



## Modelagem matemática da Covid-19 usando SIR

Marcelo Osnar Rodrigues de Abreu – Email: [osnar@outlook.com](mailto:osnar@outlook.com)

**Resumo:** Exploramos o modelo epidemiológico SIR (Suscetíveis-Infetados-Recuperados) a fim de compreender o comportamento da pandemia de Covid-19 e discutimos quais medidas podem e porque podem auxiliar no seu enfrentamento. São apresentados também os algoritmos na linguagem R para a implementação do Modelo SIR em situações reais. Além disso, ao longo do texto são mencionados os métodos matemáticos que estão por trás destes algoritmos.

**Palavras-chave:** taxa de propagação, taxa de recuperação, mínimos quadrados, Runge-Kutta, EpiDynamics.

### 1. Introdução

O modelo SIR foi proposto por [Kermack e McKendrick \(1927\)](#), onde os autores apresentam como o término de uma epidemia pode resultar de uma relação particular entre a densidade populacional e as taxas de transmissão da doença, de recuperação e de mortalidade.

Dentre as muitas variáveis que influenciam na propagação de uma epidemia, quando estamos considerando uma mesma população (ou populações comparáveis conforme citação abaixo), a taxa de infecção (ou taxa de propagação) se torna a principal variável a se determinar e as ações de enfrentamento à epidemia que visem reduzir esta taxa se torna uma das principais, senão a principal, medida para reduzir o número de infectados

[...] se considerarmos duas populações idênticas em relação às suas densidades, suas taxas de recuperação e mortalidade, mas diferentes em relação às suas taxas de infectividade, parecerá que as epidemias na população com a maior taxa de infectividade podem ser maiores em comparação com as da população com menor taxa de infectividade [...] ([KERMAK; MCKENDRICK, 1927](#)).

No decorrer deste trabalho vamos explicar o modelo SIR e explorar os parâmetros necessários para aplicar o modelo. Veremos como o sistema de equações que governam a modelagem SIR podem nos trazer informações sobre o comportamento da propagação da epidemia.

Implementações computacionais deste modelo podem servir de material didático a ser utilizados em disciplinas como introdução às equações diferenciais ordinárias e cálculo numérico.

No entanto, este modelo possui algumas limitações e é necessário um conhecimento mais profundo da situação problema na qual se pretende utilizá-lo. Uma destas limitações é o fato do modelo trabalhar com parâmetros constantes tornando-o pouco eficaz para o estudo de longo prazo de situações dinâmicas.

Os algoritmos são implementados na linguagem de programação R. R é uma linguagem multi-paradigma voltada principalmente para análise e visualização de dados.

## 2. Modelo SIR

Vamos descrever o modelo SIR na sua versão mais simples, sem considerar a taxa de nascimento e a taxa de mortalidade<sup>1</sup>, e com taxas de infecção  $\beta$  e de recuperação  $\gamma$  constantes. De acordo com [Kermack e McKendrick \(1927\)](#), muita informação pode ser obtida sobre o processo pelo qual epidemias seguem seus cursos ao se considerar  $\beta$  e  $\gamma$  constantes.

### 2.1. Definição

O modelo divide a população em três grupos, são eles:

**Figura 1** – Os três grupos populacionais do modelo SIR



**Fonte:** elaborada pelo autor.

- **Suscetíveis:** São todas as pessoas que podem, em algum momento, ser infectadas.
- **Infectados:** Este grupo é formado pelas pessoas que estão infectadas (também chamados de infecciosos) e, portanto, podem ser capazes de transmitir a infecção.
- **Recuperados:** Aqui temos que tomar um cuidado com a nomenclatura. Ao se usar a palavra recuperados dá-se a entender que as pessoas deste grupo se recuperaram da infecção. No entanto, o modelo, na versão apresentada aqui, visa estudar o comportamento da infecção e neste sentido toda pessoa que não é mais capaz de transmitir

<sup>1</sup>Desconsideramos a taxa de mortalidade, mas o número de mortes faz parte do modelo como veremos a seguir.

a infecção entra neste grupo, ou seja, o grupo dos recuperados é formado pelas pessoas que se recuperaram da infecção e também por pessoas que venham a falecer em decorrência dela.

- Todas as pessoas da população onde é realizado o estudo pertence a um dos três grupos acima e nenhuma pessoa pode pertencer, ao mesmo tempo, a dois ou mais grupos.

Uma epidemia pode se iniciar quando uma (ou mais) pessoa infectada é introduzida em uma comunidade de indivíduos suscetíveis à doença em questão. A doença se espalha do infectado para o suscetível por contato. Cada pessoa infectada passa pelo curso de sua doença e, finalmente, é removida do número de pessoas doentes, por recuperação ou por morte (KERMACK; MCKENDRICK, 1927).

Para simplificar a modelagem são feitas algumas considerações:

- O tamanho da população é constante. Não será considerado taxa de nascimento nem de mortalidade. Considerando uma população grande, em geral, nascimentos e mortes em um curto período não modificarão significativamente o tamanho da população. Então, caso não se tenha acesso a valores precisos para estas taxas o mais prudente é não as utilizar na modelagem.
- A interação entre as populações dos compartimentos se dá de forma homogênea.
- Após se recuperar o indivíduo se torna imune e não é capaz de infectar outras pessoas.
- O contágio se dá através do contato efetivo entre suscetíveis e infectados.

Novas infecções ocorrem como resultado do contato entre infectados e suscetíveis. Neste modelo a taxa na qual novas infecções ocorrem, taxa de transmissão ( $\beta$ ), é uma constante positiva<sup>2</sup>. Quando uma nova infecção ocorre, o indivíduo infectado passa do grupo das pessoas suscetíveis para o grupo dos infectados. Estamos supondo que não há outra maneira dos indivíduos entrarem ou saírem do grupo dos suscetíveis; portanto, temos a nossa primeira equação diferencial:

$$\frac{dS}{dt} = -\beta \cdot S \cdot I, \quad (1)$$

na qual

- $\frac{dS}{dt}$ : denota a taxa de variação  $S^3$  (quantidade de pessoas suscetíveis) em função do tempo  $t$ .

---

<sup>2</sup>No decorrer deste trabalho vamos explorar em mais detalhes esta taxa.

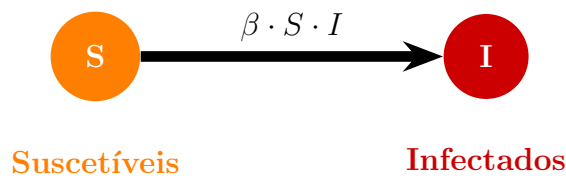
<sup>3</sup> $S$  denota uma função da variável  $t = \text{tempo}$ . Como estamos interessados no comportamento da infecção ao longo do tempo e, em caso de epidemia, a quantidade  $S$  de pessoas suscetíveis varia em função de  $t$ , é apropriado considerar  $S$  como uma função de  $t$ .

- $\beta$ : é a taxa de propagação da infecção. Esta constante é obtida pelo produto da taxa de contato entre pessoas pela probabilidade de uma pessoa suscetível ser contaminada ao ter contato com uma pessoa do grupo dos infectados.
- $S$ : quantidade de pessoas suscetíveis.
- $I$ : quantidade de pessoas do grupo infectados.

Vamos entender por que o produto  $-\beta \cdot S \cdot I$  representa a taxa de variação da população  $S$ . Primeiramente o que seria taxa de variação de  $S$ ? Esta taxa mede quão rápido a população do grupo  $S$  aumenta ou diminui. Como estamos simulando um processo de propagação de uma infecção, esta taxa deve ser negativa; uma vez que uma pessoa passa de suscetíveis para infectados a população em  $S$  diminui, isto justifica o sinal negativo no membro à direita da equação (1). Suponha inicialmente que  $I = 1$ , ou seja, há apenas uma pessoa infectada. Então, a variação na população de suscetíveis é dada pela probabilidade das pessoas suscetíveis se infectarem; tal probabilidade é dada pelo produto de  $\beta$  (taxa de propagação) por  $S$  (quantidade de pessoas suscetíveis). No caso geral, isto é, para um número arbitrário de  $I$ , a taxa de variação de  $S$  é  $-\beta \cdot S \cdot I$ .

A equação (1) descreve matematicamente a seguinte dinâmica populacional: pessoas passam do grupo suscetíveis para o grupo infectados.

**Figura 2** – Pessoas suscetíveis passam para o grupo dos infectados de acordo com a taxa  $\frac{dS}{dt}$



**Fonte:** elaborada pelo autor.

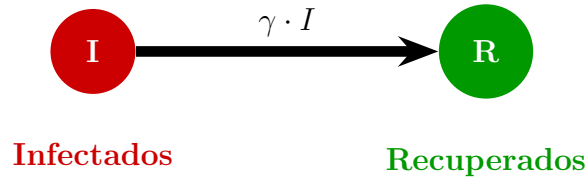
Uma outra dinâmica populacional é esperada: indivíduos infectados passam para o grupo dos recuperados. Supondo que isso aconteça com uma taxa de recuperação ( $\gamma$ ) podemos descrever esta dinâmica por meio da equação diferencial:

$$\frac{dR}{dt} = \gamma \cdot I. \quad (2)$$

A taxa de recuperação  $\gamma$  está relacionada com o tempo médio de recuperação de uma pessoa infectada. Mais precisamente,  $\gamma = \frac{1}{t_r}$ , onde  $t_r$  denota o tempo médio de recuperação de uma pessoa infectada.

Agora, a taxa de variação da população dos recuperados é dada pelo produto de  $\gamma$  (taxa de recuperação) por  $I$  (quantidade de infectados).

**Figura 3** – Pessoas infectadas passam para o grupo dos recuperados de acordo com a taxa  $\frac{dR}{dt}$

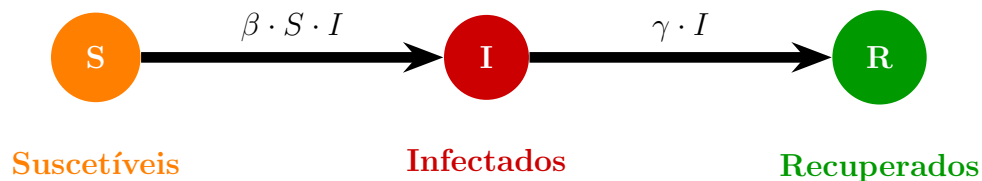


**Fonte:** elaborada pelo autor.

Por fim, precisamos descrever a taxa de variação de  $I$  (infectados): a variação dos infectados depende de quantas pessoas entram neste grupo (pessoas suscetíveis que passam para infectados), este número deve ser positivo, menos a quantidade de pessoas que saem do grupo infectados (pessoas do grupo infectados que passam para recuperados). Assim,

$$\frac{dI}{dt} = \beta \cdot S \cdot I - \gamma \cdot I. \quad (3)$$

**Figura 4** – Pessoas suscetíveis passam para o grupo dos infectados de acordo com a taxa  $\frac{dS}{dt}$



**Fonte:** elaborada pelo autor.

Considerando um estágio de tempo inicial  $t = t_0$  e as populações iniciais de cada grupo dadas por  $S(t_0) = S_0$ ,  $I(t_0) = I_0$  e  $R(t_0) = R_0$  obtemos o sistema de equações diferenciais

$$\text{Modelo SIR} = \begin{cases} \frac{dS}{dt} = -\beta \cdot S \cdot I \\ \frac{dI}{dt} = \beta \cdot S \cdot I - \gamma \cdot I \\ \frac{dR}{dt} = \gamma \cdot I \\ S(t_0) = S_0, I(t_0) = I_0, R(t_0) = R_0. \end{cases} \quad (4)$$

Uma solução do sistema de equações diferenciais dado em (4) em um intervalo  $I \ni t_0$  é constituída por funções  $S(t)$ ,  $I(t)$  e  $R(t)$  que são diferenciáveis em  $I$  e que satisfazem simultaneamente todas as equações para todo  $t \in I$ . Estudos sobre soluções exatas do sistema 4 podem ser encontradas em (SHABBIR; KHAN; SADIQ, 2010) e em (BOHNER; STREIPERT; TORRES, 2019).

## 2.2. Informações iniciais sobre a propagação da infecção

A partir das equações diferenciais (1), (2) e (3) que descrevem a dinâmica populacional é possível extrair algumas informações sobre o possível comportamento da propagação da epidemia, e até mesmo saber se haverá ou não uma epidemia.

Primeiramente note que se  $I = 0$ , ou seja, não houver pessoas no grupo infectados, então as equações (1), (2) e (3) são nulas, isto é, nenhuma pessoa muda de grupo. Da mesma forma, se  $S = 0$ , ou seja, não há pessoas suscetíveis, então não haverá aumento dos infectados.

Para tornar as coisas interessantes, no sentido matemático, precisamos começar com  $S > 0$  e  $I > 0$ .

Observe que estamos supondo que  $S + I + R$  contempla a população total. Como

$$\frac{dS}{dt} + \frac{dI}{dt} + \frac{dR}{dt} = -\beta \cdot S \cdot I + (\beta \cdot S \cdot I - \gamma \cdot I) + \gamma \cdot I = 0, \quad (5)$$

então  $S + I + R = 0$ , ou seja, a população total é constante. Por conveniência consideramos a população total igual a 1, isto é,  $S + I + R = 1$ .

No estágio inicial é aceitável assumir que  $S \approx 1$ . Vamos usar este fato para ampliar nosso estudo.

De (3), assumindo  $S = 1$ , obtemos

$$\frac{dI}{dt} = \beta \cdot I - \gamma \cdot I = (\beta - \gamma) \cdot I. \quad (6)$$

A equação (6) é uma equação diferencial ordinária linear homogênea de ordem 1 cuja solução é dada por (KEELING; ROHANI, 2011, p. 50)

$$I(t) \approx I_0 e^{(\beta - \gamma)t}. \quad (7)$$

A equação (7) nos diz que o comportamento inicial da função que descreve o número de infectados ao longo do tempo é uma função exponencial. Sendo  $I_0 > 0$ , podemos concluir, pelo comportamento da função exponencial, que o grupo dos infectados irá aumentar se, e somente se,  $\beta > \gamma$ , ou seja, se a taxa de infecção ( $\beta$ ) for maior do que a taxa de recuperação ( $\gamma$ ), então a infecção se espalhará e quanto maior for a diferença  $\beta - \gamma$  mais rápido isto ocorrerá.

Em outras palavras, se queremos mitigar a infecção ou até mesmo suprimi-la é necessário buscar meios para que a diferença  $\beta - \gamma$  seja a menor possível, preferencialmente negativa.

A taxa de recuperação da infecção tende a ser mais previsível e difícil de modificar (em um curto espaço de tempo). Neste sentido, em geral, os esforços tendem a ser no sentido de diminuir a taxa de infecção ( $\beta$ ).

Lembremos que  $\beta$  é proporcional à taxa de contato entre pessoas e à probabilidade de uma pessoa suscetível ser contaminada ao ter contato com uma pessoa do grupo dos infectados. Desta forma, a quarentena pode ser uma boa aliada:

- Isolar pessoas contaminadas reduzirá a probabilidade de uma pessoa suscetível ter contato com uma pessoa do grupo dos infectados.
- Isolar apenas pessoas contaminadas pode não ser o suficiente por vários motivos: primeiro que o isolamento completo pode não ocorrer caso a pessoa fique em uma residência com outras pessoas não infectadas. Caso a pessoa fique no hospital poderá ter contato com profissionais de saúde. Temos também o fato de muitas pessoas serem assintomáticas o que impossibilita identificar todos que precisam ser isolados.
- Pelo item anterior surge a necessidade de reduzir a mobilidade das pessoas em geral; esta medida tem o poder de diminuir a taxa de infecção uma vez que diminuirá a taxa de contato.

Sejam  $t$  =tempo (instantâneo) e  $\Delta t$  um intervalo de tempo suficientemente pequeno, então a taxa de variação dos infectados dada pela equação (3) pode ser aproximada por

$$\frac{I(t + \Delta t) - I(t)}{\Delta t}.$$

Logo,

$$\begin{aligned} \beta \cdot S \cdot I - \gamma \cdot I &= \frac{dI}{dt} \approx \frac{I(t + \Delta t) - I(t)}{\Delta t} \\ I(t + \Delta t) &= I(t) + \Delta t \cdot \beta \cdot S \cdot I - \Delta t \cdot \gamma \cdot I. \end{aligned} \quad (8)$$

Considere agora  $t = 0$  e  $\Delta t = t_r$ , em que  $t_r = \frac{1}{\gamma}$  denota o tempo médio para um indivíduo infectado passar para o grupo dos recuperados. Supondo que no instante  $t = 0$  (início da infecção) o número de pessoas suscetíveis é aproximadamente o total da população, isto é,  $S(0) = S_0 \approx 1$ , então de (8) segue que

$$\begin{aligned} I(0 + t_r) &= I(0) + t_r \cdot \beta \cdot 1 \cdot I - t_r \cdot \gamma \cdot I \\ I(t_r) &= I(0) + \frac{1}{\gamma} \cdot \beta \cdot 1 \cdot I(0) - \frac{1}{\gamma} \cdot \gamma \cdot I(0) \\ I(t_r) &= \frac{\beta}{\gamma} I(0). \end{aligned} \quad (9)$$

A equação (9) nos diz que a quantidade de infectados após o intervalo de tempo  $\Delta t = t_r$  é dada por  $\frac{\beta}{\gamma}$  vezes  $I(0)$ . Em outras palavras, supondo que a quantidade de indivíduos contaminados no estágio inicial  $I(0)$  corresponde a exatamente um indivíduo da população total  $S + I + R$ , então a quantidade de novos infectados produzidos é exatamente  $\frac{\beta}{\gamma}$ .

Ou seja, podemos afirmar que um indivíduo infectado transmite a doença para  $\frac{\beta}{\gamma}$  outros indivíduos durante o período em que ele permanece infectado.

Da biologia temos a definição de um número usualmente denotado por  $R_0$

O número básico de reprodução,  $R_0$ , é definido como o número esperado de casos secundários produzido por uma única infecção em uma população completamente suscetível. É importante notar que  $R_0$  é um número sem dimensão e não uma taxa, que teria unidades de tempo<sup>-1</sup> (JONES, 2007).

Desta forma, podemos afirmar, pelo Modelo SIR, que  $R_0 = \frac{\beta}{\gamma}$  é o número médio de pessoas suscetíveis que cada indivíduo infectado irá contaminar.

Temos três situações possíveis:

1.  $R_0 = 1$  ( $\beta = \gamma$ ): temos uma situação de equilíbrio, isto é, o número de infectados na população permanecerá igual.
2.  $R_0 > 1$  ( $\beta > \gamma$ ): o número de casos aumentará.



3.  $R_0 < 1$  ( $\beta < \gamma$ ): o número de casos diminuirá.

É importante destacar que conhecer apenas o  $R_0$  não é suficiente para descrever o comportamento da propagação da infecção

Na estrutura discutida dos modelos simples de SIR e SEIR, concluímos que apenas o número reprodutivo básico fornece poucas informações sobre a duração, tempo de geração, velocidade da epidemia e tempo geral de uma doença infecciosa como a gripe. Em vez disso, propomos que os valores dos parâmetros individuais sejam mais críticos para a compreensão da dinâmica da doença e possam ser mais valiosos para as autoridades políticas na montagem de uma resposta eficaz à saúde pública (RIDENHOUR; KOWALIK; SHAY, 2018).

Retomando nossas análises, dividindo a equação (1) pela equação (2) obtemos

$$\frac{dS}{dR} = -\frac{\beta}{\gamma} \cdot S = -R_0 \cdot S, \quad (10)$$

que é uma equação diferencial linear homogênea cuja solução é

$$S(t) = cS_0 e^{-R_0 \cdot R(t)}. \quad (11)$$

Pela equação (11) podemos concluir que  $S(t) > 0$  para todo valor de  $t$ , ou seja, a população de suscetíveis nunca se esgota durante a epidemia, ou seja, quando uma epidemia se extingue é pela diminuição dos infectados, não pela falta de suscetíveis.

Da equação (3) podemos dizer que a infecção se espalhará se  $\frac{dI}{dt} > 0$ , ou seja,  $\beta \cdot S \cdot I - \gamma \cdot I > 0$ , ou equivalentemente,  $S > \frac{\gamma}{\beta} = \frac{1}{R_0}$ . Esta informação pode ser útil, por exemplo, para se determinar qual fração da população suscetível deve ser vacinada para que não haja mais epidemia. Vejamos como isto funciona.

Vamos denotar por  $R_e$  o número efetivo de casos secundários em uma população não totalmente suscetível, produzidos por um indivíduo infectado. Sendo  $S^*$  a quantidade de indivíduos suscetíveis na população, temos que

$$R_e = S^* \cdot R_0. \quad (12)$$

Observe que se  $S^* = 1$ , ou seja, toda a população é suscetível, então  $R_e = R_0$ . No estado de equilíbrio, ou seja, o número de infectados não aumenta e nem diminui, devemos ter  $R_e = 1$ , isto é,  $S^* = \frac{1}{R_0}$ .

Supondo que toda a população seja suscetível, seja  $V$  a fração da população que foi vacinada, então, no máximo, teremos  $S^* = 1 - V$  (população total menos a quantidade de

vacinados). Assim,  $R_e = (1 - V) \cdot R_0$ . Para controlar a transmissão, devemos ter  $R_e \leq 1$ , isto é,

$$\begin{aligned} R_e = (1 - V) \cdot R_0 &\leq 1 \\ V &\geq 1 - \frac{1}{R_0}. \end{aligned} \quad (13)$$

Vamos exemplificar. Suponha que  $R_0 = 2$ , ou seja, cada pessoa infectada contamina outras duas. Então, para controlar a epidemia, apenas com vacinação, será necessário vacinar ao menos 50% da população.

### 3. Como mitigar a pandemia de Covid-19?

Nesta seção vamos avaliar cada uma das variáveis do modelo SIR com o objetivo de compreender quais medidas podem mitigar a pandemia de Covid-19.

Cabe aqui ressaltar que esta análise é feita baseando-se nas propriedades do modelo SIR apresentado anteriormente e limita-se ao sentido de estimar a quantidade de casos da COVID-19. Qualquer política de combate à epidemia deve olhar não somente para o número de casos ou a velocidade de propagação da doença. É necessário analisar qual o impacto que este número de casos fará na sociedade tais como taxa de hospitalização e de mortalidade. Além disso, é necessário se atentar ao impacto econômico das medidas a serem tomadas

Mesmo que as taxas de ataque precisas pudessem ser medidas a partir das estimativas de  $R_0$ , muitas das questões mais críticas de saúde pública permaneceriam sem resposta. Por exemplo, as taxas de fatalidade e hospitalização de casos são essencialmente independentes da dinâmica da doença (e, portanto, do que pode ser derivado das estimativas de  $R_0$ ), mas são os principais determinantes da mortalidade durante um surto de doença infecciosa. A pandemia de H1N1 de 2009 ilustra apropriadamente esse ponto: altas taxas de ataque não produziram níveis correspondentemente altos de mortalidade devido à gravidade relativamente leve da infecção e às baixas taxas de ataque entre as populações tipicamente com maior risco de complicações graves da influenza, principalmente adultos com mais de 65 anos ([RIDENHOUR; KOWALIK; SHAY, 2018](#)).

Tais aspectos não estão sendo ponderados nas conclusões a seguir.

Outro ponto que gostaríamos de destacar é o seguinte: se faz necessário avaliar a evolução da epidemia quase em tempo real para ter um suporte científico mais confiável para a tomada de decisões

Se existissem mecanismos para estimar valores em tempo real ou quase em

tempo real de  $R_e > 1$ , as autoridades de saúde pública poderiam determinar se estratégias de intervenção específicas – como fechamento ou quarentena de escolas – estão sendo efetivas para alterar a dinâmica da doença e se essas intervenções deveriam ser mantidas. A comparação dos números reprodutivos antes e depois das intervenções em saúde pública foi conduzida retrospectivamente para a pandemia de influenza de 1918 (BOOTSMA; FERGUSON, 2007), mas é improvável que a eficácia de uma intervenção medida em epidemias passadas se aplique a uma epidemia contemporânea (devido a mudanças na estrutura social, ambiente, patógenos, etc.). Devemos reconhecer que, embora a capacidade de estimar números reprodutivos em tempo real possa ser vantajosa, a eficácia de uma intervenção específica pode variar temporal e geograficamente devido a mudanças no ambiente, estrutura da população, evolução viral e imunidade; portanto, estimativas feitas em uma região podem não ser aplicáveis a outra (RIDENHOUR; KOWALIK; SHAY, 2018).

De acordo com o Modelo SIR, os parâmetros necessários são:  $S$ ,  $I$ ,  $R$ ,  $\beta$  e  $\gamma$ .

Pela análise matemática do modelo, feita na seção anterior, podemos discutir as possíveis medidas a serem tomadas.

### 3.1. Suscetíveis $S$

Pela equação (1) é possível ver que quanto maior for a população, maior será a taxa de decréscimo das pessoas suscetíveis, ou seja, o número de infectados tende a crescer mais rapidamente. Neste sentido, uma medida efetiva é reduzir o número de suscetíveis.

Mas como fazer isto?

Vacinar parte (ou toda) da população suscetível seria eficaz. Mas no caso da Covid-19 esta solução ainda não está disponível.

Outra possibilidade é a quarentena para pessoas suscetíveis. Se uma pessoa fica em isolamento social, podemos dizer que ela não faz mais parte do conjunto das pessoas suscetíveis (existem modelos epidemiológicos que consideram esta situação, veja Berger, Herkenhoff e Mongey (2020)). É como se pudéssemos saltar do grupo dos suscetíveis diretamente para o grupo recuperados.

### 3.2. Infectados $I$

Uma medida em relação aos infectados é o isolamento social, uma vez que é este grupo que propaga a doença. No entanto, muitas pessoas são assintomáticas o que dificulta a identificação das pessoas deste grupo.

Além disso, um isolamento total, ou seja, quando o infectado não tem contato com nenhuma outra pessoa que não esteja devidamente protegida é muito difícil (por inúmeros

motivos que não vamos descrever aqui) de ser colocado em prática. Este fato reforça a necessidade da quarentena geral (para todos os grupos).

Por fim, segundo He et al. (2020), mais de dois quintos das transmissões de Covid-19 ocorrem antes que a pessoa infectada apresente sintomas.

### 3.3. Recuperados $R$

Aumentar o número de recuperados, ou a taxa de recuperação<sup>4</sup>, é algo muito difícil de se fazer uma vez que não há tratamento específico e/ou vacina, até a presente data, para a Covid-19.

### 3.4. Taxa de propagação $\beta$

Quanto menor for a taxa de propagação menores serão os efeitos da epidemia e torná-la menor ou igual à taxa de recuperação é a única maneira de interromper a epidemia.

A taxa de propagação  $\beta$  depende diretamente da taxa de contato entre pessoas e da probabilidade de uma pessoa suscetível se infectar ao entrar em contato com uma pessoa infectada.

Várias medidas podem ser tomadas para diminuir esta taxa:

- isolamento social: isolamento social implica em menor taxa de contato entre pessoas e conseqüentemente irá reduzir a taxa de propagação  $\beta$ .
- evitar aglomerações: medidas para reduzir o número de pessoas que estão ao mesmo tempo em um ambiente fechado e espaçamento de dois metros entre as pessoas são exemplos de medidas que evitam aglomerações e conseqüentemente a taxa de contato entre pessoas.
- higienização: lavar as mãos com água e sabão, usar álcool 70% dentre outras reduzem a probabilidade de contágio de uma pessoa suscetível ao ter contato com uma pessoa infectada (ou um ambiente onde, recentemente, esteve uma pessoa infectada).
- máscaras e outros EPI's: reduzem a probabilidade de se infectar ao ter contato com uma pessoa infectada.

### 3.5. Taxa de recuperação $\gamma$

Reduzir o tempo de infecção irá aumentar a taxa de recuperação e conseqüentemente tornará  $\gamma$  mais próximo (ou maior) de  $\beta$ .

---

<sup>4</sup>Lembrando que a taxa de recuperação está relacionada com o tempo médio para se recuperar da infecção.

Até o momento ainda não há nenhum tratamento que comprovadamente reduza o tempo de infecção.

#### 4. Implementação do Modelo SIR

Para implementar computacionalmente o Modelo SIR são necessárias algumas informações sobre a população e sobre a infecção. Os parâmetros necessários são:

- $S$ : quantidade de pessoas no grupo de pessoas suscetíveis.  $S$  representa uma fração da população e, para haver propagação da infecção, deve ser maior que zero. Assim,  $0 < S \leq 1$ .
- $I$ : quantidade de pessoas no grupo dos infectados. Para termos propagação da infecção, devemos ter  $0 < I < 1$ . No início considera-se  $I \approx 0$  (mas não igual a zero).
- $R$ : quantidade de pessoas no grupo de pessoas recuperadas. Devemos ter  $0 \leq R < 1$ .
- $\gamma$ : taxa de recuperação. Para determinar  $\gamma$  é necessário saber o tempo médio de recuperação de uma pessoa infectada.
- $\beta$ : taxa de infecção. Na Subsecção 4.1 vamos apresentar uma técnica para estimar  $\beta$ .

##### 4.1. Estimando a taxa de infecção $\beta$

Como vimos na equação (7), o comportamento inicial da função que descreve o número de infectados ao longo do tempo é dado, aproximadamente, por uma função exponencial da forma

$$I(t) \approx I_0 e^{(\beta - \gamma)t}. \quad (14)$$

Observe que o expoente desta função exponencial é dado por  $(\beta - \gamma) \cdot t$ . Estudos indicam que uma pessoa infectada é capaz de transmitir a Covid-19, em media, por 10 dias (WOLFEL et al., 2020; YOU et al., 2020). Assim, podemos assumir que  $\gamma = \frac{1}{10}$ . Logo, se determinarmos a equação de uma função exponencial que represente  $I(t)$  poderemos estimar o valor de  $\beta$ .

Para determinar a função  $I(t)$  vamos utilizar a seguinte estratégia: a partir do número de infectados num período de 10 dias ajustamos uma função exponencial que melhor representa estes dados, assim teremos uma equação (aproximada) de  $I(t)$ .

Uma regressão exponencial é o processo de encontrar a equação da função exponencial que melhor se ajusta a um conjunto de dados. Como resultado, obtemos uma equação da forma  $y = a \cdot e^{b \cdot t}$ .

O poder preditivo relativo de um modelo exponencial é denotado por  $R^2$ . O valor de  $R^2$  varia entre 0 e 1. Quanto mais próximo de 1 for o valor de  $R^2$  mais preciso é o modelo.

### Exemplo 4.1

Considere a seguinte tabela de dados que apresenta uma sequência de 10 dias e o respectivo número de casos acumulados nestes dias:

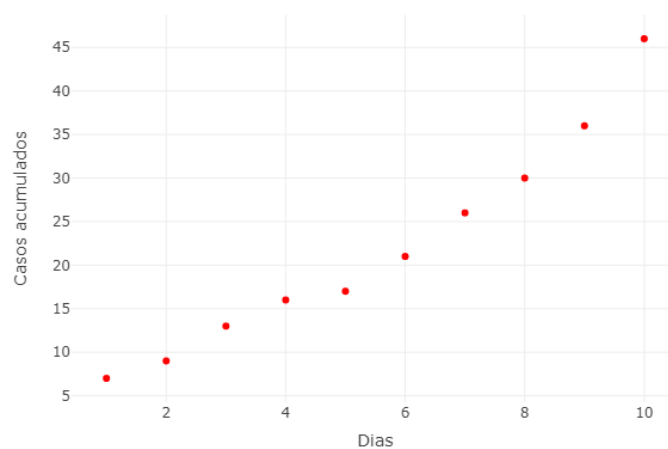
**Tabela 1** – Casos acumulados em 10 dias

dias	casos acumulados
1	7
2	9
3	13
4	16
5	17
6	21
7	26
8	30
9	36
10	46

**Fonte:** elaborada pelo autor.

Graficamente temos a seguinte representação dos dados

**Figura 5** – Casos acumulados em 10 dias



**Fonte:** Elaborada pelo autor.

O nosso objetivo é determinar uma função exponencial da forma  $a \cdot e^{b \cdot t}$  (claramente devemos ter  $a \neq 0$ ) que melhor represente o número de casos acumulados neste período de 10 dias. Assumindo que o comportamento da coluna casos acumulados da Tabela 1 seja exponencial, ou seja, casos acumulados  $\approx a \cdot e^{b \cdot t}$ , então aplicando o logaritmo natural obtemos

$$\ln(\text{casos acumulados}) \approx \ln(a) + b \cdot t, \quad (15)$$

que é uma equação de reta.

Desta forma, convertemos o problema de encontrar uma função exponencial que melhor represente a coluna casos acumulados da Tabela 1 no problema de encontrar uma reta que melhor se ajuste a  $\ln(\text{casos acumulados})$ . Assim, podemos aplicar o método dos mínimos quadrados (GUIDORIZZI, 2019, capítulo 17).

Vamos descrever o algoritmo, feito na linguagem R, para determinar os coeficientes  $a$  e  $b$  da função exponencial, a taxa de propagação  $\beta$  e gerar o gráfico da função exponencial.

A sintaxe da maior parte dos comandos será

{nome da variável que estamos criando} <- {o que será atribuído à variável criada}.

Inicialmente carregaremos a biblioteca **plotly**, que será utilizada para gerar o gráfico no final do algoritmo.

```
1 library(plotly)
```

A seguir criamos dois vetores, o primeiro formado pela sequência de 1 a 10 que representa a coluna dias da Tabela 1 e o segundo recebe os dados da coluna casos acumulados.

```
1 dias <- 1:10
2 casos <- c(7,9,13,16,17,21,26,30,36,46)
```

A seguir criamos o modelo linear com o comando **lm** que recebe o vetor casos após aplicar a função logaritmo natural (na linguagem R o logaritmo natural é definido pela função **log**) e o vetor dias.

```
1 exponential.model <- lm(log(casos)~dias)
```

A variável **exponential.model** receberá todos as informações do modelo e partir dela iremos extrair os coeficientes  $a$  e  $b$  e, conseqüentemente, a taxa de propagação  $\beta$ . O comando a seguir

```
1 summary(exponential.model)
```

retorna os dados do modelo:

```

Call:
lm(formula = log(casos) ~ dias)

Residuals:
    Min       1Q   Median       3Q      Max
-0.110775 -0.045154 -0.005971  0.015522  0.125036

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.859729   0.053047   35.06  4.79e-10 ***
dias         0.196956   0.008549   23.04  1.34e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.07765 on 8 degrees of freedom
Multiple R-squared:  0.9852,    Adjusted R-squared:  0.9833
F-statistic: 530.7 on 1 and 8 DF,  p-value: 1.338e-08

```

Observe que, neste caso,  $R^2 = 0,9833$ . Isto significa que 98,33% da variabilidade presente nos dados foi explicada pelo modelo. Além disso, considerando um nível de significância  $\alpha = 0,05$ , isto é, 95% de confiança, o modelo descreve adequadamente os dados uma vez que  $p\text{-value} = 1,338 \cdot 10^{-8} < 0,05$ .

A coluna **Estimate** apresenta, respectivamente, os valores de  $\ln(a)$  e  $b$  conforme a equação (15).

Note que se  $I(t) \approx a \cdot e^{b \cdot t}$ , então de (14) podemos assumir

$$a \cdot e^{b \cdot t} \approx I_0 e^{(\beta - \gamma) \cdot t},$$

ou seja,  $b = \beta - \gamma$ , isto é,  $\beta = b + \gamma$ . Como estamos considerando  $\gamma = 0,1$ , então  $\beta = b + 0,1$ .

Os seguintes comandos extraem a partir do modelo exponencial os valores de  $a$ ,  $b$  e  $\beta$  (note que para obtermos o coeficiente  $a$  aplicamos a função exponencial que na linguagem R é dada pelo comando **exp**).

```

1 a      <- exp(exponential.model$coefficients[[1]])
2 b      <- exponential.model$coefficients[[2]]
3 beta   <- b + 0.1

```



Obtemos os seguintes valores:

$$a = 6,421997, \quad b = 0,1969558 \quad \text{e} \quad \beta = 0,2969558.$$

Sendo  $\beta = 0,2969558$ ,  $\gamma = 0,1$  e  $R_0 = \frac{\beta}{\gamma}$ , então  $R_0 \approx 3$ . Isto significa que cada pessoa infectada transmite o vírus, em média, para três pessoas.

A seguir criamos a função exponencial.

```
1 fx <- function(x) {return(a*exp(b*x))}
```

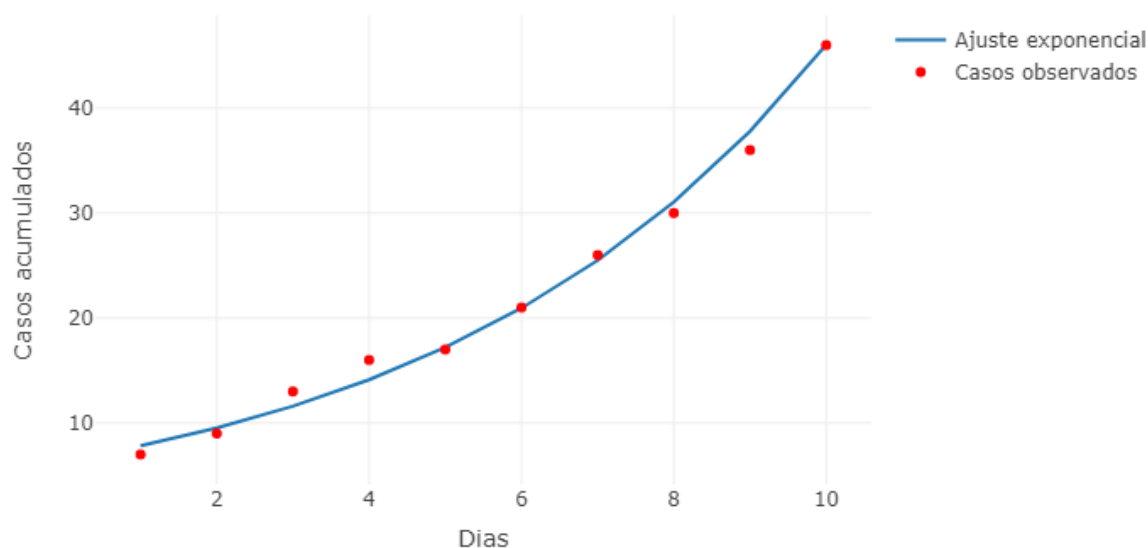
A equação da função exponencial obtida é

$$I(t) \approx 6,421997 \cdot e^{0,1969558 \cdot t}. \quad (16)$$

Por fim, os comandos a seguir geram o gráfico apresentado na Figura 6.

```
1 fig <- plot_ly(x = dias, y = fx(dias), type = "scatter",
2               mode = "lines", name="Ajuste exponencial",
3               lines=list(color='rgb(8,48,107)'))
4 fig <- fig %>% add_trace(y = casos, name="Casos observados",
5                          mode="markers", marker=list(color="red"))
6 fig <- fig %>% layout(title="", hovermode = TRUE, spikedistance = -1,
7                       xaxis=list(title="Dias"),
8                       yaxis=list(title="Casos acumulados"))
9 fig
```

**Figura 6** – Ajuste exponencial



**Fonte:** Elaborada pelo autor.

## 4.2. Resolvendo o Modelo SIR

Nesta subseção vamos implementar o Modelo SIR. Para tanto vamos considerar a seguinte situação hipotética: uma cidade com 100.000 habitantes, onde inicialmente temos sete pessoas infectadas, nenhuma pessoa recuperada e as demais todas suscetíveis. Além disso, vamos considerar que os dados apresentados na Tabela 1 representam os 10 primeiros dias de epidemia nesta cidade.

Assim, utilizando a taxa de propagação  $\beta$  obtida no Exemplo 4.1, temos os seguintes parâmetros iniciais para aplicar o Modelo SIR

$$\left\{ \begin{array}{l} S_0 = \frac{99.993}{100.000} \\ I_0 = \frac{7}{100.000} \\ R_0 = 0 \\ \beta = 0,2969558 \\ \gamma = 0,1. \end{array} \right.$$

Na linguagem R, é possível implementar as equações diferenciais do Modelo SIR apresentadas em (4) e utilizar métodos numéricos para resolver o sistema de equações diferenciais. No entanto, em R está disponível a biblioteca chamada *EpiDynamics* (SANTOS; SILVEIRA, 2015) que fornece comandos específicos para trabalharmos com modelos epidemiológicos de forma simples e intuitiva.

O sistema de equações é implementado e resolvido com o comando chamado **SIR** o qual precisa receber três vetores: o primeiro chamado **pars** contém a taxa de propagação **beta**, a taxa de recuperação **gamma** e a taxa de natalidade/mortalidade **mu** (iremos trabalhar com  $\mu=0$ ).

O segundo vetor chamado **init** deve conter os parâmetros **S**, **I** e **R** definidos pelos valores iniciais de cada grupo do Modelo SIR.

Por fim, o vetor chamado **time** dado por uma sequência numérica que descreve o período de tempo da simulação. Vamos considerar 180 dias.

```
1 sir <- SIR(pars = c(mu = 0, beta = 0.2969558, gamma = 1 / 10),
2           init  = c(S = S/N, I = I/N, R = R/N),
3           time  = 0:180)
```

As linhas de comandos acima armazenam na variável **sir** todos os dados do modelo incluindo as soluções numéricas do Modelo SIR apresentado em (4). O sistema de equações diferenciais é resolvido pelo método Runge-Kutta (ASANO; COLLI, 2009, capítulo 18). A

Tabela 3 apresenta as 10 primeiras linhas da solução.

**Tabela 3** – Primeiras 10 linhas da solução numérica do Modelo SIR

time	S	I	R
0	0.99993000	7.000000e-05	0.000000e+00
1	0.99990695	8.528938e-05	7.763815e-06
2	0.99987886	1.039146e-04	1.722189e-05
3	0.99984465	1.266044e-04	2.874457e-05
4	0.99980315	1.541242e-04	4.272089e-05
5	0.99975257	1.876721e-04	5.975982e-05
6	0.99969097	2.285213e-04	8.050884e-05
7	0.99961597	2.782537e-04	1.057726e-04
8	0.99952466	3.388009e-04	1.365341e-04
9	0.99941350	4.125108e-04	1.739887e-04

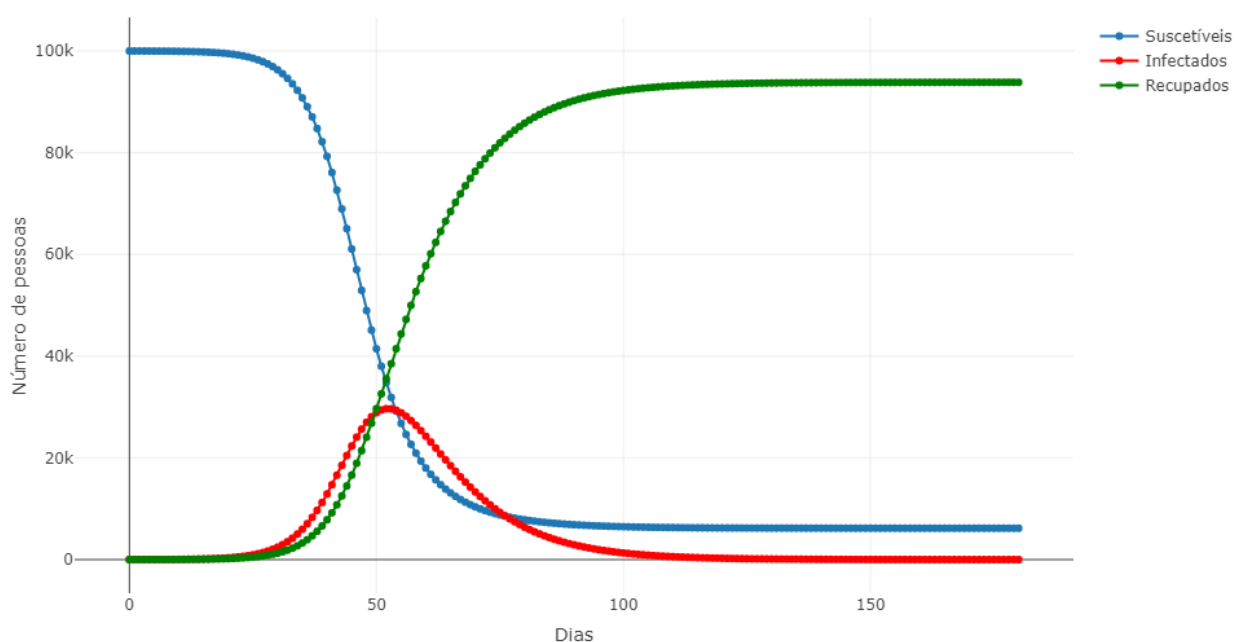
**Fonte:** elaborada pelo autor.

O resultado completo da solução pode ser acessado pelo comando `sir$results`. Os comandos a seguir geram o gráfico da solução do sistema de equações do Modelo SIR, o qual é apresentado na Figura 7.

```

1 fig <- plot_ly(x= ~sir$results$time, y=~N*sir$results$S, type = "scatter",
2               mode = "lines+markers", name="Suscetáveis")
3 fig <- fig %>% add_trace(y=~N*sir$results$I, name="Infectados",
4                           mode="lines+markers", line=list(color="red"),
5                           marker=list(color="red"))
6 fig <- fig %>% add_trace(y=~N*sir$results$R, name="Recupados",
7                           mode="lines+markers", line=list(color="green"),
8                           marker=list(color="green"))
9 fig <- fig %>% layout(title="",
10                       hovermode = TRUE, spikedistance = -1,
11                       xaxis=list(title="Dias"),
12                       yaxis=list(title="Número de pessoas"))
13 fig

```

**Figura 7** – Gráficos das soluções do Modelo SIR

Fonte: Elaborada pelo autor.

## 5. Conclusões

Do ponto de vista computacional, sobretudo na linguagem R, é simples a implementação do Modelo SIR assim como outros modelos como pode ser visto em (SANTOS; SILVEIRA, 2015). No entanto, identificar o modelo correto e, uma vez obtida a solução, interpretar corretamente as informações é uma tarefa árdua. Apenas para ilustrar este fato, observe o seguinte: na Figura 7 a curva em vermelho representa o número de pessoas infectadas ao longo do tempo. O pico atinge mais de 29 mil e ocorre em apenas 52 dias. Isto significa que em apenas 52 dias nesta cidade hipotética haveria, simultaneamente, mais de 29 mil pessoas infectadas (isto em uma população de apenas 100 mil). Além disso, pela curva em verde, dos recuperados, vemos que em 100 dias teríamos mais de 90 mil recuperados, ou seja, mais de 90% da população no grupo dos recuperados. Pensando em casos acumulados, teríamos mais de 90 mil casos em apenas 100 dias. Apenas para dar um possível contexto para este exemplo, considere a Covid-19. Não há relatos de nenhum lugar no mundo que tenha uma fração tão grande da população afetada. Isto ocorre por uma limitação do Modelo SIR, a taxa de propagação é mantida constante ao longo de todo o período o que aparentemente não acontece na realidade.

Isto não significa que o Modelo SIR não seja útil, pelo contrário, muitas informações sobre

o comportamento da pandemia e como mitigá-la podem ser extraídas do Modelo SIR como pode ser visto na Subseção 2.2. Além disso, é possível aprimorar o Modelo SIR e/ou utilizar outras estratégias e modelos em conjunto.

### Referências

- 1 ASANO, Claudio Hirofume; COLLI, Eduardo. Cálculo Numérico-Fundamentos e Aplicações. **Departamento de Matemática Aplicada, IME-USP**, v. 15, 2009.
- 2 BERGER, David W; HERKENHOFF, Kyle F; MONGEY, Simon. An seir infectious disease model with testing and conditional quarantine. **National Bureau of Economic Research**, 2020.
- 3 BOHNER, Martin; STREIPERT, Sabrina; TORRES, Delfim FM. Exact solution to a dynamic SIR model. **Nonlinear Analysis: Hybrid Systems**, Elsevier, v. 32, p. 228–238, 2019.
- 4 BOOTSMA, Martin CJ; FERGUSON, Neil M. The effect of public health measures on the 1918 influenza pandemic in US cities. **Proceedings of the National Academy of Sciences**, National Acad Sciences, v. 104, n. 18, p. 7588–7593, 2007.
- 5 GUIDORIZZI, Hamilton Luiz. **Um curso de cálculo**. 6. ed. São Paulo: Grupo Gen-LTC, 2019. v. 2.
- 6 HE, Xi et al. Temporal dynamics in viral shedding and transmissibility of COVID-19. **Nature medicine**, Nature Publishing Group, v. 26, n. 5, p. 672–675, 2020.
- 7 JONES, James Holland. Notes on R0. **California: Department of Anthropological Sciences**, 2007.
- 8 KEELING, Matt J; ROHANI, Pejman. **Modeling infectious diseases in humans and animals**. Princeton: Princeton University Press, 2011.
- 9 KERMACK, William Ogilvy; MCKENDRICK, Anderson G. A contribution to the mathematical theory of epidemics. **Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character**, The Royal Society London, v. 115, n. 772, p. 700–721, 1927.
- 10 RIDENHOUR, Benjamin; KOWALIK, Jessica M; SHAY, David K. Unraveling  $R_0$ : Considerations for Public Health Applications. **American journal of public health**, American Public Health Association, v. 108, S6, s445–s454, 2018.
- 11 SANTOS, O Baquero; SILVEIRA, F Marques. EpiDynamics: Dynamic Models in Epidemiology. **R package version 0.3.0**, 2015.
- 12 SHABBIR, G; KHAN, H; SADIQ, MA. A note on Exact solution of SIR and SIS epidemic models. **arXiv preprint arXiv:1012.5035**, 2010.
- 13 WOLFEL, R et al. Virological assessment of hospitalized cases of coronavirus disease 2019. **Nature**, 2020. DOI: <https://doi.org/10.1038/s41586-020-2196-x>.

- 14 YOU, Chong et al. Estimation of the time-varying reproduction number of COVID-19 outbreak in China. **International Journal of Hygiene and Environmental Health**, 2020.