Universidade Federal de São Carlos
Centro de Ciências Exatas e de Tecnologia
Departamento de Estatística

Análise de Período Sazonal COVID 19

Douglas de Paula Nestlehner
Raquel Malheiro Carvalho
Rodrigo Hideki Tozaki Moribayashi

Janeiro, 2023

Capítulo 1

1.1 Dados da Série Temporal

Os dados que compõe esta série temporal são referentes ao número de óbitos por COVID-19 no estado de São Paulo no ano de 2020. A série inicia-se em 25 de Fevereiro, quando a primeira medição foi feita, e continua por meio de medições diárias até o dia 31 de Dezembro, totalizando 311 medições.

Para melhor visualização, foi plotado primeiramente um gráfico da série temporal para que pudéssemos compreender como ela se comporta e encontrar indícios de suas principais características:

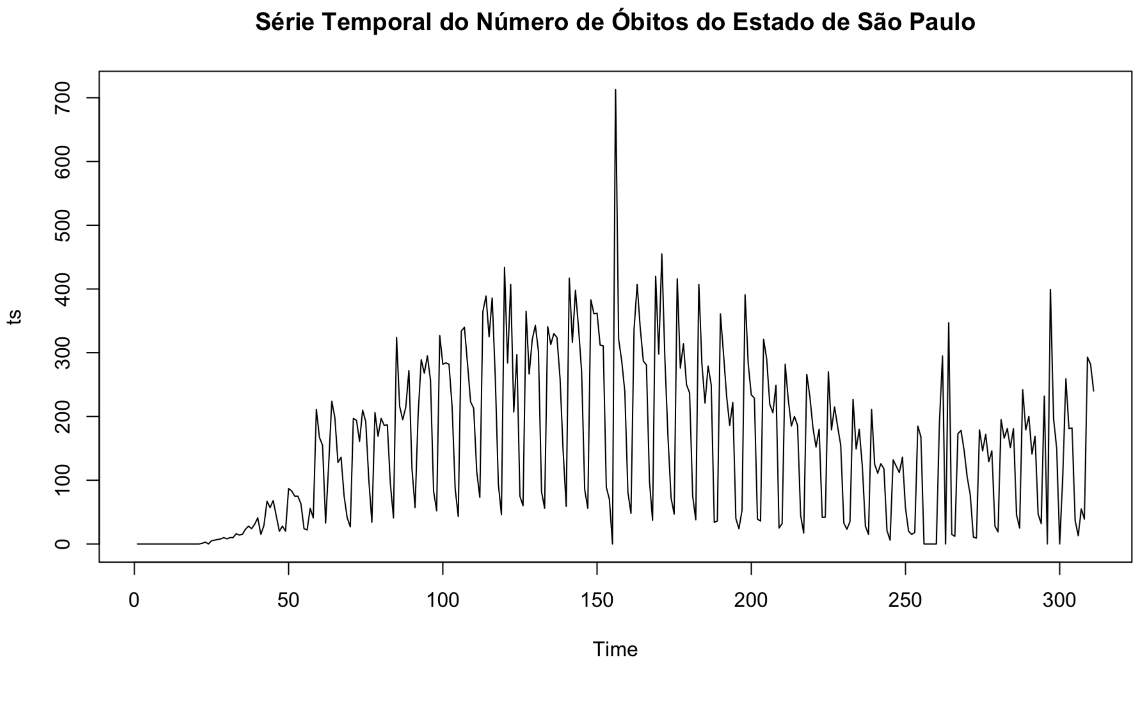

Figura 1.1: Gráfico da Série Temporal

Ao batermos o olho, temos que a série possui heterocedasticidade, ou seja, não possui variância homogênea ao longo de sua extensão. Também é possível perceber que ela possui tendência, sugerindo que não seja estacionária e que, pelo contrário, seus dados estejam variando em torno de um valor que muda ao longo do tempo. Por fim, ainda que este não esteja tão claro visualmente, é possível notar a partir dos dados um componente sazonal de 7 dias, ou seja, aos finais de semana há queda no número de óbitos registrados (vales da série).

1.2 Variabilidade

Nessa seção observamos a variabilidade da série, no intuito de verificar homocedasticidade.

1.2.1 Retirando a Variabilidade da Série

Observando a série original pudemos notar de forma clara que a variabilidade do numero de óbitos ao longo tempo não é homocedástica.

Desse modo, estabilizamos a variabilidade da série transformando os dados considerando os seguintes critérios:

$$\text{Série} = \log(\text{Série}) \tag{1.1}$$

$$\text{Série} = \sqrt{\text{Série}} \tag{1.2}$$

$$\text{Série} = \sqrt[3]{\text{Série}} \tag{1.3}$$

Na Figura 1.4, representamos a série original, a série com variabilidade estabilizada, considerando: (1.1), (1.2) e (1.3).

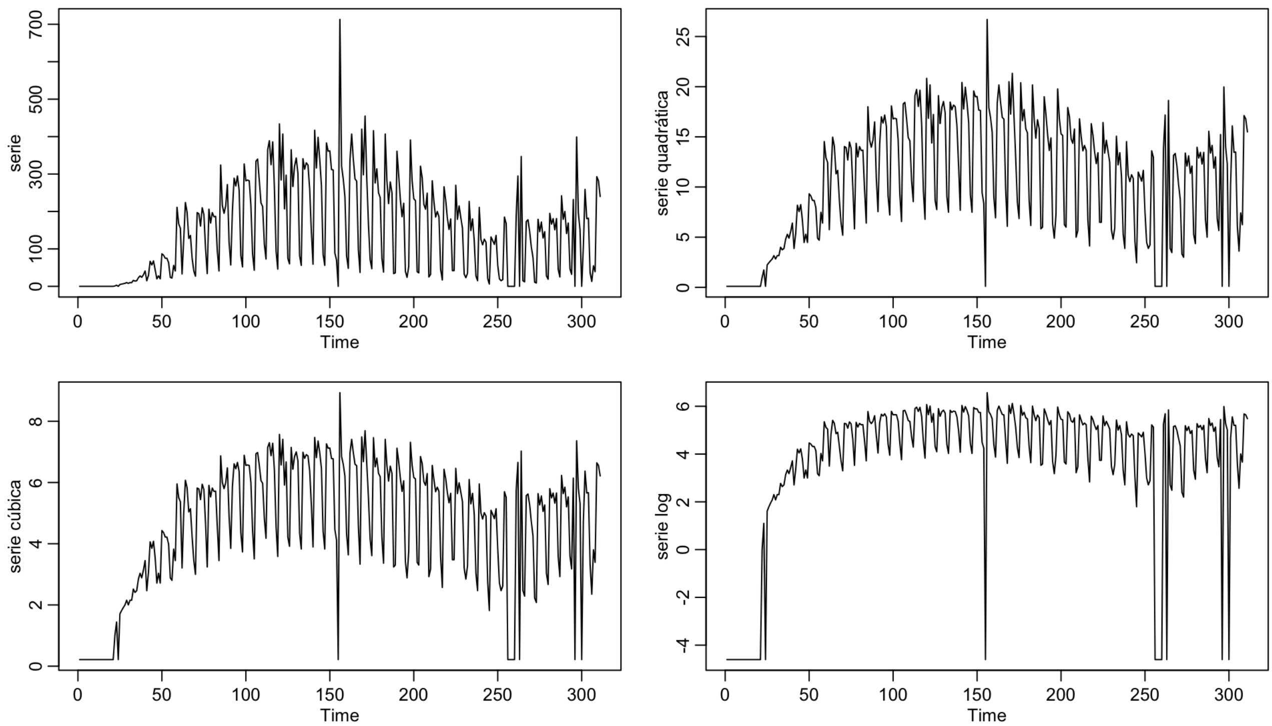

Figura 1.2: Gráfico da Série Temporal sem Variabilidade

Por meio da Figura 1.4 definimos que a transformação que aparenta estabilizar melhor a série, e a série que considera a transformação quadrática. Desse modo, para continuidade, utilizaremos a série quadrática.

1.3 Tendência

Obtendo na seção anterior, a série com variabilidade "estavel", estudamos a presença ou não de tendencia.

1.3.1 Retirando a Tendência

Observamos que a série tem tendencia, desse modo, teremos como interesse remover a tendencia para continuidade da atividade.

Para isso, utilizamos o método da diferença, no R utilizamos o comando "diff(série_v)", obtendo então a série com variabilidade estabilizada e sem tendência, representada na Figura 1.3.

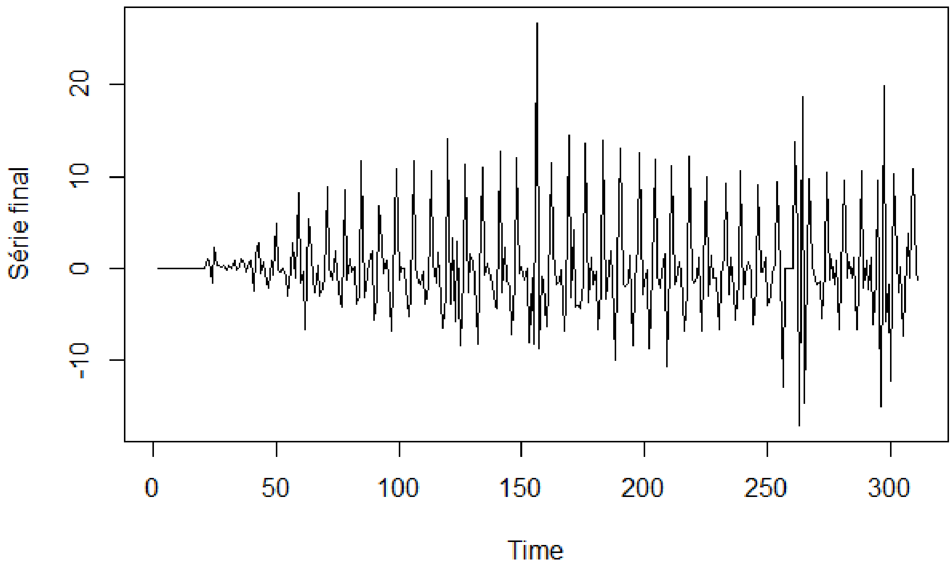

Figura 1.3: Gráfico da série temporal com variabilidade estável e sem tendencia

1.4 Sazonalidade

A partir da série com variabilidade estável e sem tendencia, estudamos a sazonalidade pelo método de fisher.

1.4.1 Verificando a presença da Sazonalidade

O Teste de Fisher verifica a presença de Sazonalidade na série por meio de um teste de hipóteses, em que a a hipótese nula (H_0) é a ausência de sazonalidade e a alternativa é sua presença (H_1). O cálculo da Estatística do Teste depende dos valores dos resíduos do modelo ajustado para a série (lembrando que a série utilizada agora é aquela que a variabilidade e a tendência foram removidas).

Assim como nos materiais de aula, o modelo ajustado nesse caso foi polinomial, em que decidimos por um polinônio de grau 3. Usando este número o modelo parece capturar as variações médias do número de óbitos que acontecem na série, levando em conta ainda a busca de um modelo parcimonioso (polinômios de graus 4,5,6,7 e 8 não apresentaram uma curva visualmente diferente da ajustada com grau 3).

A estatística do Teste, denominada Estatística g de Fisher, é calculada a seguir:

$$\begin{aligned} g = \frac{max(I_p)}{\sum_{p=1}^{311} I_p} = \\ = \frac{8.49}{60.16} = \\ = 0.14 \end{aligned}$$

O valor-p obtido é muito próximo de zero ($= 1.165745e^{-18}$). Portanto rejeitamos a hipótese nula da não presença de sazonalidade ao nível de significância $\alpha = 0.05$, ou seja, há indícios de que a séria apresenta tendência.

1.4.2 Encontrando o período sazonal

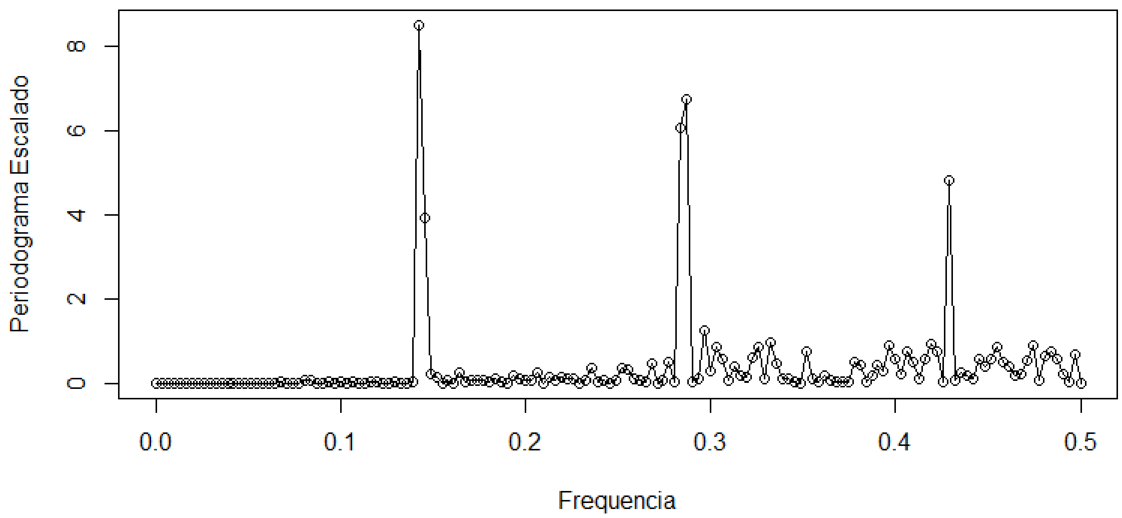

Figura 1.4: Periodograma da série com variabilidade ajustada e sem tendência.

Nota-se que o maior valor da ordenada é para a frequência de 0.14, ou seja, cada dia representa 14% de um ciclo completo. Em outras palavras, um ciclo completo ocorre a cada

$\frac{1}{0.14} = 7.04$ dias (período sazonal de 7.04 dias), como é possível observar a partir dos dados da série original.