

Estudo do Tamanho Final de uma Epidemia no Modelo SIQR CTMC

Michelle Lau de Almeida¹

PPG-CompMat/UERJ, Rio de Janeiro, RJ

Zochil González Arenas²

IME/UERJ, Rio de Janeiro, RJ

Resumo. A modelagem matemática é uma ferramenta fundamental na epidemiologia para a análise e prevenção da propagação de doenças infecciosas. Nesse âmbito, a modelagem estocástica tem ganhado destaque, pois ao considerar incerteza nos parâmetros ou nas transições entre possíveis estados epidemiológicos obtém-se uma modelagem mais realista. O modelo compartimental SIQR é uma variante do conhecido modelo SIR, que introduz um novo compartimento para considerar a possibilidade de quarentena de indivíduos infecciosos. Neste trabalho, uma versão estocástica do modelo SIQR é formulada através de Cadeias de Markov de Tempo Contínuo. Uma propriedade muito importante na modelagem epidemiológica refere-se ao tamanho final da epidemia, que representa o número total de indivíduos que contraíram a doença ao longo do tempo de duração da epidemia. Esta propriedade pode ser obtida através de diversas abordagens. Em modelos baseados em Cadeias de Markov, pode-se determinar a sua distribuição de probabilidade e, neste trabalho, apresentamos uma forma para calcular esta distribuição para um modelo SIQR CTMC.

Palavras-chave. Modelos Estocásticos, Modelagem Epidemiológica, Modelo SIQR, Cadeias de Markov de Tempo Contínuo

1 Introdução

A modelagem matemática na epidemiologia de doenças infecciosas tem entre seus objetivos melhorar as estratégias de detecção, prevenção e controle destas doenças, com o intuito de auxiliar governos e profissionais de saúde na tomada de decisões precisas. A modelagem epidemiológica tem o potencial de fornecer previsões abrangentes e estimar a incerteza associada a essas previsões [9]. Para conseguir estes objetivos, modelos são definidos, como uma representação simplificada da realidade, de modo que possam ser estudadas propriedades fundamentais da evolução e transmissão das doenças, como pode ser a duração de uma epidemia e o seu tamanho final. Esta última é uma propriedade muito importante a ser estabelecida e refere-se ao número total de indivíduos que contraíram a doença ao longo do tempo de duração da epidemia. Além da criação de modelos, realizar simulações computacionais cumpre um papel significativo na definição e teste de hipóteses, assim como para determinar a sensibilidade a mudanças nos valores dos parâmetros considerados.

Os modelos para analisar a propagação de doenças podem ser determinísticos ou estocásticos. A diferença básica entre estas abordagens é dada pela forma de acesso à informação sobre o sistema que se modela. Enquanto na abordagem determinística, uma vez definidas as condições iniciais fica univocamente determinada a evolução do sistema, numa abordagem estocástica tem-se acesso a diversas evoluções possíveis, precisando-se de uma análise estatística dessa informação para se

¹michelle.lau@ime.uerj.br

²zochil@ime.uerj.br

chegar a conclusões válidas. Assim, através da modelagem estocástica podem ser consideradas situações de incerteza nos parâmetros ou nas transições entre os estados do sistema e pode ser associada uma distribuição de probabilidade aos resultados possíveis para a evolução da doença [1]. Em alguns casos, isto permite obter uma representação mais realista na modelagem, ao considerar fenômenos não totalmente compreendidos ou influenciados por fatores aleatórios.

A utilização de modelos compartimentais é uma abordagem muito comum na área de epidemiologia. Nestes modelos, a população é dividida em compartimentos que representam diferentes estágios da doença ou estratégias que podem afetar a sua propagação. Nesse contexto, em ocasião da incidência da COVID-19 foi muito divulgado o modelo SIR, no qual divide-se a população em três compartimentos: suscetíveis (S), infecciosos (I) e recuperados (R). Para citar exemplos de compartimentos referentes a estratégias de controle, podem se considerar compartimentos como hospitalizados (H) e quarentenados (Q). Considerando a possibilidade de quarentena de indivíduos infecciosos através deste último compartimento, foi obtido o modelo compartimental SIQR, como uma variação do modelo SIR.

O modelo compartimental SIQR tem sido amplamente utilizado na modelagem epidemiológica [5, 7, 11]. Neste trabalho, será apresentada uma versão estocástica deste modelo, através do uso de Cadeias de Markov de Tempo Contínuo (CTMC) [2]. Modelos epidemiológicos estocásticos baseados em Cadeias de Markov são particularmente importantes no estudo dos estágios iniciais de contágio de uma doença, ao permitir analisar a evolução indivíduo a indivíduo, fornecendo uma compreensão quantitativa de vários fenômenos [6]. Usando esta abordagem, é possível encontrar a distribuição do tamanho final, detalhando a probabilidade de cada um dos resultados possíveis de uma epidemia. A estimação do tamanho final é um assunto de pesquisa atual e de amplo interesse, abordado através de diferentes técnicas [3, 4]. Neste trabalho, consideramos o cálculo da distribuição para o tamanho final da epidemia para o modelo SIQR CTMC.

O trabalho é organizado da seguinte forma. Na próxima seção, apresenta-se brevemente o modelo SIQR na sua versão determinística, no intuito de introduzir os conceitos básicos do modelo. A seguir, definem-se o modelo SIQR CTMC e a estratégia para o cálculo do tamanho final da epidemia utilizando este modelo. Por fim, são apresentadas as considerações finais e as principais referências bibliográficas utilizadas.

2 Modelo SIQR

Como foi introduzido antes, no modelo compartimental SIQR, considera-se um compartimento Q , denominado *quarentenado*, que é composto pelos indivíduos infectados que são isolados voluntaria ou coercivamente. Supõe-se que os indivíduos do compartimento quarentenado não tenham contato com outros, de forma a evitar a transmissão da doença para indivíduos suscetíveis [8]. Na Figura 1, ilustram-se os compartimentos deste modelo e as interações possíveis entre eles.

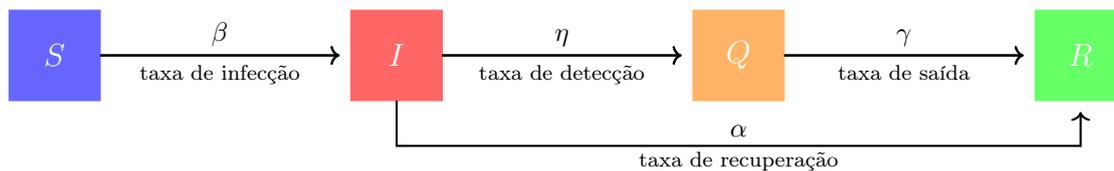


Figura 1: Diagrama do modelo $SIQR$.

Neste modelo, a população é considerada constante, desconsiderando nascimentos ou mortes, e descreve-se uma doença que produz imunidade permanente, pois indivíduos recuperados não voltam a ser suscetíveis. O tamanho da população é denotado por N , com $N = S(t) + I(t) + Q(t) + R(t)$.

O modelo SIQR, em sua versão determinística [5], é descrito pelo sistema de equações diferenciais a seguir,

$$\begin{aligned} \frac{dS}{dt} &= -\frac{\beta}{N}SI, \\ \frac{dI}{dt} &= \frac{\beta}{N}SI - (\alpha + \eta)I, \\ \frac{dQ}{dt} &= \eta I - \gamma Q, \\ \frac{dR}{dt} &= \gamma Q + \alpha I. \end{aligned} \tag{1}$$

No instante inicial $t = 0$, considera-se que não há indivíduos nos compartimentos Q e R , ou seja, $Q(0) = R(0) = 0$, pois ainda não foram isolados os indivíduos infecciosos e não há indivíduos recuperados. Além disso, $S(0) = N - I(0)$. No sistema de equações (1), β é a taxa de contato ou taxa de infecção, α é a taxa de recuperação de indivíduos assintomáticos ou que não se isolaram, η é a taxa de detecção de novos casos, ou seja, dos que testaram positivo para a doença e se isolaram e, por fim, γ é a taxa de saída de um indivíduo em quarentena para a classe de recuperados.

3 Modelo Estocástico

No modelo SIQR CTMC os compartimentos S , I , Q e R são considerados como variáveis aleatórias em cada instante de tempo. As probabilidades de transição são definidas em termos das taxas do modelo determinístico, considerando a mudança que ocorre nos compartimentos S , I e Q durante um intervalo de tempo pequeno Δt na forma $\Delta S = S(t + \Delta t) - S(t)$, $\Delta I = I(t + \Delta t) - I(t)$ e $\Delta Q = Q(t + \Delta t) - Q(t)$. Assim, as probabilidades descritas na Tabela 1 representam a variação nestes compartimentos no intervalo Δt . Cada tripla $(\Delta S, \Delta I, \Delta Q)$ representa uma transição entre os possíveis estados do sistema. O compartimento R não é necessário para avaliar essas transições, uma vez que não há uma variação decrescente neste compartimento. Através de $(\Delta S, \Delta I, \Delta Q) = (-1, 1, 0)$ representa-se a condição de um indivíduo suscetível ter sido infectado, $(\Delta S, \Delta I, \Delta Q) = (0, -1, 0)$ ocorre quando um indivíduo infectado se recuperou. No caso em que $(\Delta S, \Delta I, \Delta Q) = (0, -1, 1)$ caracteriza-se um indivíduo infectado que foi isolado e $(\Delta S, \Delta I, \Delta Q) = (0, 0, -1)$ representa a saída de um indivíduo do isolamento. Por fim, $(\Delta S, \Delta I, \Delta Q) = (0, 0, 0)$ representa que não houve alteração nos compartimentos, ou seja, nenhum indivíduo foi infectado, isolado ou recuperado.

Tabela 1: Representação das probabilidades de transição utilizadas no modelo SIQR.

$(\Delta S, \Delta I, \Delta Q)$	Probabilidade
$(-1, 1, 0)$	$\frac{\beta}{N}S(t)I(t)\Delta t + o[\Delta t]$
$(0, -1, 0)$	$\alpha I(t)\Delta t + o[\Delta t]$
$(0, -1, 1)$	$\eta I(t)\Delta t + o[\Delta t]$
$(0, 0, -1)$	$\gamma Q(t)\Delta t + o[\Delta t]$
$(0, 0, 0)$	$1 - \left[\frac{\beta}{N}S(t)I(t) + \alpha I(t) + \eta I(t) + \gamma Q(t) \right] \Delta t$

É importante destacar que a evolução do sistema, para cada instante de tempo, é dada pela transição de apenas um indivíduo entre dois compartimentos. Cada transição no sistema tem uma probabilidade associada e, portanto, não há transições simultâneas entre dois ou mais pares de

compartimentos. Por exemplo, se houver uma diminuição de um indivíduo no compartimento de suscetíveis, isso resultará no aumento correspondente no grupo de infecciosos (ou seja, um indivíduo suscetível que se tornou infeccioso), o que se encontra representado na primeira linha da Tabela 1. Ao mesmo tempo, não será possível ter uma diminuição no número de infecciosos, representado pela segunda linha da tabela, nem uma variação (positiva ou negativa) no número de indivíduos em qualquer outro compartimento (linhas 3 ou 4 da tabela), uma vez que já houve uma transição nesse intervalo de tempo. Essa transição é estritamente definida por um indivíduo que sai de S e entra em I . Qualquer outra transição possível deverá acontecer em outro instante de tempo.

Neste modelo, os estados do sistema são considerados na ordem $(S(t), I(t), Q(t)) = (s, i, q)$, indicando o número de indivíduos nos compartimentos S , I e Q para cada estado, em um determinado instante de tempo t , de modo que $0 \leq s + i + q \leq N$. Define-se a distribuição inicial do sistema, para $t = 0$, como $(S(0), I(0), Q(0)) = (s_0, i_0, q_0)$, onde $s_0 + i_0 = N$ e $q_0 = 0$, pois considera-se que no instante inicial não há indivíduos quarentenados. Como descrito no modelo determinístico, considera-se também vazio o compartimento de indivíduos recuperados no instante inicial, $R(0) = 0$ e se cumpre que $s_0 \geq 0$ e $i_0 > 0$.

Nos $N + 1$ estados dados pela tripla $(s, 0, 0)$, onde $s = 0, 1, 2, \dots, N$, não ocorrem transições para qualquer outro estado do sistema, pois não há indivíduos infectados nem quarentenados e, portanto, esses estados são fechados. Esses estados são denominados absorventes pois não há mais transições possíveis a serem realizadas quando um deles é alcançado. No modelo SIQR CTMC trabalhamos com as probabilidades conjuntas dos compartimentos S , I e Q . Dessa forma, o número total de estados possíveis associados ao modelo SIQR CTMC corresponde ao número de soluções inteiras e não negativas da inequação $0 \leq S + I + Q \leq N$. Assim, neste modelo, temos um total de $(N + 1)(N + 2)(N + 3)/6$ estados possíveis [10].

3.1 Tamanho final da Epidemia

A gravidade de uma epidemia pode ser avaliada pelo número total de casos que ocorreram, conhecido como tamanho final da epidemia. O tamanho final de uma epidemia pode ser definido, então, como o número total de pessoas que contraem a infecção ao longo do surto da doença. Essa quantidade é tipicamente chamada pelos epidemiologistas de taxa de ataque e é expressa como uma porcentagem da população em questão. Ao considerarmos $t \rightarrow \infty$, o tamanho final de uma epidemia é calculado, nos modelos determinísticos, por meio do $R(\infty)$ sendo $R(\infty) = N - S(\infty)$, pois representa todos os indivíduos que foram infectados e se recuperaram. No modelo estocástico, determina-se esta propriedade através da mesma relação dos modelos determinísticos. Porém, em modelos CTMC pode-se obter uma distribuição de probabilidade para esse tamanho final e assim, analisar características específicas do sistema. Para calcular essa distribuição de probabilidade para o tamanho final da epidemia no modelo SIQR CTMC, é necessário determinar a matriz de transição da Cadeia de Markov Incorporada (EMC³) do sistema, que será denotada por P_N . Esta distribuição ficará definida a partir de uma dada potência de P_N , como mostraremos a seguir.

A matriz de transição P_N depende de como os estados do sistema são ordenados. Como visto na seção anterior, no modelo SIQR CTMC trabalhamos com as probabilidades conjuntas dos compartimentos S , I e Q e, portanto, existem $(N + 1)(N + 2)(N + 3)/6$ estados possíveis. Uma ordenação qualquer desse total de estados pode tornar desafiadora a obtenção dessas matrizes ou resultar em uma estrutura de matriz que prejudica a eficiência dos cálculos subsequentes. Encontrar uma ordenação apropriada é, de fato, um aspecto crucial nesses modelos [1].

Considerando uma dada ordenação das triplas de estados do modelo, denotando por $p_{(s,i,q)}$ a probabilidade do estado formado por s indivíduos suscetíveis, i indivíduos infecciosos e q indivíduos

³Embedded Markov Chain.

quarentenados, representamos o conjunto das probabilidades de estado através do vetor

$$p(t) = (p_{(N,0,0)}, p_{(N-1,0,0)}, \dots, p_{(0,0,N)})^T, \tag{2}$$

onde T denota a transposta.

Para organizarmos esses estados na matriz de transição, cada um deles também será denotado por um número inteiro, na seguinte forma:

$$\begin{array}{lll} (N, 0, 0) & \rightarrow & 1 \\ (N - 1, 0, 0) & \rightarrow & 2 \\ \vdots & \dots & \vdots \\ (0, 0, N) & \rightarrow & N_F = \frac{(N + 1)(N + 2)(N + 3)}{6} \end{array} \tag{3}$$

A matriz de transição P_N é uma matriz quadrada de dimensão N_F , construída usando-se as probabilidades de transição de um estado l para outro estado k , \tilde{p}_{kl} . Os primeiros $N + 1$ estados são classificados como estados absorventes, pois a partir deles não há transição possível para outros estados, dado que não há indivíduos infectados nem quarentenados, ou seja, $(s, i, q) = (s, 0, 0)$. Isso significa que a matriz de transição $P_N = (\tilde{p}_{kl})$ da EMC satisfaz $\tilde{p}_{ll} = 1$, para $l = 1, 2, \dots, N + 1$.

Considerando $(S(t), I(t), Q(t)) = (s, i, q)$, podem-se determinar os elementos da matriz P_N a partir das probabilidades de transição dos estados possíveis, sabendo que no modelo é analisada a variação de somente um indivíduo a cada instante de tempo. A transição do estado (s, i, q) para o estado $(s - 1, i + 1, q)$ representa a infecção de um indivíduo suscetível, ao passo que a transição do estado (s, i, q) para o estado $(s, i - 1, q)$ significa a recuperação de um indivíduo infectado. Já a transição do estado (s, i, q) para o estado $(s, i - 1, q + 1)$ representa o isolamento de um indivíduo infectado e a transição do estado (s, i, q) para o estado $(s, i, q - 1)$ representa a recuperação de um indivíduo que estava em quarentena.

A matriz de transição P_N da EMC é muito útil no cálculo do tamanho final da epidemia. Em geral, no modelo SIQR CTMC, para qualquer população de tamanho N e começando com um indivíduo infectado, $p_{(N-1,1,0)}(0) = 1$ (probabilidade de estado, dada pela expressão (2)), o número máximo de transições até ocorrer a absorção, ou seja, quando não há mais indivíduos infectados, é $3N - 1$. Este resultado decorre do presente trabalho e a sua demonstração está sendo construída através de uma abordagem de teoria de grafos, da qual apresentamos a ideia geral a seguir. O modelo SIQR CTMC pode ser considerado como um grafo orientado sem ciclos, chamado de digrafo acíclico. Cada estado do sistema representa um vértice do grafo e as transições possíveis entre os estados são dadas pelas arestas. Nesse contexto, é possível definir as trajetórias, em termos das transições possíveis entre os estados inicial e final, que determinam o tamanho do caminho máximo. Dessa forma, encontra-se o valor do caminho máximo até qualquer estado absorvente como estado final. Dado que ainda estamos trabalhando na formalização desta demonstração, neste trabalho apresentamos este resultado como uma conjectura.

Conjectura 3.1. *Para qualquer população de tamanho N , começando com um indivíduo infeccioso na população, ou seja, $p_{(N-1,1,0)}(0) = 1$, no modelo SIQR CTMC, o número máximo de transições até a absorção é $3N - 1$.*

Considerando que a Conjectura 3.1 seja verdadeira, fica estabelecido o número máximo de passos até alcançar um estado absorvente. Como foi definido acima, os estados absorventes são aqueles onde não há infectados nem quarentenados, ou seja, $(s, 0, 0)$. Eles representam a situação na qual a epidemia acabou, de modo que a partir desse momento, não há outros estados possíveis. Assim sendo, como consequência da normalização da distribuição de probabilidade, a soma das

probabilidades dos estados absorventes é 1. Além disso, dado que, inicialmente, $I(0) = 1$, o somatório só pode ser considerado até $N - 1$. Isto fica representado através da seguinte expressão:

$$\lim_{t \rightarrow \infty} \sum_{s=0}^{N-1} \tilde{p}_{(s,0,0)}(t) = 1. \tag{4}$$

A Equação (4) estabelece que os estados absorventes constituem uma distribuição de probabilidade para um tempo suficientemente grande, $t \rightarrow \infty$, no qual a epidemia tenha acabado. Estas probabilidades podem ser calculadas através da potência $3N - 1$ da matriz de transição P_N [1, 2].

A distribuição de probabilidade associada ao tamanho final da epidemia $\{p_i^f\}$, $i = 1, \dots, N$, pode ser encontrada a partir do cálculo das probabilidades de absorção [1]. Estas probabilidades de tamanho final satisfazem

$$\lim_{t \rightarrow \infty} \tilde{p}_{(s,0,0)}(t) = p_{N-s}^f, \tag{5}$$

para $s = 0, 1, 2, \dots, N - 1$. Havendo s indivíduos suscetíveis uma vez que não há mais indivíduos infecciosos, o tamanho final da epidemia é $N - s$. Dessa forma, começando com apenas um indivíduo infeccioso e dado que no instante $t = 3N - 1$ foram alcançados todos os estados absorventes, obtém-se

$$\lim_{t \rightarrow \infty} p(t) = p(3N - 1) = (P_N)^{3N-1} p(0), \tag{6}$$

onde $p(t)$ é o vetor de probabilidade, dado pela expressão (2). O vetor $p(0)$ tem a particularidade de ter todas as suas componentes nulas, excetuando a componente $N + 2$, que representa a posição do estado inicial $(s - 1, 1, 0)$. A expressão (6) fornece o resultado que estava-se procurando, que é o vetor final associado à distribuição de probabilidade para o tamanho final de uma epidemia. Pode-se notar que este vetor corresponde à coluna $N + 2$ da potência $3N - 1$ da matriz de transição do sistema.

A distribuição de probabilidade do tamanho final permite avaliar riscos associados à possível evolução temporal de uma epidemia como, por exemplo, a identificação de cenários de alto risco epidêmico e a avaliação da probabilidade de ocorrência de surtos em grande escala. Ao conhecer esta distribuição de probabilidade, podem ser desenvolvidos planos de contingência mais robustos, definindo uma alocação eficiente de recursos e o estabelecimento de estratégias para antecipação de potenciais cenários de transmissão da doença [3].

4 Considerações Finais

Os modelos estocásticos epidemiológicos propostos através de Cadeias de Markov de Tempo Contínuo permitem inserir uma certa aleatoriedade nas transições entre os estados epidemiológicos e encontrar a distribuição de probabilidade associada aos resultados possíveis para a dinâmica evolutiva de uma doença. Dessa forma, podem ser explorados diversos cenários de propagação, permitindo avaliar a possível evolução de uma epidemia.

No presente trabalho, apresentamos uma versão estocástica do modelo compartimental SIQR através do uso de Cadeia de Markov de Tempo Contínuo (CTMC). Como uma primeira aproximação ao tema utilizando este modelo, na abordagem proposta não foram consideradas outras estratégias ou medidas de controle da propagação da doença além da quarentena de indivíduos infecciosos. Para este modelo foi estudada uma propriedade fundamental na área epidemiológica, referente ao tamanho final de uma epidemia. No trabalho foram introduzidos alguns conceitos e ferramentas básicas que permitiram chegar nos resultados procurados e, por fim, apresentamos uma expressão para o vetor de tamanho final de uma epidemia para o modelo SIQR CTMC.

Como trabalho futuro, está sendo desenvolvida a implementação deste método visando a ampliar a aplicação dos resultados obtidos através da simulação computacional.

Agradecimentos

O presente trabalho foi realizado com apoio da Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ) e da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001. M.L.A. é bolsista de doutorado da CAPES.

Referências

- [1] L. J. S. Allen. **An introduction to stochastic processes with applications to biology**. 2a. ed. Lubbock: CRC Press, 2010. ISBN: 9781439894682.
- [2] M. L. Almeida. “Modelagem Epidemiológica com um Modelo SIR Estocástico utilizando Cadeia de Markov de Tempo Contínuo”. Dissertação de mestrado. Universidade do Estado do Rio de Janeiro, 2020, p. 75.
- [3] A. J. Black e J. V. Ross. “Computation of epidemic final size distributions”. Em: **Journal of theoretical biology** 367 (2015), pp. 159–165.
- [4] F. Brauer. “A final size relation for epidemic models of vector-transmitted diseases”. Em: **Infectious Disease Modelling** 2.1 (2017), pp. 12–20.
- [5] N. Crokidakis. “Modeling the early evolution of the COVID-19 in Brazil: Results from a Susceptible–Infectious–Quarantined–Recovered (SIQR) model”. Em: **International Journal of Modern Physics C (IJMPC)** 31.10 (2020), pp. 2050135/1–2050135/7. DOI: 10.1142/S0129183120501351.
- [6] D. J. Daley e J. Gani. **Epidemic modelling: an introduction**. 1a. ed. 15. New York: Cambridge University Press, 2001. ISBN: 978-0521014670.
- [7] M. Erdem, M. Safan e C. Castillo-Chavez. “Mathematical analysis of an SIQR influenza model with imperfect quarantine”. Em: **Bulletin of Mathematical Biology** 79 (2017), pp. 1612–1636. DOI: 10.1007/s11538-017-0301-6.
- [8] H. Hethcote, M. Zhien e L. Shengbing. “Effects of quarantine in six endemic models for infectious diseases”. Em: **Mathematical biosciences** 180.1-2 (2002), pp. 141–160. DOI: 10.1016/S0025-5564(02)00111-6.
- [9] H. W. Hethcote. “The Mathematics of Infectious Diseases”. Em: **SIAM Review** 42.4 (2000), pp. 599–653. DOI: 10.1137/S0036144500371907.
- [10] M. Lau, Z. G. Arenas e D. Nicodemos. “Modelagem Estocástica da Evolução de Epidemias Através do Modelo SIQR CTMC”. Em: **Proceeding Series of the Brazilian Society of Computational and Applied Mathematics** 10.1 (2023), pp. 2–6.
- [11] T. Odagaki. “Exact properties of SIQR model for COVID-19”. Em: **Physica A: Statistical Mechanics and its Applications** 564 (2021), pp. 1–9. DOI: 10.1016/j.physa.2020.125564.