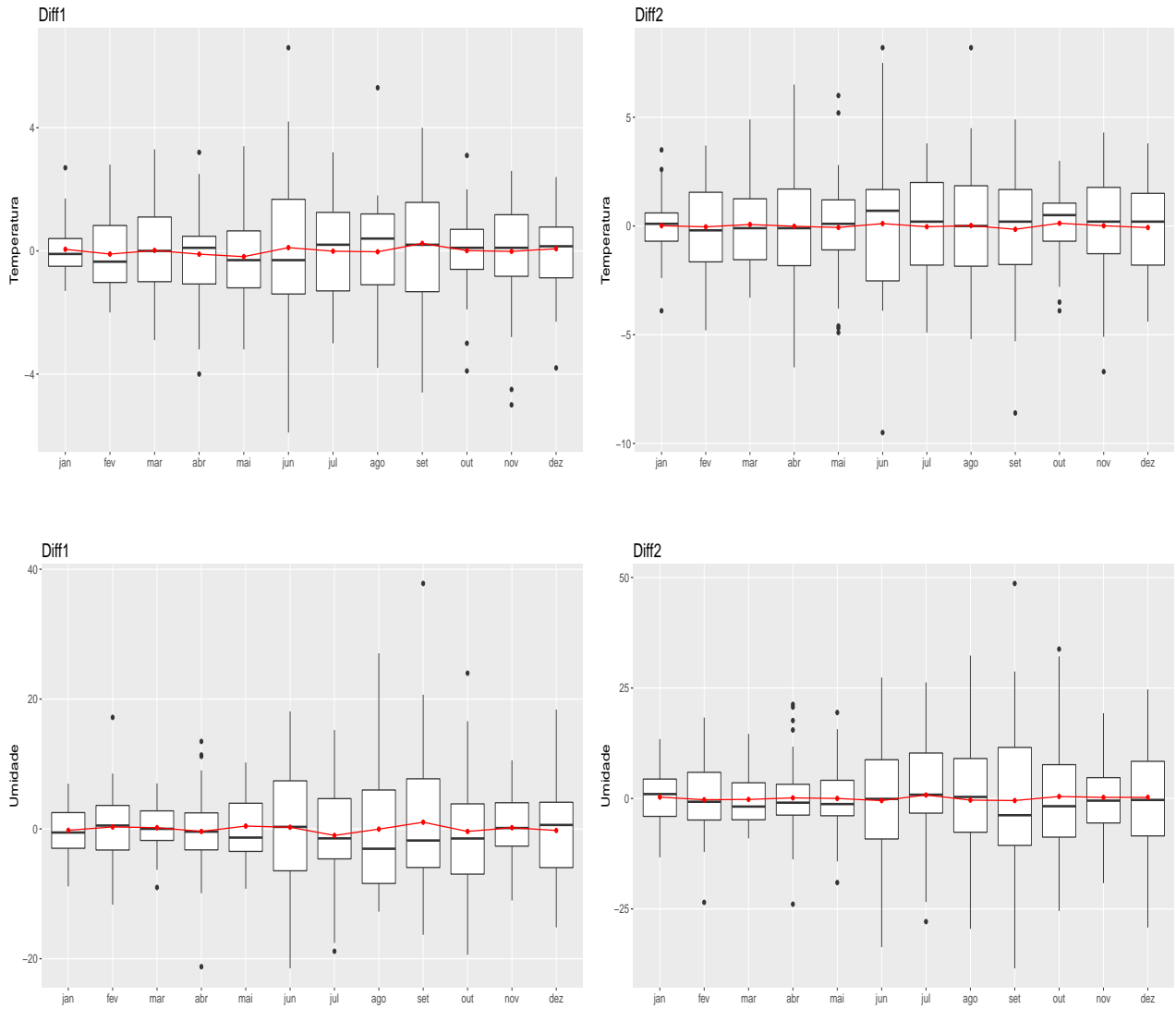


Figura 3: Boxplots das séries de diferenciações da temperatura e umidade diária do ano de 1997.

2. (a) Para verificar a estacionariedade da série temporal apresentamos na Figura 4 os gráficos das observações ao longo do tempo, o boxplot das vendas conforme o mês e o boxplot das vendas anuais. Nota-se no primeiro gráfico que a série apresenta um comportamento sazonal. Já nos boxplots pode-se notar que as maiores vendas ocorrem no mês de Maio e que existe uma leve tendência linear de crescimento das vendas com o decorrer dos anos. Assim, fica evidente que a série não é estacionária.

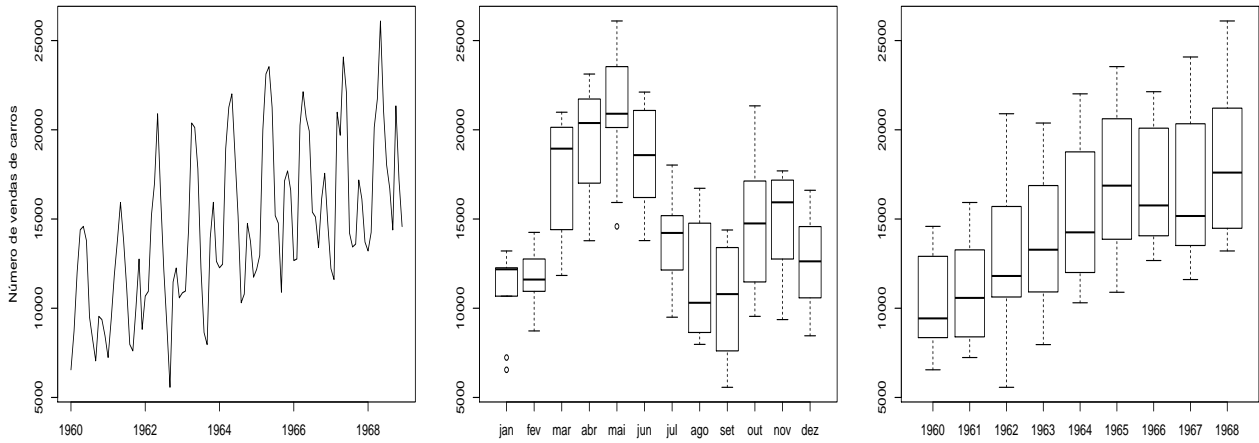


Figura 4: Vendas mensais de carros em Quebec 1960–1968

- (b) Os mesmos gráficos apresentados para a série original estão expostos para as séries de diferenciação de 1° e 2° ordem na Figura 5. Nota-se que nestas séries também há fortes evidências de que elas não são estacionárias. Além disto pode-se observar que a tendência anual de crescimento desapareceu das séries, no entanto ainda é perceptível a sazonalidade da série.

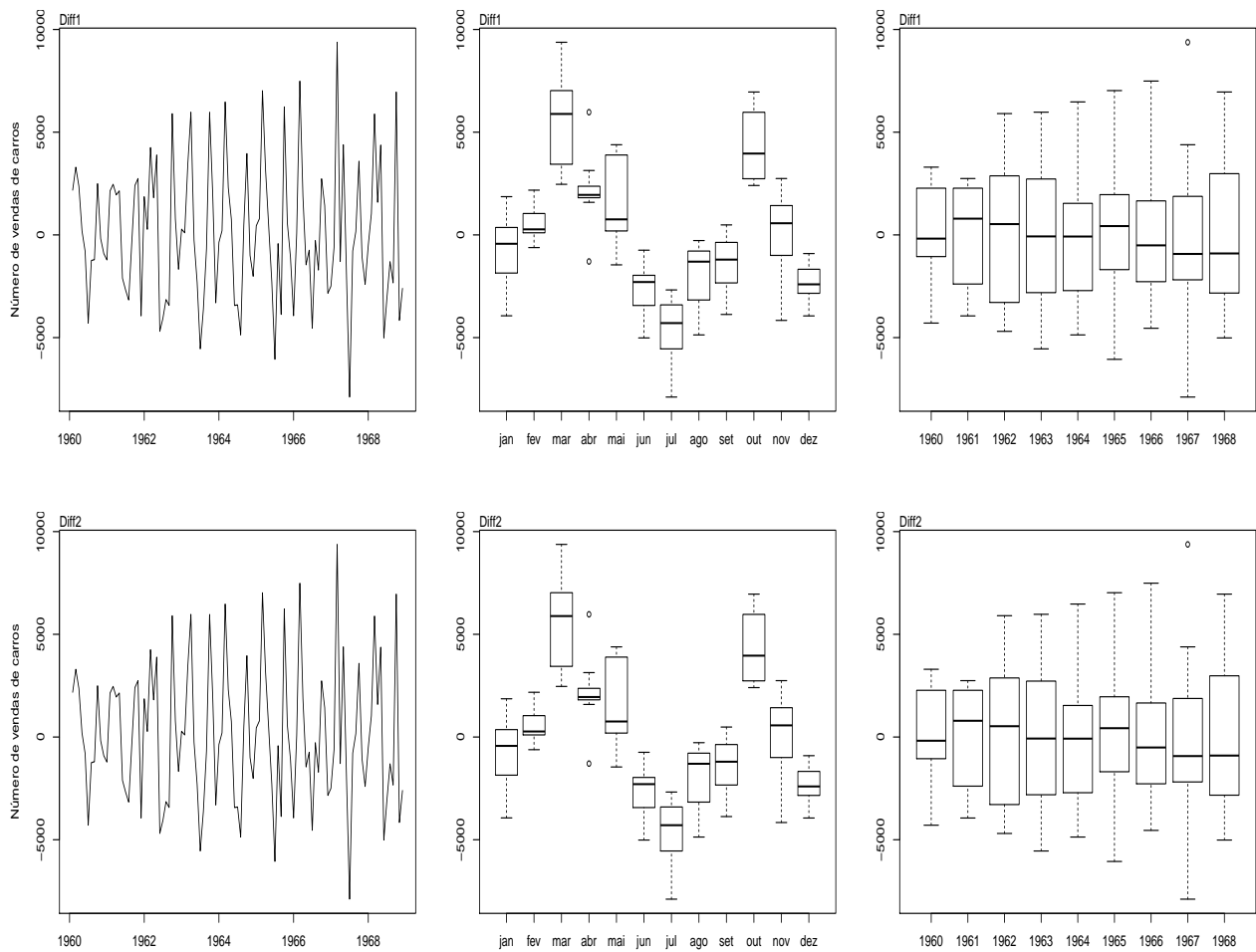


Figura 5: Séries da 1° e 2° diferenciação das vendas mensais de carros em Quebec 1960–1968

- (c) Na Figura 6 temos os gráficos para o logaritmo das vendas, pode-se notar que mesmo utilizando uma transformação para os dados a série ainda apresenta uma tendencia de crescimento ao decorrer do ano, bem como uma sazonalidade.

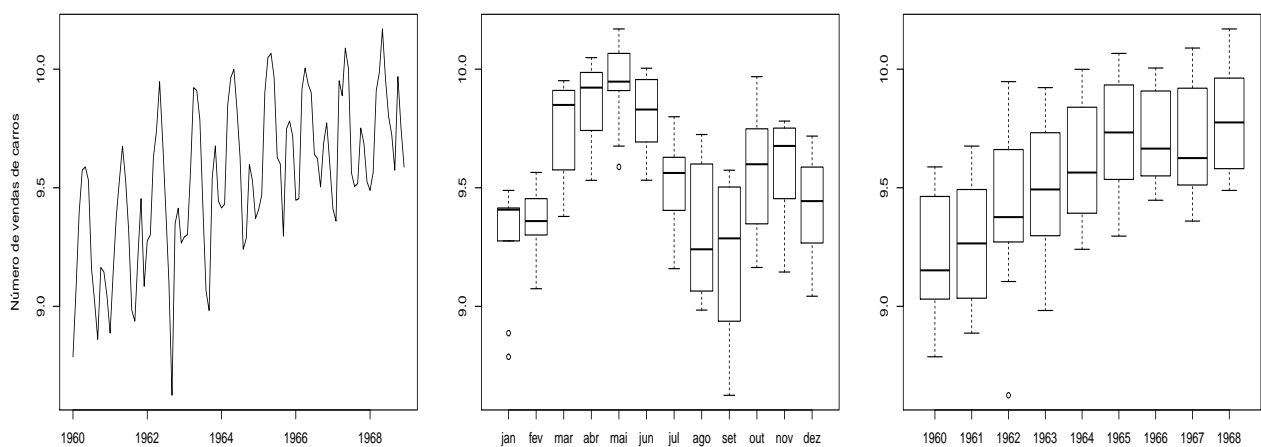


Figura 6: Logaritmo das vendas mensais de carros em Quebec 1960–1968

- (d) Na Figura 7 temos os gráficos para a diferenciação de primeira ordem da série utilizando o logaritmo da variável resposta. Pode-se notar que a tendência desapareceu da série, no entanto a variância ainda é muito alta especialmente entre os meses, indicando portanto que a série não é estacionária.

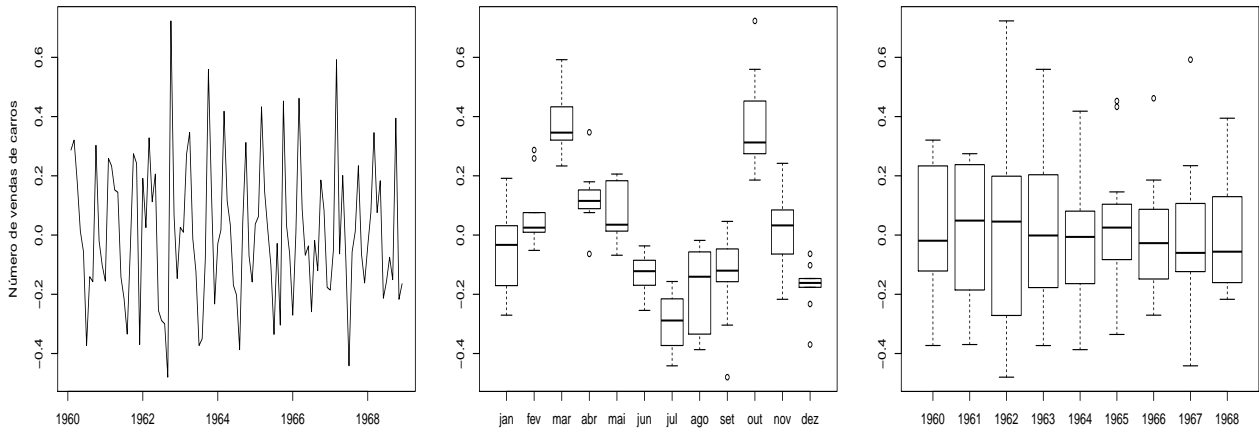


Figura 7: Séries de diferenciação do logaritmo das vendas mensais de carros em Quebec 1960–1968

- (e) Os mesmos gráficos utilizando considerando a raiz quadrada da resposta estão expostos na Figura 8. Observa-se que não há diferenças significativas entre a transformação $\log(X_t)$ e $\sqrt{X_t}$. Ou seja, a série permanece sendo não estacionária.

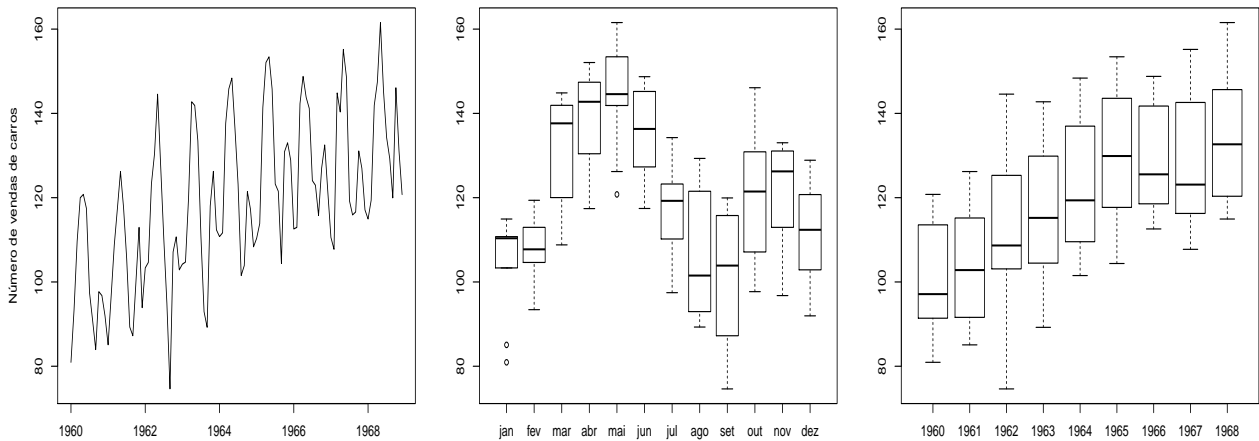


Figura 8: Raiz quadrada das vendas mensais de carros em Quebec 1960–1968

- (f) Em relação a primeira diferenciação considerando a raiz quadrada da variável resposta apresentamos os gráficos na Figura 9. Assim, como o logaritmo da variável resposta esta transformação mostrou que a série das diferenças não possui mais uma tendência linear anual. No entanto, fica evidente que a variância não é constante. Logo, mesmo com esta transformação a série continua não sendo estacionária.

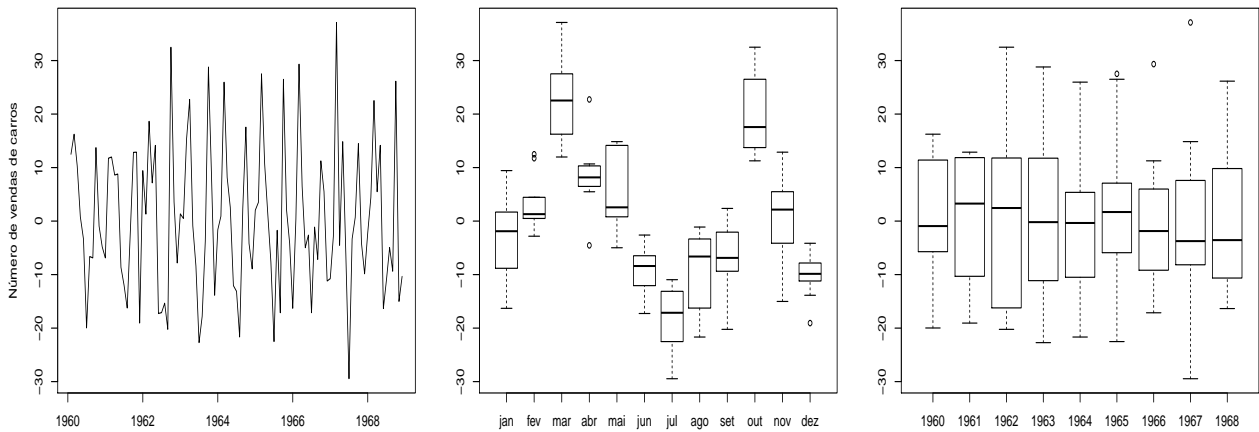


Figura 9: Séries de diferenciação da raiz quadrada das vendas mensais de carros em Quebec 1960–1968

- (g) Como discutido em todos os itens acima, a série temporal mensal das vendas mensais de carros em Quebec entre os anos de 1960–1968 demonstrou uma tendência de crescimento anual, bem como uma sazonalidade. Mesmo realizando transformações sobre a variável resposta considerando as séries de diferenciações não foi possível afirmar empiricamente que a série é estacionária. O principal motivo foi a variância, ou seja, fica claro pelos boxplots anuais e mensais que estas séries não possuem mesma variância.
3. (a) Na Figura 10 apresenta-se o gráfico da série temporal das observações mensais da variável ICV no município de São Paulo de janeiro de 1970 a junho de 1980. Pode-se verificar que existe uma tendência não linear de crescimento do índice ao decorrer do tempo. Em específico, têm-se um crescimento exponencial do ICV mensal.

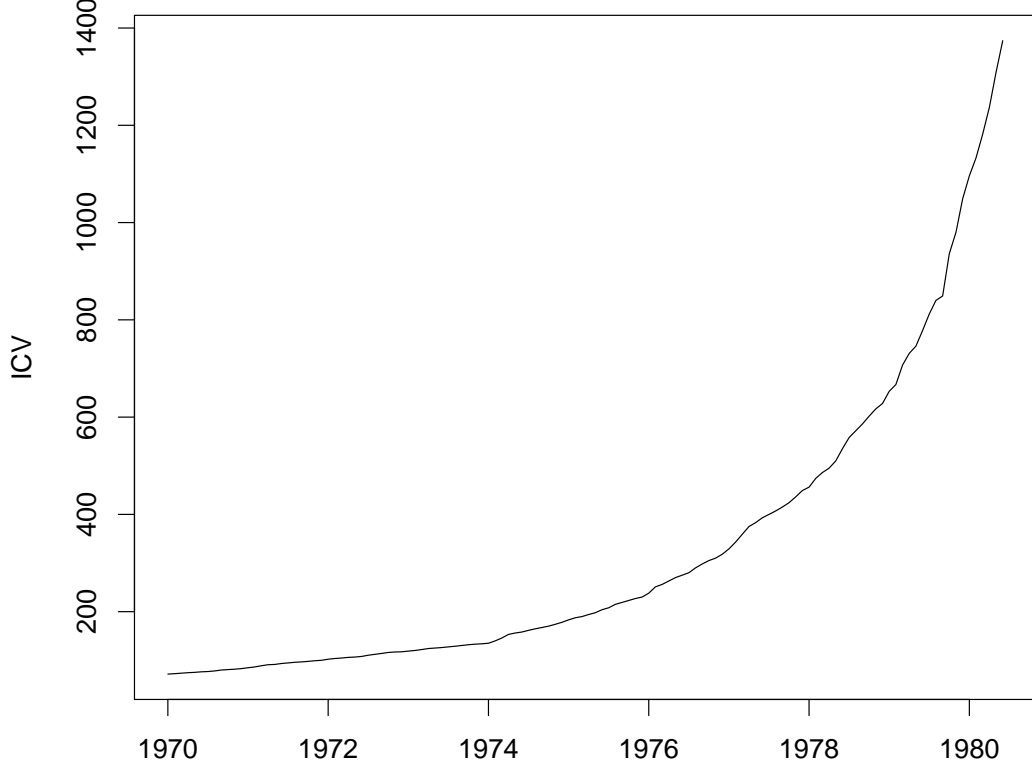


Figura 10: Índice do custo de vida (ICV) do município de São Paulo.

- (b) Para estimar a tendência dois métodos serão utilizados. O primeiro método consiste em ajustar um modelo não linear, definido por:

$$\mathbf{y} = f(\mathbf{x}, \boldsymbol{\theta}) + \boldsymbol{\varepsilon} \quad (1)$$

em que $\mathbf{y} = (y_1, \dots, y_n)^\top$ e $\mathbf{x} = (x_1, \dots, x_n)^\top$ são os vetores de variáveis resposta (ICV) e explicativa (tempo), respectivamente, $\boldsymbol{\theta}$ é o vetor de parâmetros desconhecidos, $\boldsymbol{\varepsilon} \sim N(0, 1)$ é o vetor de erros aleatórios e $f(\mathbf{x}, \boldsymbol{\theta})$ é uma função das variáveis regressoras e dos parâmetros chamada de função esperança. Tendo em vista o comportamento apresentado na Figura 10 a função esperança escolhida foi:

$$f(\mathbf{x}, \boldsymbol{\theta}) = \theta_1 \exp(\theta_2 x_i) \quad (2)$$

As estimativas de θ_1 e θ_2 foram obtidas utilizando a função `nls` do software R. Um gráfico com a curva ajustada é apresentado na Figura 11

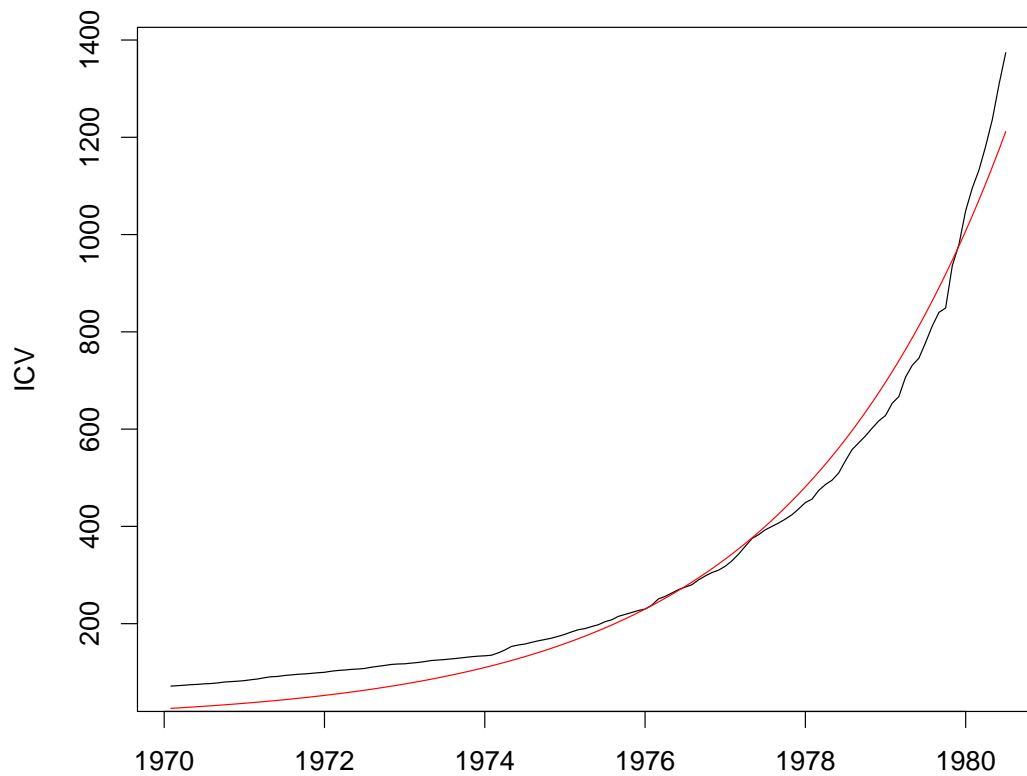


Figura 11: Curva ajustada para o índice do custo de vida (ICV) do município de São Paulo.

O segundo método utilizado para estimar a tendência consiste na diferenciação de segunda ordem da série original.

- (c) Representações gráficas das séries sem a tendência estimada são apresentadas na Figura 12.

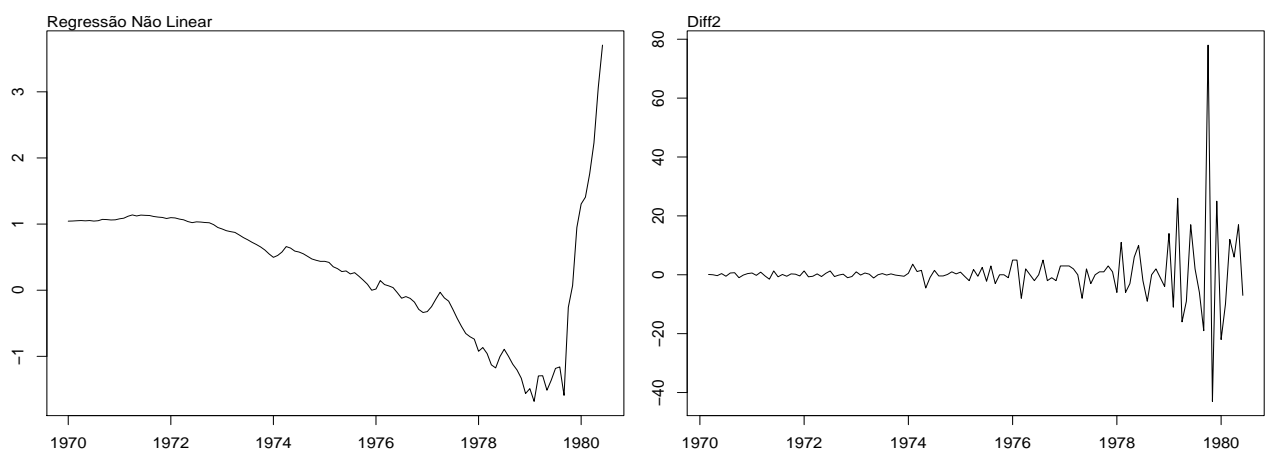


Figura 12: Séries sem tendência do índice do custo de vida (ICV) do município de São Paulo.

- (d) Nota-se que os métodos apresentaram séries bem distintas. Em particular, os resíduos produzidos pelo modelo de regressão não linear indicam que a série não é estacionária.

Por outro lado, a série baseada no método de diferenciação sucessiva indica que a mesma também não estacionária, uma vez que a variância aumenta ao longo do tempo.

4. É importante mencionar que séries temporais que exibem um comportamento que tende a se repetir a cada determinados períodos de tempo chamamos a este suposto comportamento de sazonalidade.

- (a) Na Figura 13 temos o gráfico das vendas de uma companhia X em períodos de 4 semanas sucessivas durante 1967-1970. Observa-se que o número de vendas tem um comportamento sazonal aditivo, isto é, em geral as vendas atingem um pico (máximo) na quinta (V) semana e posteriormente decrescem.

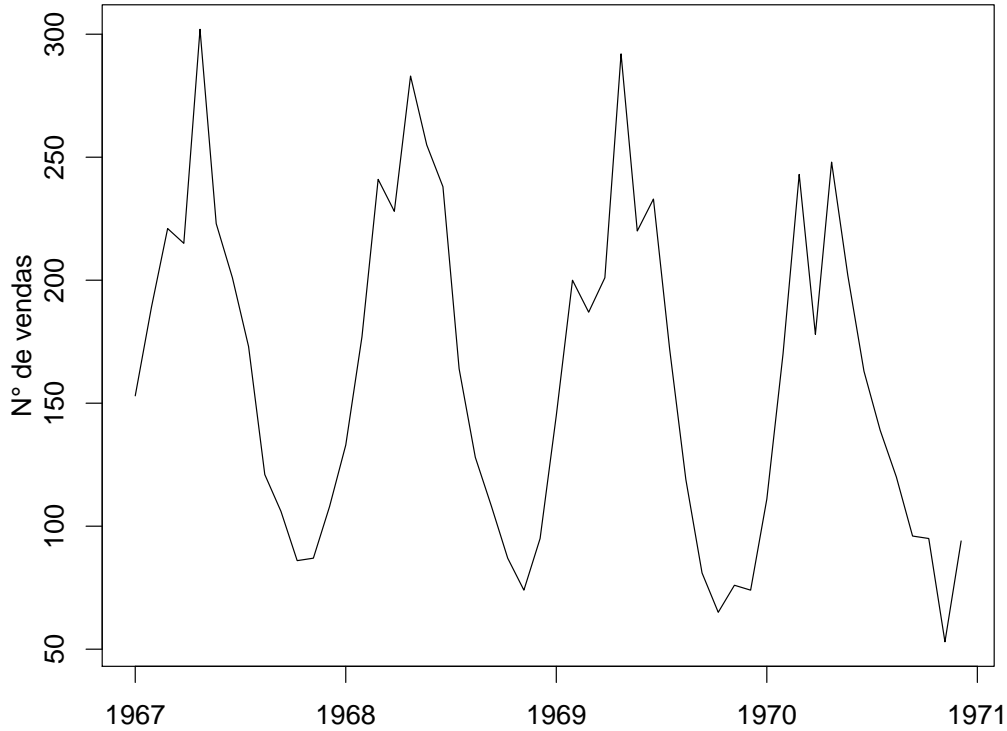


Figura 13: Vendas de uma companhia X em períodos de 4 semanas sucessivas durante 1967-1970.

- (b) Para estimar a sazonalidade bem como a tendência iremos utilizar os filtros lineares. Um filtro linear converte um série $\{x_t\}$ em outra $\{y_t\}$ através da seguinte operação linear:

$$y_t = \sum_{j=-q}^s a_j x_{t+j} \quad (3)$$

em que $\{a_t\}$ é um conjunto de pesos. Como queremos estimar a média local os pesos devem ser tais que $\sum_{j=-q}^s a_j = 1$, garantido assim que $\min x_t < y_t < \max x_t$.

O caso mais simples quando todos os pesos a_j tem o mesmo valor e devido à restrição de soma igual a um segue que $a_j = 1/(2q + 1)$, para $j = -q, \dots, q$. Assim, o valor suavizado de x_t é dado por:

$$y_t = \frac{1}{2q + 1} \sum_{j=-q}^q x_{t+j}. \quad (4)$$

Na Figura 14 temos o gráfico da série temporal juntamente com a sazonalidade estimada via médias moveis, utilizando $q = 2$ e $q = 4$.

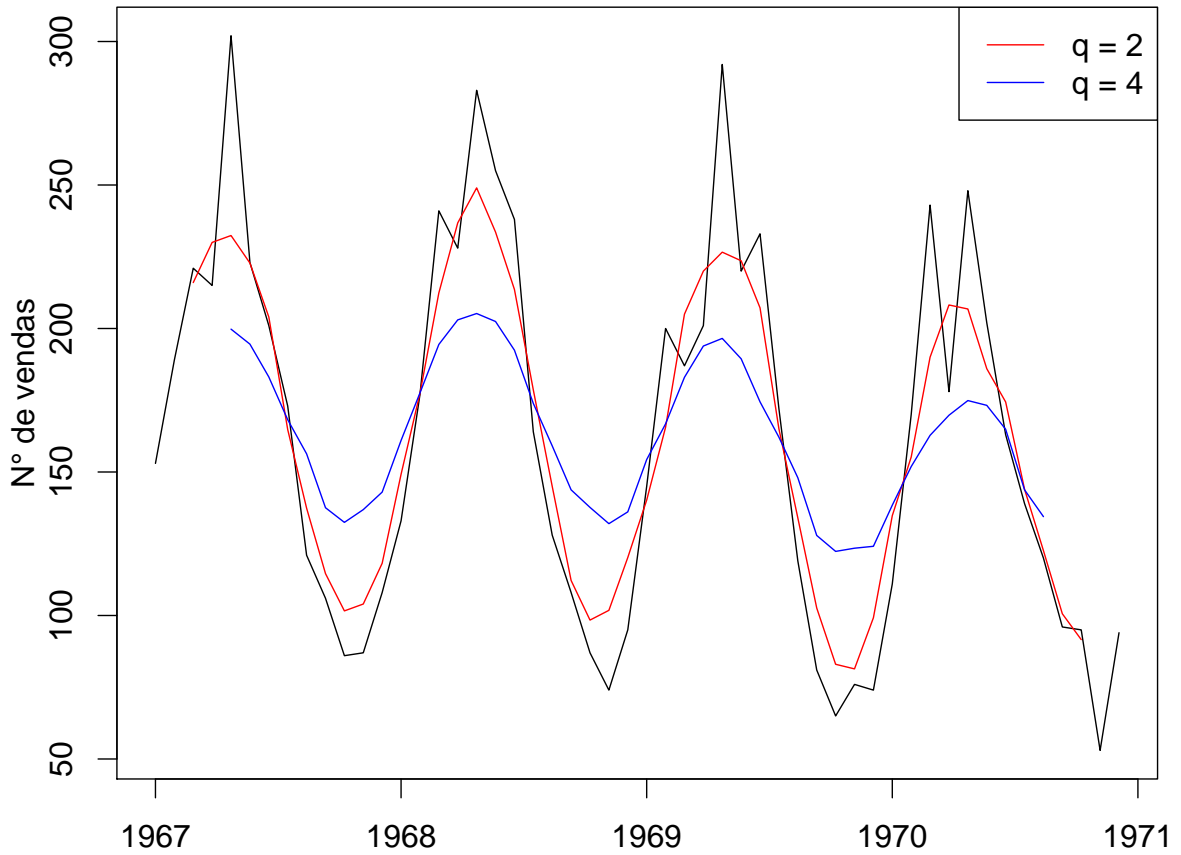


Figura 14: Vendas de uma companhia X em períodos de 4 semanas sucessivas durante 1967-1970.

Por fim, para obter uma decomposição da série, em termo de tendencia e sazonalidade o comando `decompose` foi utilizado. Na Figura 15 apresentamos separadamente a estimativa da tendência, sazonalidade e a parte aleatória que sobrou da série. Nota-se no último gráfico que a série continua não sendo estacionária, uma vez que ela não varia aleatoriamente em torno de sua média (linha pontilhada).

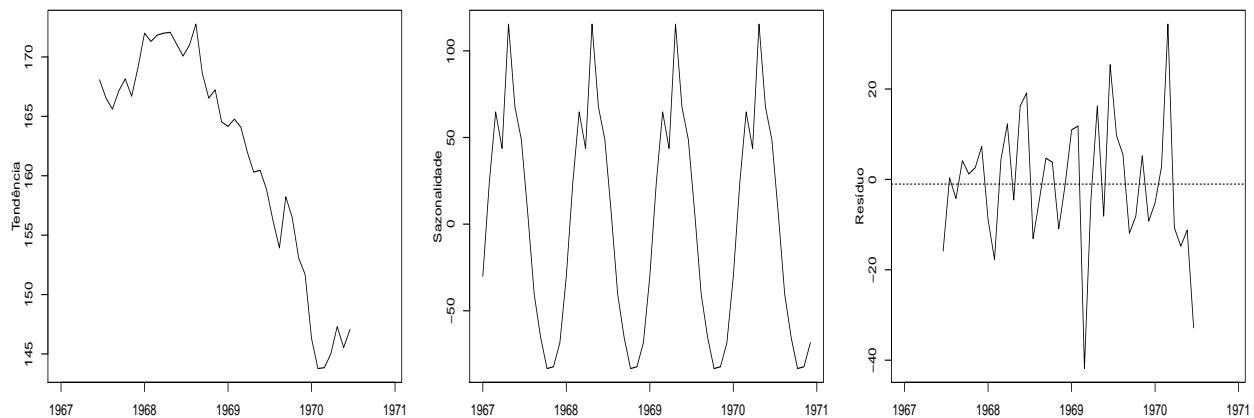


Figura 15: Tendência, sazonalidade estimada e resíduo para a série de vendas da companhia X.

Referências

Pérez, L. F., 2017. Análise de séries temporais. Universidade Federal do Paraná.
 URL <http://topespest.com.br/AST/AST.html>