

UNIVERSIDADE ESTADUAL DE MARINGÁ
CENTRO DE CIÊNCIAS EXATAS - CCE
DEPARTAMENTO DE FÍSICA - DFI

ADRIANO HENRIQUE DANHONI NEVES

Sistemas Dinâmicos de Transmissão para Doenças Infectocontagiosas

Maringá
Abril, 2021

ADRIANO HENRIQUE DANHONI NEVES

Sistemas Dinâmicos de Transmissão para Doenças Infectocontagiosas

Trabalho de conclusão de curso apresentado ao Departamento de Física da Universidade Estadual de Maringá.

Orientador:
Prof. Dr. Denner Serafim Vieira
Coorientadora:
Profa. Dra. Hatsumi Mukai

Maringá
Abril, 2021

Dedico este trabalho aos meus pais, José Adriano e Marlene, pelo carinho e apoio durante toda essa jornada.

Agradecimentos

Primeiramente, quero agradecer meu orientador, Professor Denner Serafim Vieira, por desde cedo se mostrar entusiasmado com o trabalho e aceitar ser meu orientador. Obrigado pelas sugestões, paciência, dedicação e pelas conversas amigáveis ao longo destes mais de nove meses de orientação. Tenho muito o que levar de aprendizado deste trabalho.

À minha co-orientadora, Professora Hatsumi Mukai, por garantir que esta orientação fosse realizável. Por seus esforços, guiando não só a mim, mas diversos outros colegas com relação aos trâmites deste trabalho, além de toda a dedicação e disposição demonstradas ao longo deste período.

Aos meus pais, por sempre estarem presentes em minha vida, fornecendo amor e carinho, e pelo apoio e incentivo desde o momento em que escolhi cursar física.

Aos meus grandíssimos amigos Leonardo e Demétrio, pela companhia e amizade durante toda a graduação, amizade essa que espero levar para a vida toda. Sou muito grato por todos os momentos juntos.

A todos os meus amigos, com os quais partilho longos anos de amizade e que estiveram presentes nos momentos mais marcantes de minha vida. Agradeço muito por isso.

Agradeço também aos Professores Renio, Haroldo e Breno que, de certo modo, tiveram grande influência no caminho que decidi tomar para o meu futuro, não somente profissional, mas como pessoa em si.

Living is worthwhile if one can contribute in some small way to this endless chain of progress.

Paul Dirac.

Resumo

Neste trabalho, é estudado o modelo de Kermack-McKendrick, que usa a lei de ação das massas e separa a população hospedeira em compartimentos: suscetíveis (S), infectados (I) e removidos (R), conhecido como modelo SIR, cuja dinâmica é regida por um conjunto de equações diferenciais não lineares. São estudadas algumas variações desse modelo SIR, tal qual os modelos SEIR, SIRS e SIS, que levam em consideração outras características da doença, como a possibilidade de reinfeção e incubação da mesma. Estuda-se também esse modelo SIR exposto à uma abordagem estocástica, com o uso de cadeias de Markov, tratando os compartimentos como variáveis estocásticas discretas. Nesse cenário, a matriz de transição rege as probabilidades de mudança de estado do sistema. Com isso, tratam-se os dados analisados da COVID-19 para cinco estados brasileiros (São Paulo, Rio de Janeiro, Paraná, Santa Catarina e Rio Grande do Sul) e três países (Estados Unidos da América, Índia e Brasil), de modo que realiza-se uma comparação entre o número de infectados e recuperados com o total de casos dos dados empíricos. Essa abordagem permite obter a evolução das taxas de transmissão (β) e de recuperação (γ), além do número de reprodução dependente do tempo ($\mathcal{R}(t)$) da doença, sendo esse último o principal indicador de invasão de uma enfermidade em uma população. Os resultados obtidos desta comparação evidenciam a acurácia dessa teoria estudada frente à modelagem de surtos epidêmicos reais. Comparando o número de reprodução dependente do tempo $\mathcal{R}(t)$ empírico e o fornecido pelo modelo SIR e pelo SIR com dinâmicas vitais, conclui-se que este ajuste se torna mais preciso a medida que são adicionados mais parâmetros e efeitos dinâmicos presentes na sociedade.

Palavras-chave: Matemática epidemiológica, Modelo de Kermack-McKendrick, COVID-19, Número de reprodução, modelo SIR DTMC

Abstract

In this work, the Kermack-McKendrick model is studied, which uses the law of mass action and separates the host population into compartments: susceptibles (S), infected (I) and removed (R), known as SIR model, whose dynamics is dictated by a set of nonlinear differential equations. Some variations of the SIR model are studied too, such as the SEIR, SIRS and SIS models, which take into account other characteristics of the diseases, such as the possibility of reinfection and incubation. It is also studied the SIR model exposed to a stochastic approach, through Markov chains, treating the compartments as discrete stochastic variables. In this scenario, the transition matrix governs the probabilities of change in the state of the system. With this, the analyzed data from COVID-19 for five Brazilian states (São Paulo, Rio de Janeiro, Paraná, Santa Catarina and Rio Grande do Sul) and three countries (United States of America, India and Brazil) are treated, so that a comparison is made between the number of infected and recovered with the total number of cases in the empirical data. This approach allows to obtain the evolution of the transmission (β) and recovery (γ) rates, in addition to the time-dependent number of reproduction ($\mathcal{R}(t)$) of the disease, with the latter being the main indicator of disease invasion in a population. The results obtained from this comparison show the accuracy of this theory studied in relation to the modeling of real epidemic outbreaks. Comparing the empirical time-dependent number of reproduction $\mathcal{R}(t)$ and that provided by the SIR model and the SIR with demographic effects, it is concluded that this adjustment becomes more precise with the addition of more parameters and dynamical effects present in society.

Key words: Mathematical epidemiology, Kermack-McKendrick model, COVID-19, Reproduction number, SIR DTMC model

Lista de Figuras

1.1	O problema do crescimento populacional, por Fibonacci . . .	4
1.2	Fluxograma do modelo compartimental SIR	8
1.3	Simulação do modelo SIR	12
1.4	Gráficos do modelo SIR	13
1.5	Fluxograma do modelo compartimental SIRM	14
1.6	Fluxograma do modelo compartimental SIR apresentando dinâmicas vitais	16
1.7	Fluxograma do modelo compartimental SIRS	17
1.8	Simulação do modelo SIRS	18
1.9	Gráficos do modelo SIRS	19
1.10	Gráficos do modelo SIRS para longos períodos de tempo . . .	20
1.11	Plano de fase SIRS para longos períodos de tempo	20
1.12	Fluxograma do modelo compartimental SEIR	21
1.13	Simulação do modelo SEIR	21
1.14	Gráficos do modelo SEIR	22
1.15	Fluxograma do modelo compartimental SEIRM	23
1.16	Fluxograma do modelo compartimental SEIR apresentando dinâmicas vitais	24
1.17	Fluxograma do modelo compartimental SIS	24
1.18	Simulação do modelo SIS	26
1.19	Fluxograma do modelo compartimental SIS apresentando di- nâmicas vitais	27
1.20	Gráficos do modelo SIS com dinâmicas vitais	28

2.1	Simulação de uma caminhada aleatória com passos uniformes	30
2.2	Simulação de uma caminhada aleatória com passos não uniformes	31
2.3	Gráfico do modelo SIS DTMC	36
2.4	Gráfico do modelo SIR DTMC	38
3.1	Dados de ocorrências globais da COVID-19 (OWID)	40
3.2	Os três países com maior número de casos (OWID)	41
3.3	Mapa referente à distribuição de casos acumulados da COVID-19 (OWID)	42
3.4	Mapa referente à distribuição de mortes da COVID-19 (OWID)	42
3.5	Distribuições de casos da COVID-19 nos estados brasileiros .	43
3.6	Distribuições de mortes acumuladas da COVID-19 nos estados brasileiros	44
4.1	Ajuste do modelo SIR aos dados globais	48
4.2	Ajuste do modelo SIR aos estados brasileiros	49
4.3	Ajuste do modelo SIR com dinâmicas vitais aos dados globais	51

Lista de Tabelas

1.1	Tabela de órbitos agrupada por Graunt (1662)	4
4.1	Resultados do modelo SIR à nível mundial	47
4.2	Resultados do modelo SIR com dinâmicas vitais à nível mundial	52
A.1	Ocorrências globais disponibilizadas pela OMS dos 30 países com maior número de casos registrados	59
A.2	Ocorrências globais disponibilizadas pela OMS, com países ordenados de 31° a 60°.	60
A.3	Ocorrências globais disponibilizadas pela OMS, com países ordenados de 61° a 90°.	61
A.4	Ocorrências globais disponibilizadas pela OMS, com países ordenados de 91° a 120°.	62
A.5	Ocorrências globais disponibilizadas pela OMS, com países ordenados de 121° a 150°.	63
A.6	Ocorrências globais disponibilizadas pela OMS, com países ordenados de 151° a 180°.	64
A.7	Ocorrências globais disponibilizadas pela OMS, com países ordenados de 181° a 210°.	65
A.8	Ocorrências globais disponibilizadas pela OMS, com países ordenados de 211° a 237°.	66
A.9	Dados das ocorrências nacionais.	67

Introdução	1
1 Modelos Compartmentais em Epidemiologia	3
1.1 Modelos Deterministas em Epidemiologia	6
1.1.1 Modelo de Kermack-McKendrick	6
1.1.2 Modelo de Kermack-McKendrick com taxa de contato gene- ralizada	11
1.1.3 Modelo de Kermack-McKendrick com dinâmicas vitais	15
1.1.4 Modelo SIRS	16
1.1.5 Modelo SEIR	19
1.1.6 Modelo SIS	24
2 Modelos Estocásticos em Epidemiologia	29
2.1 Cadeias de Markov	29
2.2 Modelo DTMC	33
3 Dados da COVID-19	39
3.1 Dados mundiais	39
3.2 Brasil	40
4 Modelagem dos dados	45
4.1 Ajustando o modelo SIR	45
4.2 Ajustando o modelo SIR com dinâmicas vitais	50
Conclusões	53

Referências	55
A Tabelas de dados	58

INTRODUÇÃO

Em dezembro de 2019, foi detectado e reportado em Wuhan, província de Hubei, China, o primeiro caso da doença denominada COVID-19, causada pelo coronavírus da síndrome respiratória aguda grave 2 (SARS-CoV-2), levando a um surto da doença na província. Este surto rapidamente se espalhou para as províncias próximas, acarretando em uma epidemia de escala estadual, o que futuramente evoluiria para o nível nacional. Em janeiro de 2020, a organização mundial da saúde (OMS) declarou que a epidemia da doença era uma emergência de saúde pública de interesse internacional (PHEIC¹) [1] e, em março do mesmo ano, a epidemia foi declarada como uma pandemia, uma vez que a mesma alastrou-se por diversos países, de continentes distintos. Até abril de 2021, o número de casos reportados por 237 países e territórios excedeu 140 milhões, resultando em mais de 3 milhão de óbitos até o presente momento [2].

Coronavírus é um grupo de vírus que causam doenças em mamíferos e aves, mais especificamente. Causam infecções no trato respiratório que podem ser desde brandas até fatais. Infecções leves em humanos causam sintomas parecidos com os da gripe comum, enquanto casos mais graves podem causar a síndrome respiratória aguda grave (SARS), a síndrome respiratória do oriente médio (MERS) e a doença do coronavírus 2019 (COVID-19). Este vírus espalha-se por meio de aerossóis e gotículas de saliva produzidas por uma pessoa infectada. Ao espirrar, tossir, falar ou até mesmo respirar, os SARS-CoV-2 são carregados por estes veículos de transmissão. Sintomas comuns da doença incluem febre, tosse, coriza, dor de garganta, fadiga, dispnéia, distúrbios gastrintestinais, alteração do paladar, perda de olfato e diminuição do apetite [3–5]. Complicações podem levar à pneumonia, à síndrome

¹Public Health Emergency of International Concern

da angústia respiratória, podendo levar ao óbito.

Em 1927, W. O. Kermack e A. G. McKendrick criaram uma hipótese para a predição do número de casos de uma doença infectocontagiosa [6]. Isso culminou em um modelo epidemiológico que descreve a dinâmica de transmissão da enfermidade em uma população ao longo do tempo, consolidando o que seria a fonte para o modelo epidemiológico conhecido hoje como SIR. Esse modelo discrimina a população em três compartimentos, os suscetíveis à contração da doença (S), os infectados pela mesma (I) e os removidos/recuperados (R), e leva em consideração taxas de infecção e recuperação. Nesse modelo, a dinâmica de indivíduos de cada compartimento é regida por uma equação diferencial. Suas publicações subsequentes trataram de modelos para doenças endêmicas e epidemias recorrentes [7, 8], levando em conta taxas de natalidade, migração e mortalidade com respeito aos modelos anteriores. A partir desses estudos, emergiram modelos epidemiológicos mais sofisticados. Esses levam em conta outras particularidades presentes na natureza e/ou na sociedade, adicionados outros compartimentos e parâmetros. Essa tentativa de mimetizar a natureza complexa da realidade da melhor forma possível aumenta a complexidade das equações que descrevem o sistema.

Para realizar este estudo, a presente monografia está estruturada da seguinte forma: nos Capítulos 1 e 2, é realizada uma revisão com respeito à dinâmica dos modelos epidemiológicos encontrados na literatura, tal qual o modelo de Kermack-McKendrick e suas variantes, pertinentes às características da COVID-19, contemplando abordagens deterministas e estocásticas. Em seguida, no Capítulo 3, são introduzidos os dados utilizados neste estudo, cuja escala destes varia de estadual, para o caso do Brasil, a global, desde as primeiras ocorrências reportadas para a então declarada pandemia. À posteriori, no Capítulo 4, há a união entre os Capítulos 1 e 3, por meio da modelagem destes dados. São realizados diversos ajustes para o modelo SIR usual e com a presença de dinâmicas vitais para os países com maior número de casos registrados e alguns estados brasileiros. No Capítulo 4.2 apresentam-se as conclusões a respeito dos ajustes realizados e uma comparação entre os modelos adotados. Finaliza-se o trabalho com as referências bibliográficas utilizadas neste trabalho, seguido pelo Apêndice A, apresentando algumas tabelas com os casos totais reportados mundialmente e nacionalmente.

Modelos Compartimentais em Epidemiologia

O intuito da modelagem matemática de sistemas dinâmicos está atrelado à compreensão de como os mesmos evoluem conforme o passar do tempo. Dentre os vários ramos originados dos sistemas dinâmicos, há aquele de importância fundamental para o estudo dos organismos vivos e a descrição de seus comportamentos, chamado de matemática biológica. Desde meados do século XIII, a matemática já estava sendo empregada no estudo dos organismos vivos, por Leonardo Fibonacci, quando o mesmo empregou a série de Fibonacci para calcular o crescimento populacional de coelhos [9]. Em seu trabalho, ele assumiu que cada casal de coelhos nesta população, após nascerem, demorariam um certo tempo para amadurecerem. Após este período, o casal geraria outro casal de coelhos todo mês e assim por diante. A Figura 1.1 descreve este processo.

Posteriormente, em 1662, John Graunt publicou o livro *Natural and Political Observations made upon the Bills of Mortality*¹, um dos trabalhos pioneiros no emprego de uma abordagem quantitativa para a descrição de doenças humanas e mortes causadas por estas [10]. John analisou os boletins mortuários semanais das paróquias de Londres, listando os números e as causas de morte dos indivíduos. Nele, Graunt discutiu sobre uma variedade de problemas demográficos da Grã-Bretanha do século XVII, com quase 20 anos de dados acumulados destes boletins semanais, em que separou todas as 229.250 mortes em uma tabela contendo a causa da morte, número de mortos e o risco destas causas (Tabela 1.1).

No século XVIII, o físico e matemático Daniel Bernoulli, fazendo uso de uma

¹Observações naturais e políticas feitas sobre os boletins de mortalidade. Tradução do autor.

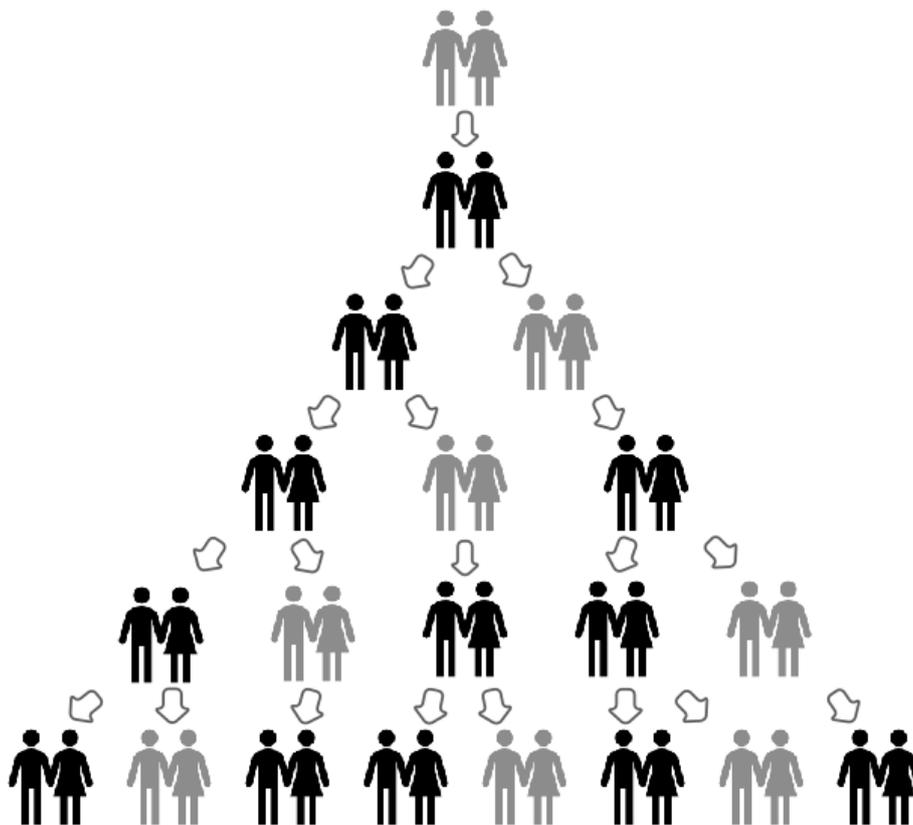


Figura 1.1: O problema do crescimento populacional, por Fibonacci, apresentado em seu livro *Liber Abaci*. O casal em cinza representa aqueles que ainda não atingiram a maturidade, e que devem aguardar uma unidade de tempo para amadurecerem. Porém, o casal em preto representa aqueles que atingiram a maturidade, e que geram um par de descendentes a cada unidade de tempo. O número de casais em cada linha forma os elementos da sequência de Fibonacci. Imagem feita pelo autor.

	Causas	Óbitos	Risco
1.	Candidíase, convulsão, raquitismo, dentes e vermes; abortivos, bebês, aumento do fígado.	71124	0,310
2.	Doenças crônicas: febre, febre aguda e tuberculose.	68271	0,298
3.	Doenças agudas e diversas.	49505	0,216
4.	Peste bubônica.	16384	0,071
5.	Variola, variola suína, sarampo e verminoses.	12210	0,053
6.	Doenças notórias: apoplexia, gota, lepra, paralisia, cálculo renal e estranguria, etc.	5547	0,024
7.	Câncer, fístula, feridas, úlcera, inflamação, prurido, escrófula, cisto.	3320	0,014
8.	Casualidades: afogamento, acidente, assassinato.	2889	0,013

Tabela 1.1: Tabela de óbitos agrupada por Graunt (1662), ao longo dos anos 1629 – 1636 e 1647 – 1658, com 81 causas de morte diferentes, para os 229.250 óbitos registrados. A tabela está organizada em grupos, baseado no número de óbitos ocasionados devido às causas e pelo risco (chance) de ser acometido pelas mesmas. Tabela traduzida da referência [11], com a tradução feita pelo autor.

abordagem mais teórica com relação ao estudo dos efeitos das doenças, como a

varíola², amplamente disseminada em diversas partes da Europa na época. Seu estudo, *Essai d'une nouvelle analyse de la mortalité causée par la petite vérole et des avantages de l'inoculation pour la prévenir*³, apresentado à Academia Francesa Real de Ciências de Paris, promovia uma forma de diminuir a letalidade da doença. Em seu trabalho, Bernoulli demonstrou a partir de cálculos de um modelo que propôs, que a inoculação com vírus ainda vivos obtidos de um paciente com quadro leve de varíola usualmente conferiria imunidade ao indivíduo, reduzindo desta forma a letalidade da doença.

Posteriormente, no início do século XX, William Hamer realizou contribuições significativas para a modelagem matemática de doenças infecciosas, devido a sua busca por explicações para a recorrência de sarampo na Inglaterra [13]. Sua contribuição foi de grande importância em detrimento de ser a primeira aplicação da lei de ação em massa para a modelagem de surtos epidêmicos, assunto que será tratado na próxima seção.

Ainda neste século, Sir Ronald Ross realizou trabalhos pioneiros com respeito à malária, descobrindo que esta transmite-se entre humanos e mosquitos. Devido à importância do estudo, Ross foi agraciado com o prêmio Nobel em fisiologia ou medicina, de 1902. Em 1911, publicou o livro *The prevention of Malaria*⁴. Nessa obra, ele desenvolveu modelos matemáticos para a transmissão da doença e deduziu um parâmetro chamado de quantidade limite, conhecido hoje como número de reprodução básico. Esse número determina se haverá um surto epidêmico em uma população ou não.

Apesar de tudo, a modelagem matemática para a descrição da dinâmica de doenças não era uma área de pesquisa reconhecida na época. Uma teoria mais aceita só foi surgir em 1927, trazendo o devido reconhecimento para esta área de estudo, por meio dos trabalhos de William Ogilvy Kermack e Anderson Gray McKendrick. Em seu artigo *A contribution to the mathematical theory of epidemics*⁵ eles descrevem um modelo que futuramente seria conhecido como modelo de Kermack-McKendrick.

Nos tempos atuais existem modelos mais complexos, que utilizam a teoria de redes complexas ou modelos baseados em indivíduos (IBM⁶), que atribuem pesos às

²Doença infecciosa, transmitida no contato entre pessoas ou com objetos contaminados, declarada erradicada em 1980. Encontram-se presentes como sintomas febre, vômito, úlceras na boca e erupções cutâneas na pele. Com o passar do tempo, estas erupções evoluíam para bolhas, com uma depressão em seu centro. Tais bolhas formavam crostas e desprendiam-se da pele, formando cicatrizes. A letalidade da doença foi estimada em 30% [12].

³Teste de uma nova análise de mortalidade causada pela varíola e as vantagens de inoculação para preveni-la. Tradução do google tradutor.

⁴A prevenção da malária.

⁵Uma contribuição para a teoria matemática das epidemias

⁶*Individual-based models*

relações sociais. Contudo, o modelo de Kermack-McKendrick ainda possui enorme relevância, devido à simplicidade e elegância com que é construído. Essa maleabilidade permite facilmente a adição de certas peculiaridades inerentes às doenças ou até mesmo às populações estudadas na dinâmica de transmissão de doenças, por meio do conjunto de equações diferenciais que fundamentam este modelo.

1.1 Modelos Deterministas em Epidemiologia

Um modelo determinista é aquele em que todo conjunto de estados é determinado unicamente pelos parâmetros do modelo e pelo estado inicial das variáveis. Todos os modelos apresentados no trabalho serão não-lineares e dinâmicos.

Para o estudo de modelos epidemiológicos deterministas, leva-se em conta a lei de ação em massa. Essa lei diz que, se os indivíduos de uma população são misturados homogeneamente, a taxa de interação entre dois subconjuntos diferentes da população é proporcional ao produto dos tamanhos de cada subconjunto, chamado de compartimento ou classe, respectivamente.

É plausível que, em qualquer população, diversos processos ocorram concomitantemente, de forma que os efeitos nos tamanhos dos compartimentos possam ser considerados aditivos. Para a modelagem epidemiológica, esta lei é aplicada às taxas de transição dos indivíduos entre dois compartimentos interagentes. Um exemplo seria membros suscetíveis que entram em contato com indivíduos infectados, tornando-se assim infectados também. Outro processo comum em matemática epidemiológica que ocorre simultaneamente está relacionado à transição de infectados para recuperados, uma vez que a doença é curada.

Grande parte dos modelos epidemiológicos deterministas dividem a população hospedeira em compartimentos, em que cada um contém indivíduos idênticos em termos de status com respeito à doença em questão. Após a separação da população em compartimentos, necessita-se de um conjunto de equações que especifiquem suas dinâmicas. A variável independente nesses modelos compartimentais é o tempo t , assim, as taxas de transferência entre os compartimentos são expressas, matematicamente, como derivadas dos tamanhos dos mesmos com respeito ao tempo.

1.1.1 Modelo de Kermack-McKendrick

Ao formular os modelos compartimentais, assume-se que o número de membros em um compartimento é uma função diferenciável no tempo, o que é uma aproximação razoável se houverem muitos membros em um compartimento.

Os modelos compartimentais mais básicos que descrevem a transmissão de doenças infecciosas foram formulados em uma sequência de artigos [6–8]. O primeiro descreve os modelos epidêmicos. O que é comumente chamado de modelo epidêmico de Kermack-McKendrick ou modelo SIR é um caso à parte do modelo geral introduzido em seu artigo. O modelo geral inclui uma dependência dos parâmetros com a idade dos indivíduos, isto é, os indivíduos não são idênticos em termos de status com relação à doença. Estes modelos que levam em conta isso tornaram-se muito úteis no estudo do vírus da imunodeficiência humana (HIV).

O caso particular do modelo proposto por Kermack e McKendrick, em 1927, é descrito como um modelo epidemiológico compartimental. No modelo em questão, têm-se os seguintes compartimentos:

- *Suscetíveis (S)* - Indivíduos não imunes ao agente infeccioso que podem ser infectados se expostos.
- *Infectados (I)* - Indivíduos infectados podendo transmitir a infecção para indivíduos suscetíveis com quem possuem contato.
- *Removidos (R)* - Indivíduos imunes à infecção que não afetam a dinâmica de transmissão quando entram em contato com outros indivíduos.

Nesse cenário, o tamanho total da população hospedeira é $N = S + I + R$.

Após essa divisão, faz-se necessário um conjunto de equações que especificam como o tamanho desses compartimentos alteram-se com o passar do tempo. As soluções dessas equações carregam consigo o tamanho dos compartimentos em um tempo t . Portanto, o sistema de equações que rege este modelo é descrito por

$$\frac{dS}{dt} = -\beta SI, \tag{1.1.1}$$

$$\frac{dI}{dt} = \beta SI - \gamma I, \tag{1.1.2}$$

$$\frac{dR}{dt} = \gamma I, \tag{1.1.3}$$

em que β e γ são as taxas de transmissão e de recuperação, respectivamente, cujo período médio de infecção é $1/\gamma$. A Figura 1.2 descreve o fluxograma deste modelo baseado nas seguintes premissas:

- (i) Um indivíduo infectado da população realiza contato o suficiente para transmitir a infecção à uma taxa βN para outrem por unidade de tempo, com N representando o tamanho total da população.

- (ii) Infectados são exonerados da classe infectiva com taxa γI por unidade de tempo.
- (iii) Não há entrada ou saída de indivíduos da população total, exceto por morte ocasionada pela doença.

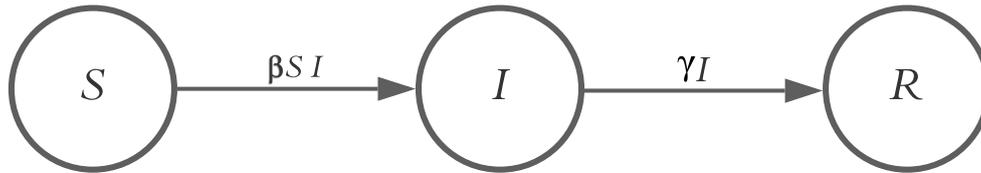


Figura 1.2: Fluxograma do modelo compartimental SIR, demonstrando a saída de indivíduos do compartimento dos suscetíveis e dos infectados, relacionada por meio das taxas de transmissão β , e recuperação γ da doença, respectivamente. Fonte: o autor, 2021.

De acordo com (i), a probabilidade de que um contato aleatório de um indivíduo pertencente à classe dos infectados I ser com um indivíduo pertencente à classe dos suscetíveis S é $\frac{S}{N}$. Assim, o número de novas infecções por unidade de tempo por infectado é $(\beta N)(\frac{S}{N})$. Portanto, a taxa de novas infecções é $(\beta N)(\frac{S}{N}I) = \beta SI$.

De forma alternativa, pode-se argumentar que, a probabilidade de um indivíduo suscetível realizar contato com um indivíduo infectado é $\frac{I}{N}$. Logo, a taxa de novas infecções por suscetível é denotada por $(\beta N)(\frac{I}{N})$, fornecendo a taxa de novas infecções, descrita por $(\beta N)(\frac{I}{N}S) = \beta SI$. Para um resultado normalizado do tamanho dos compartimentos, divide-se a taxa de novas infecções pelo tamanho total da população, resultando em $\beta SI/N$.

Para a compreensão da premissa (ii), considera-se o conjunto de todos os membros que estavam infectados em um tempo t e denota-se $u(s)$ o número destes que continuam infectados s unidades de tempo após serem infectados. Se uma fração de γ destes membros é exonerada da classe infectiva por unidade de tempo, então, a taxa de variação de u com relação a s é

$$\frac{du}{ds} = u' = -\gamma u, \quad (1.1.4)$$

e a solução para esta equação diferencial elementar é

$$u(s) = u(0) e^{-\gamma s}. \quad (1.1.5)$$

Portanto, a fração de infectados que permanecem assim s unidades de tempo após se tornarem infectivos é $e^{-\gamma s}$. Logo, o módulo deste período infeccioso é distribuído exponencialmente com média $\int_0^{\infty} e^{-\gamma s} ds = \frac{1}{\gamma}$.

Neste modelo, R é determinado uma vez que S e I são conhecidos. Desta forma, é possível obter a equação que descreve a derivada de R em função do tempo no modelo, resultando no seguinte sistema de duas equações

$$S' = -\beta SI, \quad (1.1.6)$$

$$I' = (\beta S - \gamma)I, \quad (1.1.7)$$

cujos termos estão representados com uma “linha” (S' e I') representam a derivada temporal da variável em questão.

Não era conhecido que o sistema poderia ser resolvido analiticamente, porém, sabia-se que o modelo só faz sentido enquanto $S(t)$ e $I(t)$ forem não-negativos. Logo, se qualquer um destes dois atingir zero, é considerado que o sistema encerrou-se. Portanto, é tido que $S' < 0$ para todo t e $I' > 0$ se, e somente se, $S > \frac{\gamma N}{\beta}$. Então, I aumenta contanto que $S > \frac{\gamma N}{\beta}$, mas, como S decresce para todo t , I decresce até aproximar-se de zero. Se $S(0) < \frac{\gamma N}{\beta}$, I decresce à zero, enquanto que, se $S(0) > \frac{\gamma N}{\beta}$, I primeiramente aumenta a um máximo quando $S = \frac{\gamma N}{\beta}$ e, posteriormente, decresce a zero.

Com isto, são obtidas as informações se haverá uma epidemia ou não, por meio de uma quantidade chamada *número de reprodução básico* \mathcal{R}_0 . Esta quantidade \mathcal{R}_0 é dada por $\frac{\beta S(0)}{\gamma N}$, e, caso $\mathcal{R}_0 < 1$ a infecção acabará por cessar-se, enquanto para $\mathcal{R}_0 > 1$, certamente haverá uma epidemia.

A definição de \mathcal{R}_0 é o valor médio do número de infecções secundárias causadas por um único infectado ao incorporá-lo em uma população de suscetíveis de tamanho $J \approx S(0)$.

De modo a determinar I , é necessário realizar os seguintes passos

$$\frac{I'}{S'} = \frac{dI}{dS} = \frac{(\beta S - \gamma)I}{-\beta SI} = -1 + \frac{\gamma}{\beta S}. \quad (1.1.8)$$

Integrando para encontrar o plano de fase, têm-se

$$I = -S + \frac{\gamma}{\beta} \ln S + c = I(0) + S(0) - S(t) + \frac{\gamma}{\beta} \ln \frac{S}{S(0)}. \quad (1.1.9)$$

O número máximo de infectados ocorre para um valor particular de S . Para encontrá-lo, basta derivar a Eq. (1.1.9) com relação a S e igualá-la a zero. Seja S^*

o valor que maximiza o número de infectados:

$$\left. \frac{dI}{dS} \right|_{S=S^*} = -1 + \frac{1}{S^*} \frac{\gamma}{\beta} = 0 \quad (1.1.10)$$

de maneira que

$$S^* = \frac{\gamma}{\beta}. \quad (1.1.11)$$

Portanto, o número máximo de infectados é

$$I(S^*) = I_{\text{máx}} = -\frac{\gamma}{\beta} + \frac{\gamma}{\beta} \ln \left(\frac{\gamma}{\beta} \right) + S(0) + I(0) - \frac{\gamma}{\beta} \ln S(0). \quad (1.1.12)$$

Apesar de ser uma solução exata, I está escrito como função de S ao invés de t . Acreditava-se que até mesmo para um modelo simples como o SIR, não era possível obter $I(t)$ analiticamente, porém, estudos recentes apontam o contrário [14, 15], apresentando soluções complicadas para o modelo. Inicialmente, o número de infectados cresce exponencialmente, pois a equação que descreve I pode ser aproximada para

$$I' = (\beta J - \gamma)I, \quad (1.1.13)$$

com I sendo descrito por

$$I(t) = I_0 e^{(\beta J - \gamma)t}, \quad (1.1.14)$$

e a taxa de crescimento inicial, isto é, a constante presente na exponencial da Eq. (1.1.14), é descrita por

$$r = \beta J - \gamma = \gamma(\mathcal{R}_0 - 1). \quad (1.1.15)$$

É possível determinar experimentalmente esta taxa de crescimento inicial r no começo de uma epidemia. Portanto, uma vez que J e γ podem ser medidos, estima-se β da seguinte forma

$$\beta = \frac{r + \gamma}{J}. \quad (1.1.16)$$

Entretanto, é possível também que esta estimativa não seja muito precisa, devido à incompletude de dados e às subnotificações de casos. A acurácia é afetada de forma ainda mais significativa para surtos de uma doença anteriormente desconhecida, visto que os casos iniciais tendem a ser diagnosticados erroneamente.

Kermack e McKendrick encontraram uma solução aproximada de $I(t)$ para o modelo SIR básico, porém, esta só é válida para os momentos iniciais de um surto. Desta forma, ao invés de buscar uma equação explícita para $I(t)$, pode-se obter uma aproximação numérica da solução usando o método de Euler. Ao longo de um intervalo de tempo suficientemente pequeno Δt , pode-se realizar a aproximação $\frac{dS}{dt} \simeq \frac{\Delta S}{\Delta t}$, em que $\Delta S = S(t + \Delta t) - S(t)$. A solução para o número de suscetíveis S em um tempo Δt no futuro é

$$S(t + \Delta t) = S(t) - \beta S(t)I(t)\Delta t. \quad (1.1.17)$$

De modo similar, é possível aproximar o número de infectados e recuperados em $t + \Delta t$ como

$$I(t + \Delta t) = I(t) + [\beta S(t)I(t) - \gamma I(t)]\Delta t, \quad (1.1.18)$$

$$R(t + \Delta t) = R(t) + \gamma I(t)\Delta t. \quad (1.1.19)$$

As equações (1.1.17)-(1.1.19) juntas providenciam um esquema para aproximar as soluções do modelo SIR básico. As Figuras 1.3 e 1.4 descrevem graficamente o comportamento deste modelo epidemiológico. A simulação da Figura 1.3 foi realizada utilizando a linguagem de programação *Python*, por meio das soluções numéricas [16] do conjunto de equações diferenciais do modelo SIR, que foram agregadas para todos os instantes de tempo t e expostas graficamente na imagem.

1.1.2 Modelo de Kermack-McKendrick com taxa de contato generalizada

Pode-se generalizar a premissa (i), apresentada na seção 1.1.1, de modo a obter uma taxa de contato que seja uma função crescente do tamanho populacional. Para isso, substitui-se a premissa (i) pelo fato de que um membro médio da população realiza $C(N)$ contatos por unidade de tempo com $\frac{dC(N)}{dt} \geq 0$ e define-se

$$\beta(N) = \frac{C}{N}. \quad (1.1.20)$$

Assume-se também que $\frac{d\beta(N)}{dt} \leq 0$ para expressar a saturação no número de contatos.

Com o tamanho total da população fazendo-se presente no modelo, o mesmo é incluso no sistema de equações que descreve o modelo. Assim, é necessário distin-

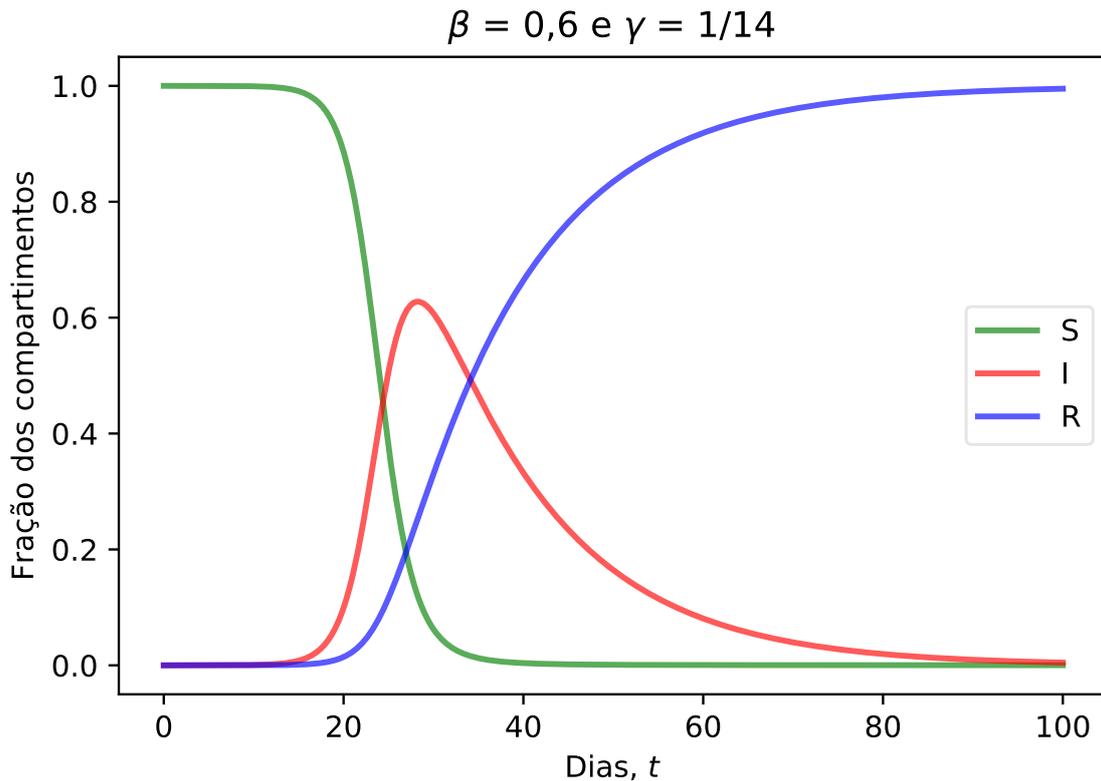


Figura 1.3: Simulação do modelo SIR, fundamentada nas premissas (i), (ii) e (iii) e no conjunto de equações diferenciais (1.1.1)-(1.1.3). Introduce-se um único infectado em uma população totalmente suscetível como condição inicial e mantém-se a taxa de transmissão $\beta = 0,6$ e a taxa de recuperação $\gamma = 1/14$ constantes durante toda a simulação. Os compartimentos estão representados pelas curvas, com verde representando os indivíduos suscetíveis, vermelho os infectados e azul os recuperados. Fonte: o autor, 2021.

guir os membros da população que tornaram-se parte dos recuperados R e os que faleceram. Se f é a fração dos membros que partem da classe infecciosa com γI em um tempo t para o compartimento dos recuperados R , a letalidade é fornecida pela fração restante $(1 - f)$. Essa taxa quantifica a proporção de indivíduos mortos em detrimento da doença. Usando as variáveis S , I e N em conjunto com $N = S + I + R$, obtendo-se o seguinte conjunto de equações

$$S' = -\beta(N)SI, \tag{1.1.21}$$

$$I' = \beta(N)SI - \gamma I, \tag{1.1.22}$$

$$R' = f\gamma I, \tag{1.1.23}$$

$$M' = (1 - f)\gamma I = -N', \tag{1.1.24}$$

em que M' condiz com o fluxo de indivíduos que adentram o compartimento dos

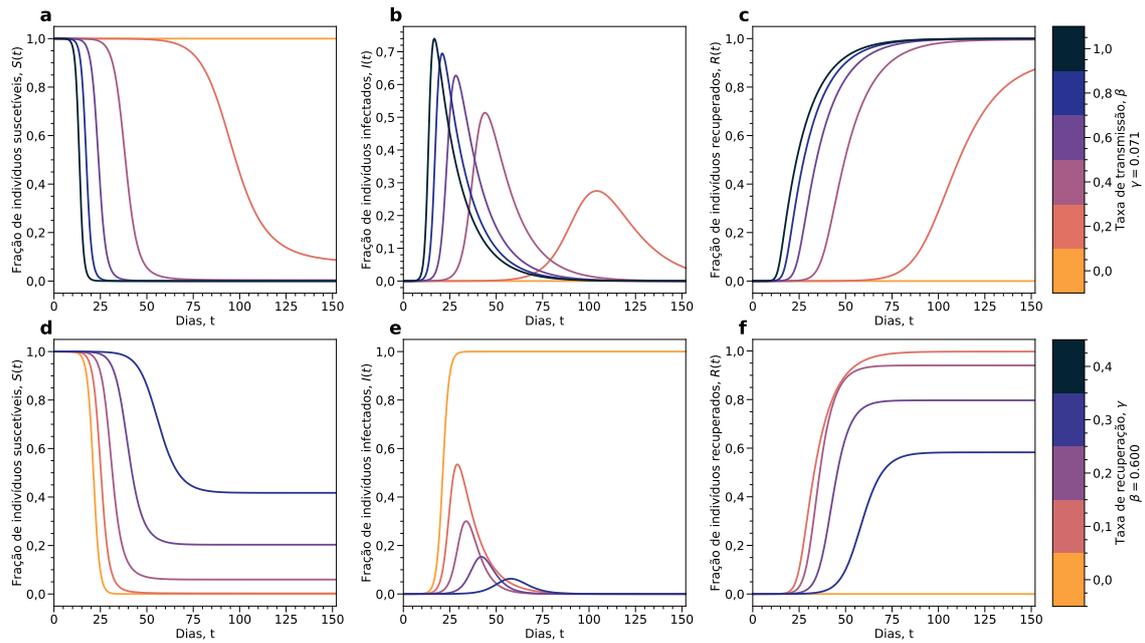


Figura 1.4: Gráficos do modelo SIR, fundamentados nas premissas (i), (ii) e (iii) e no conjunto de equações diferenciais (1.1.1)-(1.1.3), com um conjunto de valores para os parâmetros β e γ . Nos gráficos **a**, **b** e **c**, o valor da taxa de recuperação é mantido constante, representando um período médio de infecção de 14 dias, para demonstrar como o aumento ou decréscimo na taxa de transmissão afeta a dinâmica dos compartimentos. O mesmo é realizado para os gráficos **d**, **e** e **f**, porém, mantendo $\beta = 0,6$, o que representa uma média de 0,6 novas infecções causadas para cada contato entre um infectado e um suscetível, variando os valores de γ . Fonte: o autor, 2021.

óbitos. É importante notar que esta variante do modelo SIR leva em conta as fatalidades da doença, cuja variação com relação ao tempo é $M' = -N'$. A Figura 1.5 descreve o fluxo de indivíduos da população ao longo dos compartimentos do modelo SIRM. Para $f = 1$, não há óbitos provenientes da doença e recupera-se o modelo de Kermack-McKendrick usual, isto é, o modelo SIR. Dessa forma, usa-se o número de indivíduos inicial $J = N(0) \approx S(0)$ mencionado na seção 1.1.1, reduzindo o modelo ao de Kermack-McKendrick com $\beta = \beta(J)$. De modo análogo ao caso anterior, pode-se determinar o número de reprodução básico para este modelo:

$$\mathcal{R}_0 = \frac{J\beta(J)}{N\gamma}, \quad (1.1.25)$$

possuindo o mesmo comportamento, qualitativamente, do modelo de Kermack-McKendrick, com respeito aos valores possíveis de \mathcal{R}_0 . Além de \mathcal{R}_0 , há também o número de reprodução dependente do tempo \mathcal{R}^* , representado pelo número de infecções secundárias causadas por um único indivíduo infectado na população para

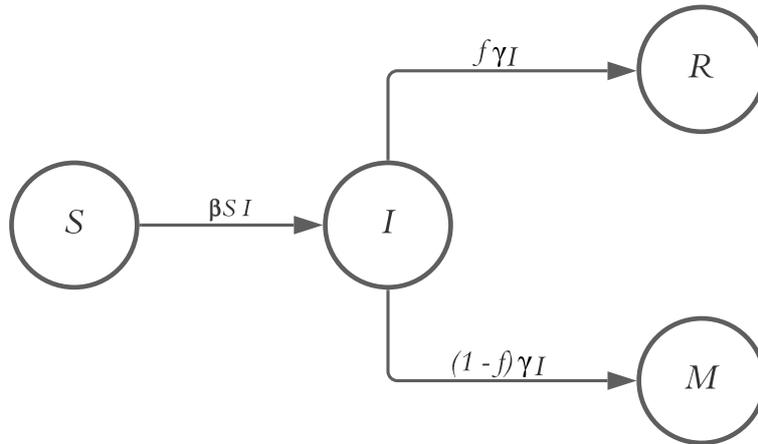


Figura 1.5: Fluxograma do modelo compartimental SIRM, apresentando o fluxo de indivíduos do compartimento dos infectados em direção aos compartimentos dos recuperados e dos mortos, descritos pela taxa de recuperação γ e pela letalidade $(1 - f)$ da doença, respectivamente. Fonte: o autor, 2021.

qualquer tempo t . Neste sentido, um infectado realiza $C(N) = N\beta(N)$ contatos por unidade de tempo, e a fração S/N destes contatos são realizados com indivíduos suscetíveis, gerando, desta forma, novas infecções. Portanto, este número de reprodução é dado por

$$\mathcal{R}^* = \frac{S\beta(N)}{N\gamma}. \quad (1.1.26)$$

Somando as equações (1.1.21) e (1.1.22), encontram-se

$$S' + I' = -\gamma I, \quad (1.1.27)$$

$$N' = -\gamma(1 - f)I. \quad (1.1.28)$$

Integrando-se ambas as equações com os limites de integração definidos em 0 e t , têm-se

$$S(t) + I(t) - S(0) - I(0) = \gamma \int_0^t I(s) ds, \quad (1.1.29)$$

$$N(t) - N(0) = -\gamma(1 - f) \int_0^t I(s) ds. \quad (1.1.30)$$

Ao dividir a expressão (1.1.29) pela expressão (1.1.30), de modo a retirar a integral

da equação, e tomando-se $N(0) = J = S(0) + I(0)$, é obtido

$$J - N(t) = (1 - f) [J - S(t) - I(t)] . \quad (1.1.31)$$

Ao tomar-se o limite desta expressão com o tempo $t \rightarrow \infty$, $S(t)$ e $N(t)$ decrescem até os limites S_∞ e N_∞ , enquanto $I(t) \rightarrow 0$. Nesse limite

$$J - N_\infty = (1 - f) [J - S_\infty] . \quad (1.1.32)$$

A equação (1.1.32) fornece algumas informações importantes, tal qual $J - N_\infty$, que representa a mudança no tamanho total da população, mediante às mortes ao longo da epidemia, enquanto $J - S_\infty$ representa o número de casos acumulados da doença ao longo do surto epidêmico.

1.1.3 Modelo de Kermack-McKendrick com dinâmicas vitais

É possível expandir o modelo SIR de modo a incluir efeitos demográficos à população, como a taxa de natalidade B e a taxa de mortalidade natural μ . Nesse cenário, têm-se o sistema de equações

$$S' = BN - \beta SI - \mu S , \quad (1.1.33)$$

$$I' = \beta SI - \gamma I - \mu I , \quad (1.1.34)$$

$$R' = \gamma I - \mu R . \quad (1.1.35)$$

Para epidemias de curto período, pode-se assumir que o tamanho da população seja constante (dessa forma $BN = \mu N$, de modo que haja somente um único parâmetro novo nas equações (1.1.33) a (1.1.35)). Porém, a presença de dinâmicas vitais em modelos epidemiológicos torna-se mais relevante para o estudo de epidemias de longa duração ou recorrentes, e.g., sarampo e doenças similares. Para curtos períodos de tempo, a manifestação das taxas B e μ são muito sutis. A Figura 1.6 apresenta o fluxo de indivíduos nos compartimentos do modelo, denotados pelo sistema de equações (1.1.33) a (1.1.35). Neste modelo, \mathcal{R}_0 sofre uma alteração em seu denominador por um fator μ [17]. Logo, este é determinado por

$$\mathcal{R}_0 = \frac{S\beta}{N(\gamma + \mu)} . \quad (1.1.36)$$

De maneira análoga ao proposto, usa-se o método de Euler para gerar soluções aproximadas das equações apresentadas. Em vista disso, ao longo de um intervalo de tempo Δt suficientemente pequeno, consideram-se as aproximações para o número

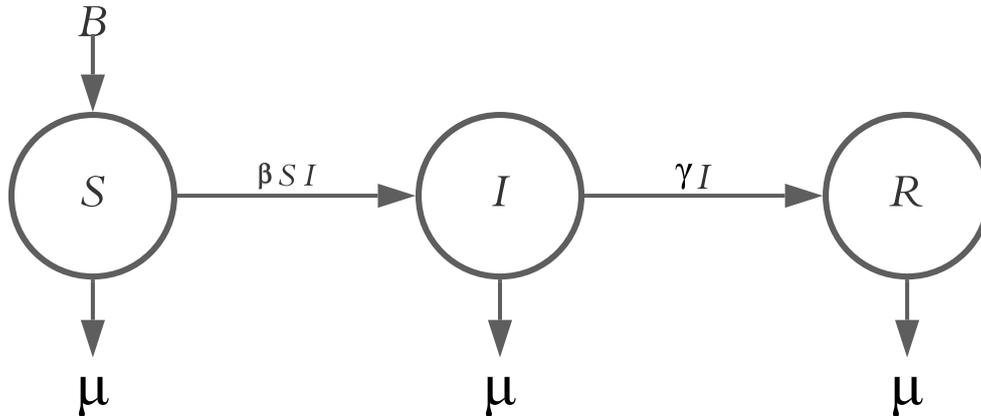


Figura 1.6: Fluxograma do modelo compartimental SIR. Nesse modelo, indivíduos suscetíveis são adicionados à população com uma taxa de natalidade B . Também é possível que indivíduos sejam exonerados de qualquer compartimento com uma taxa de mortalidade μ . Essa combinação garante a presença de dinâmicas vitais no modelo. Fonte: o autor, 2021.

de suscetíveis, infectados e recuperados em um tempo Δt no futuro

$$S(t + \Delta t) = S(t) + [BN - \beta S(t)I(t) - \mu S(t)]\Delta t, \quad (1.1.37)$$

$$I(t + \Delta t) = I(t) + [\beta S(t)I(t) - \gamma I(t) - \mu I(t)]\Delta t, \quad (1.1.38)$$

$$R(t + \Delta t) = R(t) + [\gamma I(t) - \mu R(t)]\Delta t. \quad (1.1.39)$$

1.1.4 Modelo SIRS

O modelo SIR e suas variações compreendem a dinâmica de infecção de doenças que, após sua fase infecciosa, conferem imunidade permanente aos indivíduos da população. Um exemplo desse cenário seria a varicela, também conhecida como catapora. Contudo, há também doenças que não atribuem este tipo de imunidade ao seu hospedeiro, ao invés disso, tem-se um período cujo sistema imunológico do indivíduo retém informações sobre o patógeno e sobre como neutralizá-lo. Este processo confere um período imune ao sujeito, uma vez que estas informações podem se perder com o tempo. Com isto, chega-se no modelo determinístico chamado de SIRS, caracterizado pelo parâmetro $\sigma = 1/T$, em que σ recebe o nome de taxa de imunidade, enquanto T representa o tempo médio de imunidade do sistema imunológico para esta doença. Portanto, o sistema de equações que descreve este modelo

é descrito por

$$S' = -\beta SI + \sigma R, \quad (1.1.40)$$

$$I' = \beta SI - \gamma I, \quad (1.1.41)$$

$$R' = \gamma I - \sigma R. \quad (1.1.42)$$

A Figura 1.7 descreve a dinâmica do modelo SIRS por meio de um fluxograma.

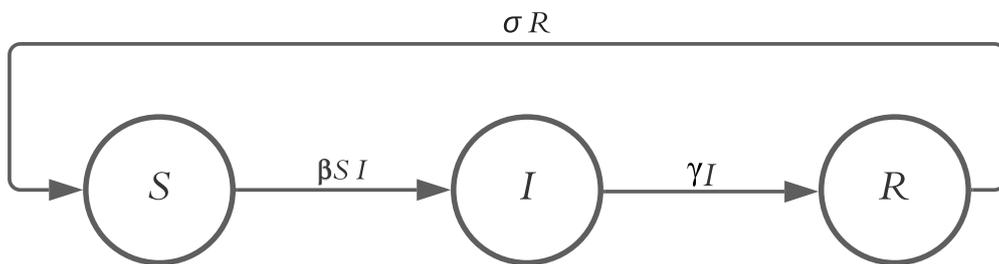


Figura 1.7: Fluxograma do modelo compartimental SIRS. Indivíduos suscetíveis aparecem continuamente na população em detrimento da taxa de imunidade σ . Isso pode ocasionar em uma configuração final com estabilidade assintótica, com $I \neq 0$. Fonte: o autor, 2021.

Uma primeira análise deste modelo evidencia o incremento contínuo de “combustível” ao sistema, de modo que os indivíduos permaneçam em um ciclo sem fim; sendo infectados, imunizados e, novamente, suscetíveis. Pelas Figuras 1.8 e 1.9, nota-se que o equilíbrio do sistema não será atingido para períodos curtos de tempo, uma vez que é visível a oscilação nas frações dos três compartimentos. A Figura 1.10 já evidencia o equilíbrio assintótico do modelo, em que a amplitude destas oscilações passam a ser cada vez menores, tendendo a um valor específico. Os pontos de equilíbrio deste sistema são descritos a partir de algumas manipulações das equações (1.1.41) e (1.1.42). Então, tomando o máximo da segunda equação, têm-se

$$\frac{dI}{dt} = \beta S_m I_m - \gamma I_m = 0 \quad (1.1.43)$$

$$\beta S_m - \gamma = 0 \quad (1.1.44)$$

$$S_m = \frac{\gamma}{\beta} \quad (1.1.45)$$

em que S_m representa o ponto de equilíbrio assintótico do compartimento suscetível [18]. Tomando o máximo da primeira equação, isolando R em $N = S + I + R$ e

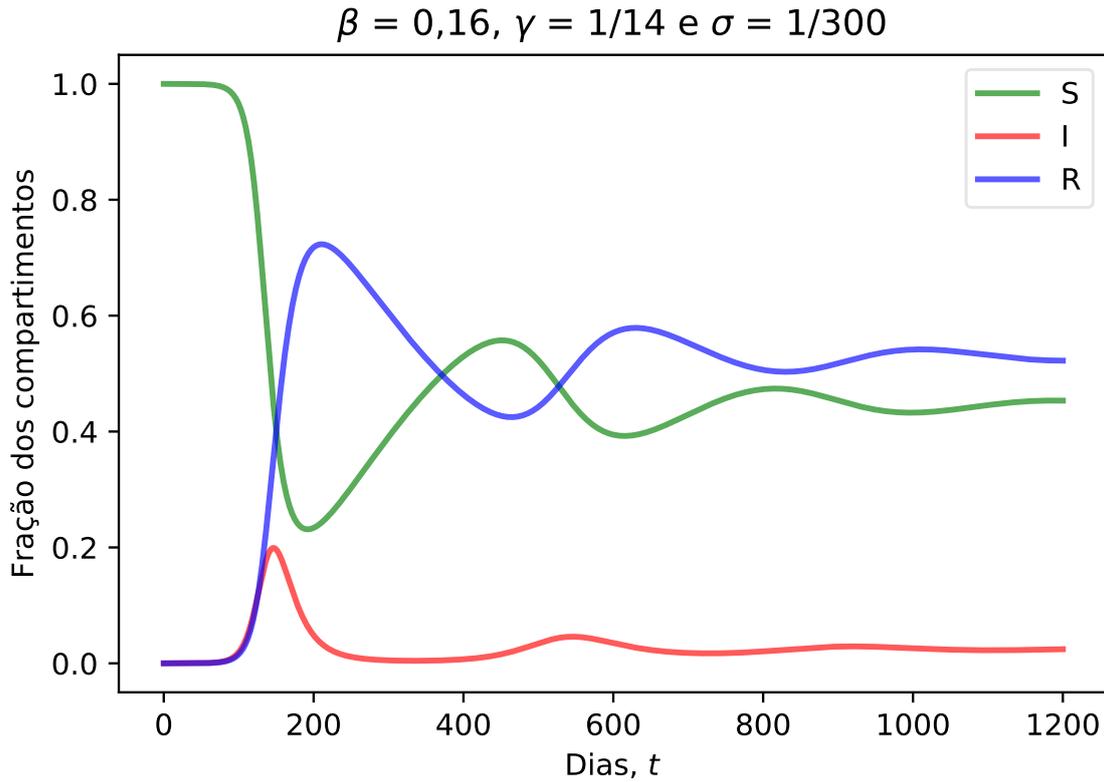


Figura 1.8: Simulação do modelo SIRS, em que novamente introduz-se apenas um indivíduo infectado numa população completamente suscetível e mantêm-se as taxas de transmissão $\beta = 0,16$, de recuperação $\gamma = 1/14$ e de imunidade $\sigma = 1/300$ constantes durante toda a simulação. Nota-se, nesta simulação, a presença de surtos epidêmicos recorrentes na população. Fonte: o autor, 2021.

utilizando o resultado de (1.1.45), têm-se

$$\frac{dS}{dt} = -\beta S_m I_m + \sigma R_m = 0 \quad (1.1.46)$$

$$-\frac{\beta\gamma}{\beta} I_m + \sigma(N - S_m - I_m) = 0 \quad (1.1.47)$$

$$-\gamma I_m + \sigma N - \frac{\sigma\gamma}{\beta} - \sigma I_m = 0 \quad (1.1.48)$$

$$I_m = \frac{\sigma(N - \frac{\gamma}{\beta})}{\gamma + \sigma}. \quad (1.1.49)$$

A partir disto, torna-se simples determinar R_m , sabendo que $N = R_m + I_m + S_m$. Esta estabilidade assintótica torna-se evidente nas Figuras 1.10 e 1.11, em que a última descreve a evolução temporal do sistema nos planos de fase $F(S, I)$, $G(S, R)$ e $H(I, R)$.

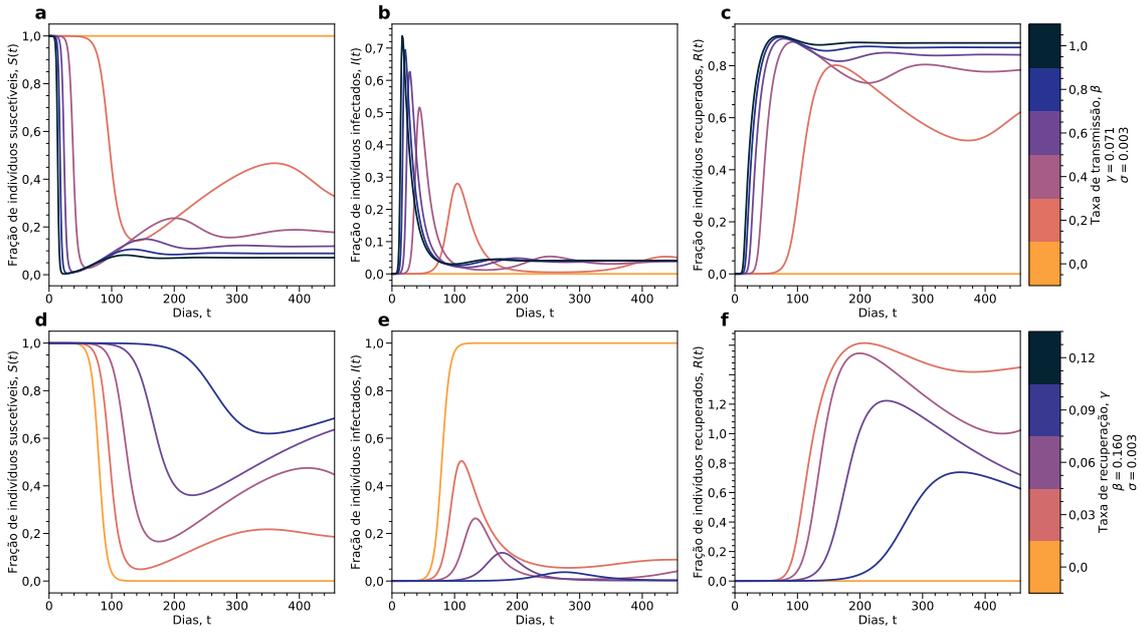


Figura 1.9: Gráficos do modelo SIRS que descrevem a dinâmica dos respectivos compartimentos, utilizando uma variedade de valores para os parâmetros β e γ , de modo a facilitar a visualização de como estes afetam a dinâmica do modelo. Os painéis **a** e **d** compreendem o compartimento dos suscetíveis, os painéis **b** e **e** representam o compartimento dos infectados, enquanto os últimos dois, **c** e **f**, correspondem ao compartimento dos recuperados. Fonte: o autor, 2021.

1.1.5 Modelo SEIR

O modelo SEIR é uma extensão do modelo SIR que inclui um compartimento E para indivíduos expostos à doença. Esses são os indivíduos que deixam de ser suscetíveis, porém ainda não transmitem a enfermidade. Dessa maneira, esses indivíduos atuam como intermediários entre os compartimentos S e I . A equação que descreve a dinâmica desse novo compartimento está associada a taxa de exposição κ . Esse número traz consigo o tempo médio de exposição $\tau = 1/\kappa$, equivalente ao período médio que um indivíduo leva para ascender da classe exposta E para a classe infectiva I . A população total passa a ser representada por $N = S + E + I + R$, e o sistema de equações que rege a dinâmica do sistema é

$$S' = -\beta SI, \quad (1.1.50)$$

$$E' = \beta SI - \kappa E, \quad (1.1.51)$$

$$I' = \kappa E - \gamma I, \quad (1.1.52)$$

$$R' = \gamma I. \quad (1.1.53)$$

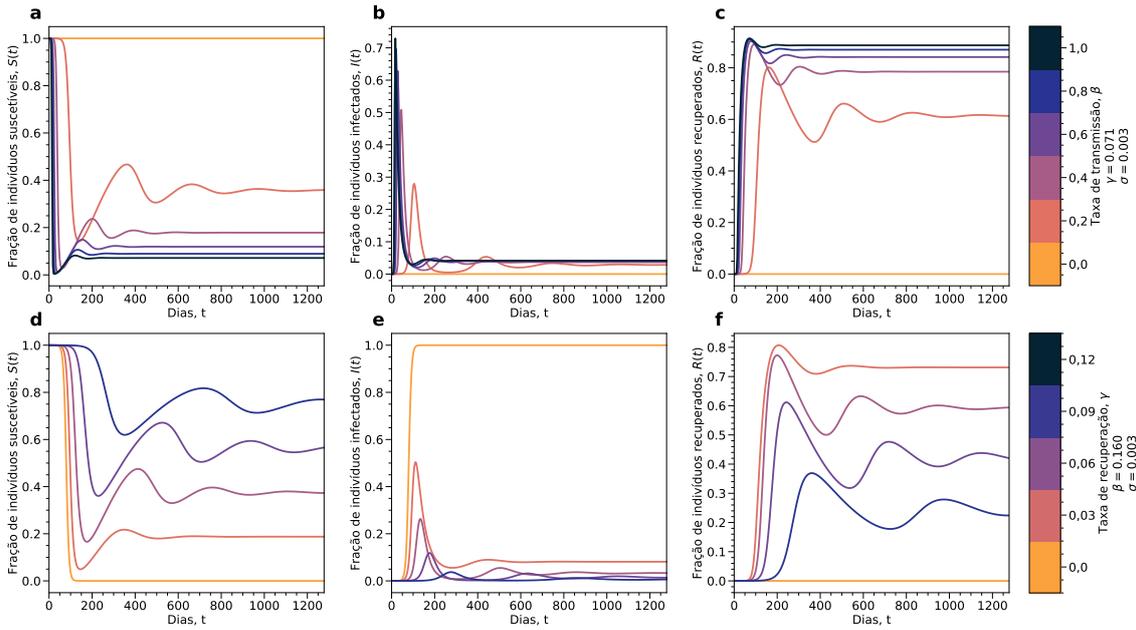


Figura 1.10: Gráficos do modelo SIRS para longos períodos de tempo. É notável que a variação na população, ao longo dos compartimentos, tende a ser cada vez menor com o passar do tempo. Isso ocorre até chegar ao seu ponto de equilíbrio $(S_m, I_m, R_m) = (\frac{\gamma}{\beta}, \frac{\sigma(N-\frac{\gamma}{\beta})}{\gamma+\sigma}, 1 - \frac{\gamma}{\beta} - \frac{\sigma(N-\frac{\gamma}{\beta})}{\gamma+\sigma})$, descrevendo assim, sua estabilidade assintótica. Fonte: o autor, 2021.

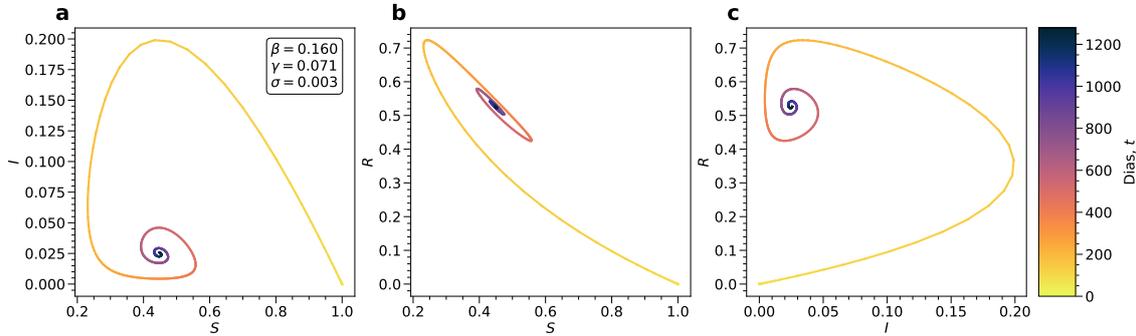


Figura 1.11: Plano de fase SIRS para longos períodos de tempo, em que a progressão temporal é denotada pelo degradê de cores, cuja escala é presente ao lado. Considerando os parâmetros $\beta = 0,160$, $\gamma = 0,071$ e $\sigma = 0,003$, têm-se que o ponto de equilíbrio do sistema está situado em $(S_m, I_m, R_m) = (0,44375, 0,02255, 0,5337)$. Fonte: o autor, 2021.

A Figura 1.12 descreve o fluxograma deste modelo enunciado pelo sistema de equações presente, enquanto as Figuras 1.13 e 1.14 apresentam uma simulação e um conjunto de soluções numéricas para o modelo SEIR.

O número \mathcal{R}_0 é descrito neste modelo assim como é descrito no modelo SIR. Portanto, encontra-se

$$\mathcal{R}_0 = \frac{S\beta}{N\gamma}, \tag{1.1.54}$$

visto que o período latente da doença não afeta o número de infecções secundárias causadas pela mesma. Assim como o modelo SIR, o modelo SEIR pode ser

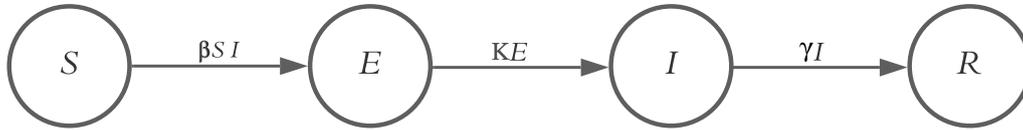


Figura 1.12: Fluxograma do modelo compartimental SEIR. A dinâmica do modelo é alterada de modo que o fluxo de indivíduos do compartimento dos Suscetíveis ao compartimento dos Infectados é mediado pelo compartimento dos Expostos, realizado por meio da taxa de exposição κ da doença. Fonte: o autor, 2021.

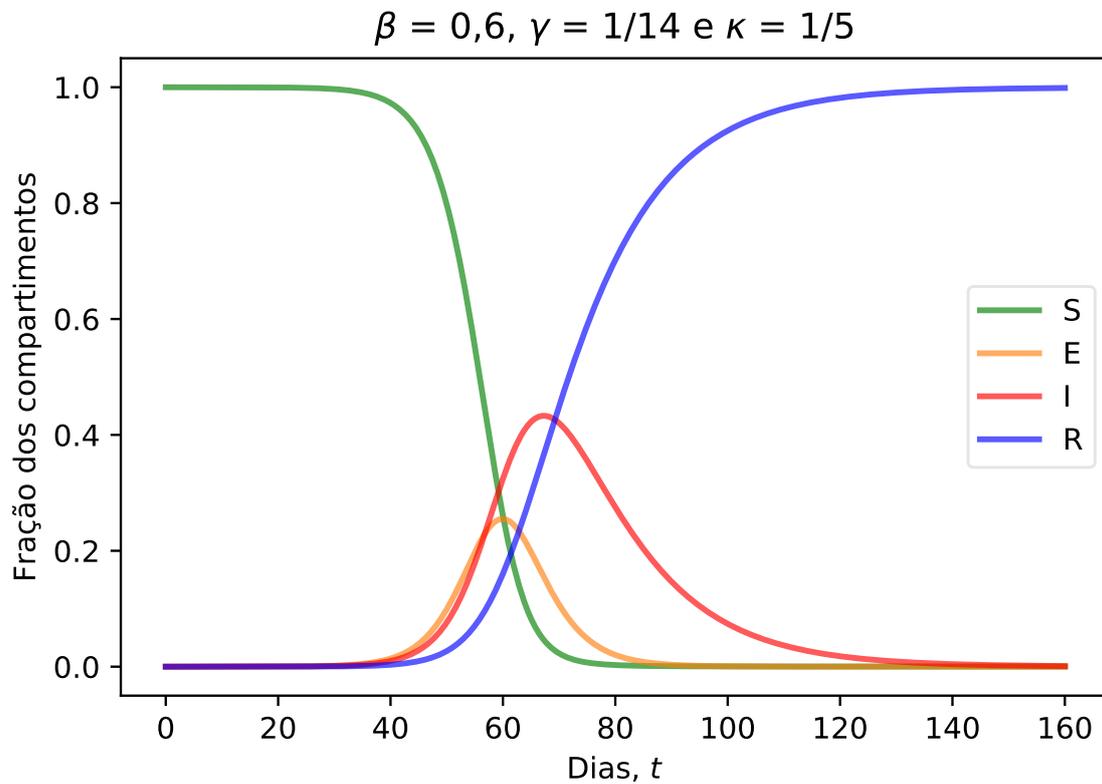


Figura 1.13: Simulação do modelo SEIR, em que introduz-se um único indivíduo infectado numa população completamente suscetível como condição inicial, mantendo as taxas de transmissão $\beta = 0,6$, recuperação $\gamma = 1/14$ e exposição $\kappa = 1/5$. Nesta simulação, há a introdução do compartimento dos indivíduos expostos (E), denotado pela cor laranja. Fonte: o autor, 2021.

expandido para incluir o compartimento dos óbitos. Para tal, assume-se que uma fração dos indivíduos infectados recuperam-se, enquanto a remanescente refere-se à fração dos indivíduos mortos pela doença. Portanto, assim como proposto na seção

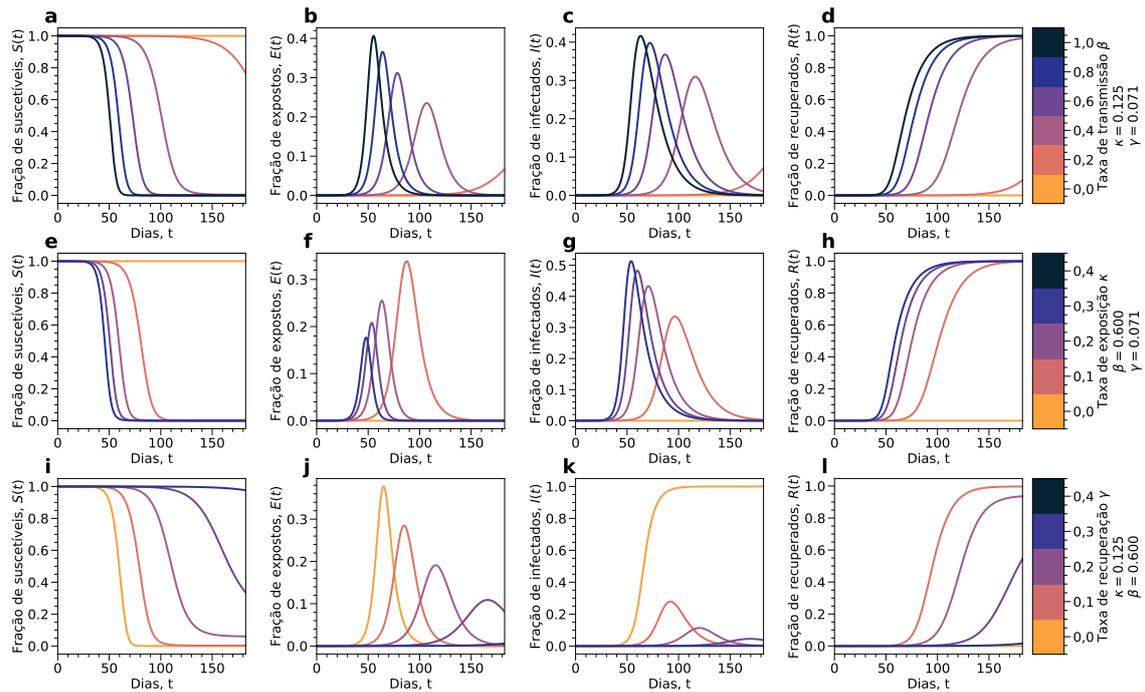


Figura 1.14: Gráficos do modelo SEIR, que demonstra a defasagem entre a curva dos expostos e a dos infectados. Para diferentes valores de κ , é visível que o aumento desta quantidade se traduz em uma curva mais baixa no compartimento dos expostos, uma vez que isto representa uma quantidade menor de dias dentro do mesmo. Fonte: o autor, 2021.

do modelo de Kermack-McKendrick com taxa de contato generalizada, o sistema de equações do modelo é modificado

$$S' = -\beta SI, \quad (1.1.55)$$

$$E' = \beta SI - \kappa E, \quad (1.1.56)$$

$$I' = \kappa E - \gamma I, \quad (1.1.57)$$

$$R' = f\gamma I, \quad (1.1.58)$$

$$M' = (1 - f)\gamma I. \quad (1.1.59)$$

A Figura 1.15 apresenta o fluxo de indivíduos da população ao longo dos compartimentos do modelo SEIRM. Quando $f = 1$ não há óbitos, recuperando-se o modelo SEIR usual. Quando um surto se estende por meses ou anos, é válido adicionar a presença de dinâmicas vitais ao modelo. Isso permite um fluxo de novos indivíduos no modelo, relacionados à taxa de natalidade B e à taxa de mortalidade natural μ , assim como é descrito por meio de (1.1.33)-(1.1.35). Portanto, o sistema de equações (1.1.50)-(1.1.53) devem ser modificados:

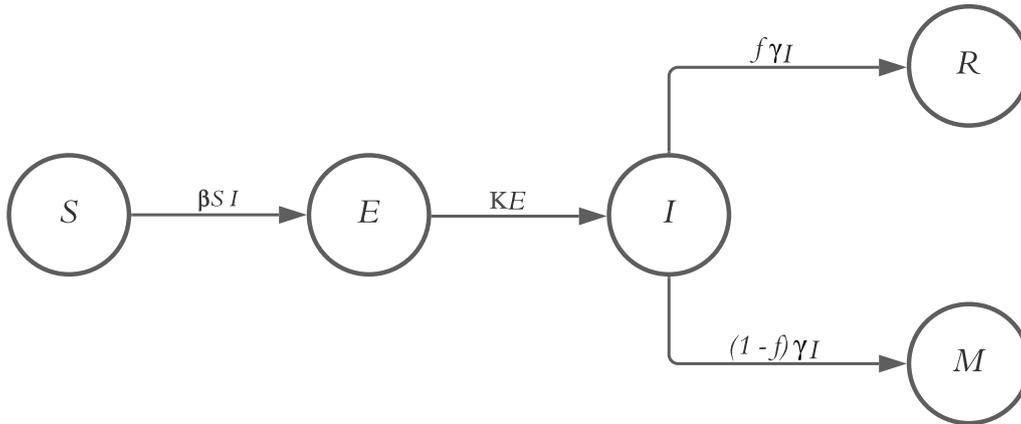


Figura 1.15: Fluxograma do modelo compartimental SEIRM, apresentando o compartimento dos óbitos acometidos pela doença, em que a proporção de infectados que tornam-se vítimas é relacionada por sua taxa de letalidade $(1 - f)$. Fonte: o autor, 2021.

$$S' = BN - \beta SI - \mu S, \quad (1.1.60)$$

$$E' = \beta SI - \kappa E - \mu E, \quad (1.1.61)$$

$$I' = \kappa E - \gamma I - \mu I, \quad (1.1.62)$$

$$R' = \gamma I - \mu R. \quad (1.1.63)$$

A Figura 1.16 expõe a representação da dinâmica desse modelo por meio do fluxo de indivíduos em seus compartimentos.

Diferentemente do modelo SEIR sem dinâmicas vitais, o número de reprodução básico \mathcal{R}_0 sofre certas alterações. Devido à taxa de mortalidade natural, um indivíduo que adquiriu a doença e está presente no compartimento dos expostos pode vir a óbito, desta forma, não contribuindo para o número de infecções secundárias causadas por um infectado. Portanto, seja $\frac{\kappa}{\kappa + \mu}$ a fração dos indivíduos que passam do compartimento E para I , é notável que ocorre uma modificação no número de reprodução básico, dada pela multiplicação desta fração de indivíduos [17], da seguinte forma

$$\mathcal{R}_0 = \frac{\kappa}{\kappa + \mu} \frac{S\beta}{N(\gamma + \mu)}, \quad (1.1.64)$$

de modo que, ao adotar $\mu = 0$, \mathcal{R}_0 regride ao modelo SIR e SEIR usuais.

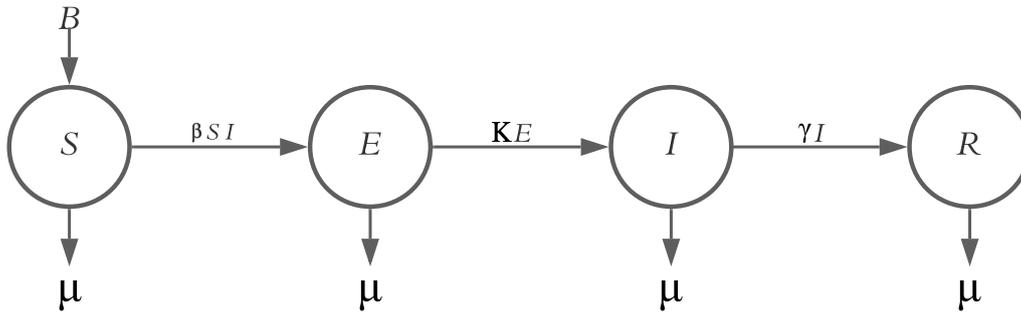


Figura 1.16: Fluxograma do modelo compartimental SEIR. A presença de dinâmicas vitais inclui indivíduos suscetíveis adicionados à população a uma taxa de natalidade B . Nesse cenário, os indivíduos de qualquer compartimento podem falecer, seguindo uma taxa de mortalidade natural μ . Fonte: o autor, 2021.

1.1.6 Modelo SIS

No cenário em que os indivíduos de uma população não adquirem imunidade após o período médio de infecção $1/\gamma$, estes são exonerados do compartimento I com a mesma taxa de recuperação γ encontrada nos modelos anteriores. Tais indivíduos retornam à classe suscetível S , estando aptos a contraírem novamente a doença, alterando sua classe de suscetível para infectado. A dinâmica desse modelo é descrita pelo seguinte sistema de equações diferenciais

$$S' = -\beta SI + \gamma I, \quad (1.1.65)$$

$$I' = \beta SI - \gamma I, \quad (1.1.66)$$

com $N = S + I$ representando a população total do mesmo. A Figura 1.17 exibe o fluxograma do modelo em questão. Esse modelo epidemiológico é chamado de SIS, visto a sua característica de reinfecção.

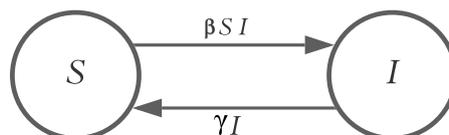


Figura 1.17: Fluxograma do modelo compartimental SIS. Nesse modelo a recuperação da doença não confere imunidade aos indivíduos. Fonte: o autor, 2021.

Como o tamanho total da população é constante, tem-se que o sistema de equações do modelo SIS pode ser reduzido a uma única equação. Ao substituir $S = N - I$ na equação (1.1.66), obtém-se

$$I' = \beta I(N - I) - \gamma I \quad (1.1.67)$$

$$= rI \left(1 - \frac{I}{K}\right), \quad (1.1.68)$$

com $r = \beta N - \gamma$ referindo-se a uma taxa de crescimento e $K = r/\beta$. A equação (1.1.68) representa a equação logística, um modelo clássico em dinâmica de populações. Por meio do método da separação de variáveis, pode-se analisar a equação (1.1.68), de modo a extrair a equação explícita que descreve o compartimento dos infectados em função do tempo. Logo, tem-se

$$\frac{dI}{I(1 - \frac{I}{K})} = r dt \quad (1.1.69)$$

$$\int_{I_0}^I \frac{dI'}{I'(1 - \frac{I'}{K})} = \int_{t=0}^t r dt' \quad (1.1.70)$$

$$\frac{I}{I - K} = \frac{I_0}{I_0 - K} e^{rt} \quad (1.1.71)$$

$$I(t) = \frac{I_0 K e^{rt}}{I_0 - K} \frac{1}{1 - \frac{I_0}{I_0 - K} e^{rt}} \quad (1.1.72)$$

$$I(t) = \frac{K}{1 + (\frac{K}{I_0} - 1) e^{-rt}}, \quad (1.1.73)$$

em que I_0 é o número de infectados inicial. Assim, a forma explícita para I é

$$I(t) = \frac{\beta N - \gamma}{\beta + [(\frac{\beta N - \gamma}{I_0}) - \beta] e^{-(\beta N - \gamma)t}}. \quad (1.1.74)$$

Ao tomar o limite $t \rightarrow \infty$, nota-se que a doença usada no modelo permanece na população indefinidamente, pois $\lim_{t \rightarrow \infty} I(t) = K = N - \gamma/\beta$. Como $I = N - S$, determina-se que $\lim_{t \rightarrow \infty} S(t) = \frac{\gamma}{\beta}$, que é o resultado obtido em (1.1.45). Esses limites são demonstrados de acordo com a Figura 1.18. A condição limite $r > 0$ é reescrita em termos do número \mathcal{R}_0

$$\mathcal{R}_0 = \frac{S\beta}{N\gamma}. \quad (1.1.75)$$

Para $\mathcal{R}_0 > 1$, a doença difunde-se e permanece na população, com o número de

infectados estabilizando-se no valor K . Isso é chamado de equilíbrio endêmico do sistema. Quando $\mathcal{R}_0 < 1$, a doença tende a desaparecer do sistema estudado. Para intervalos de tempo muito longos, leva-se em conta as taxas de natalidade

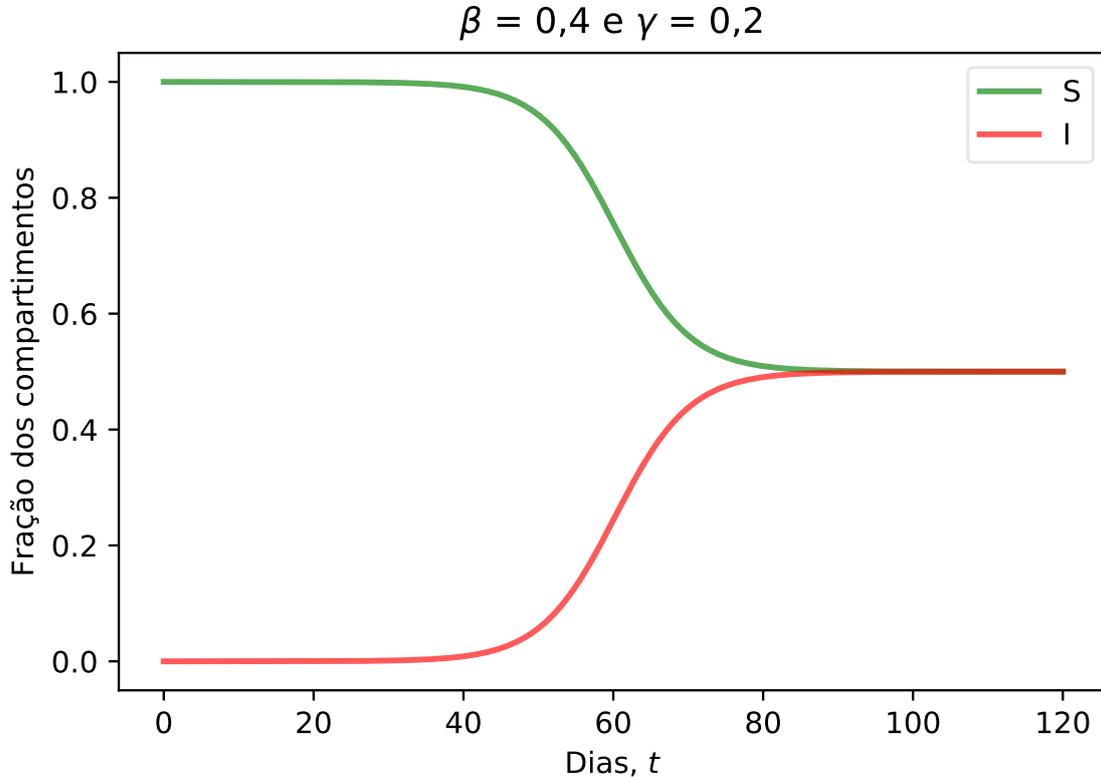


Figura 1.18: Simulação do modelo SIS, em que introduz-se um único indivíduo infectado numa população completamente suscetível como condição inicial, mantendo as taxas de transmissão $\beta = 0,4$ e recuperação $\gamma = 0,2$. Considerando os limites $\lim_{t \rightarrow \infty} S = \gamma/\beta = 1/2$ e $\lim_{t \rightarrow \infty} I = N - \gamma/\beta = 1/2$, faz-se notável o resultado da equação (1.1.74). Fonte: o autor, 2021.

e mortalidade natural da população. Desse modo, o tamanho total da população é compelido a alterar-se, a menos que ambas as taxas assumam o mesmo valor. Portanto, ao introduzir estas dinâmicas vitais no modelo, têm-se

$$S' = BN - \beta SI + \gamma I - \mu S, \quad (1.1.76)$$

$$I' = \beta SI - \gamma I - \mu I, \quad (1.1.77)$$

$$N' = BN - \mu N, \quad (1.1.78)$$

em que B e μ correspondem às taxas de natalidade e mortalidade natural, respectivamente. O fluxo de indivíduos por meio dos compartimentos deste modelo é descrito pela Figura 1.19.

Ao tomar-se o mesmo tratamento denotado pelas equações (1.1.68)-(1.1.74), têm-

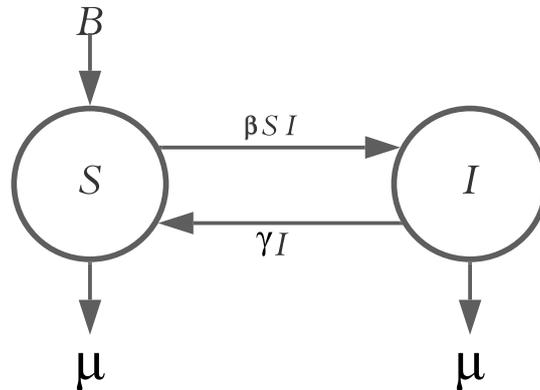


Figura 1.19: Fluxograma do modelo compartimental SIS, ilustrando a entrada e saída de indivíduos da população, descritas pelas taxas de natalidade B e mortalidade natural μ , respectivamente. Fonte: o autor, 2021.

se que $r \rightarrow \beta N - \gamma - \mu$ e $K = r/\beta$. Portanto, a equação explícita para o compartimento dos Infectados com a presença de dinâmicas vitais no modelo é

$$I(t) = \frac{\beta N - \gamma - \mu}{\beta + \left[\frac{(\beta N - \gamma - \mu)}{I_0} - \beta\right] e^{-(\beta N - \gamma - \mu)t}}. \quad (1.1.79)$$

A estabilidade do modelo pode ser analisada tomando o limite da equação (1.1.79) tendendo ao infinito. Portanto, $\lim_{t \rightarrow \infty} I(t) = N - \frac{\gamma + \mu}{\beta}$ e $\lim_{t \rightarrow \infty} S(t) = \frac{\gamma + \mu}{\beta}$. Porém, este limite é válido apenas para $B = \mu$, uma vez que assume-se N constante. Para compreender $B \neq \mu$, é necessário considerar $N = N(t)$ e a equação (1.1.78). Novamente, determina-se o número básico de reprodução para este modelo

$$\mathcal{R}_0 = \frac{S\beta(N)}{N(\gamma + \mu)}. \quad (1.1.80)$$

A Figura 1.20 demonstra graficamente a dinâmica deste modelo e a permanência da doença na população para $t \rightarrow \infty$.

Neste capítulo, foi realizada uma revisão com respeito aos modelos compartimentais encontrados na literatura e foi visto como epidemias podem ser modeladas a partir destes modelos, com cada um deles carregando consigo particularidades específicas de doenças, como a presença de um período latente antes do período infeccioso e a ausência de imunidade duradoura. Em suma, modelos compartimentais são muito úteis para descrever a dinâmica de transmissão de doenças em uma população muito grande, fornecendo um indicador se haverá um surto epidêmico

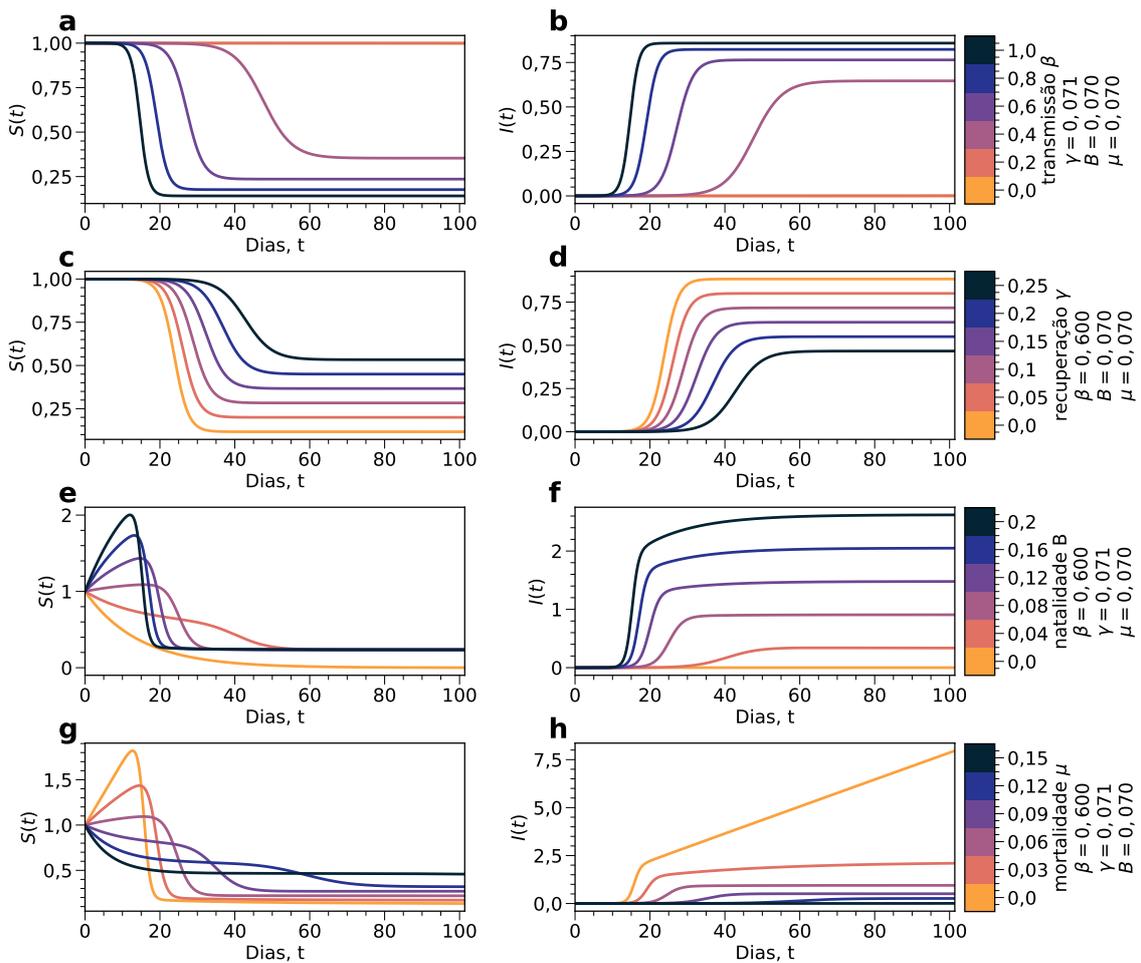


Figura 1.20: Gráficos do modelo SIS com dinâmicas vitais. Para os gráficos **a**, **b**, **c** e **d**, a estabilidade assintótica de $I \neq 0$ descrita pelo resultado da equação (1.1.79) é muito clara, em que $B = \mu$. No gráfico **e**, é notável que o compartimento suscetível tende a $S_m = \frac{\gamma + \mu}{\beta}$ para quaisquer valores de $B \neq 0$, ao passo que valores limites de I no gráfico **f** são descritos por (1.1.79), tomando $N = N(t)$. O mesmo ocorre para os gráficos **g** e **h**. Fonte: o autor, 2021.

ou não, dados os parâmetros necessários. Outro ponto importante descrito é com relação ao equilíbrio dos modelos, em especial a estabilidade assintótica dos compartimentos destes. Contudo, existem outras vertentes para o estudo de tais sistemas. Dentre elas, será estudado modelos descritos por meio de processos estocásticos no capítulo seguinte, e será visto como é possível empregar estes modelos estocásticos para quantificar esta dinâmica epidemiológica.

Modelos Estocásticos em Epidemiologia

Os modelos deterministas vistos anteriormente são excelentes para descrever o comportamento da transmissão de doenças em uma população muito grande. Contudo, para populações menores, as premissas adotadas à priori deixam de fazer sentido. Isso ocorre pois os efeitos contínuos da população perante às equações, como o decréscimo de suscetíveis ou o aumento de recuperados, deixam de se manifestar, revelando seu caráter discreto. Não obstante, os modelos deterministas não lidam com os possíveis comportamentos aleatórios presentes em um surto epidêmico real. Um exemplo disso seria uma variação muito maior do que o esperado no número de novos infectados entre dois intervalos de tempo consecutivos. Dessa forma, convém adotar modelos epidemiológicos estocásticos para a descrição da dinâmica de transmissão de doenças em uma população, que levam em conta a presença de variáveis estocásticas. Essas são funções aleatórias que dependem do tempo, concomitantemente com propriedades markovianas, que surgem de um caso particular de processo estocástico com estados discretos, chamado de cadeias de Markov.

2.1 Cadeias de Markov

Para introduzir o conceito de cadeias de Markov, será apresentado brevemente o processo estocástico denominado caminhada aleatória (*random walk*). Em sua forma mais simples, este processo estocástico é unidimensional, discreto no tempo e só pode assumir determinadas posições específicas descritas por uma variável aleatória x_t . Consiste em um caminho descrito por uma partícula, por meio de uma sucessão de

passos aleatórios, dados por uma certa probabilidade [19]. Os estados desse sistema são representados pela sua posição em determinado tempo $t = N\Delta t$, em que N é o número de passos dados pela partícula e Δt é o intervalo de tempo entre esses passos, vide a Figura 2.1. Para uma caminhada aleatória cujo tamanho do passo é

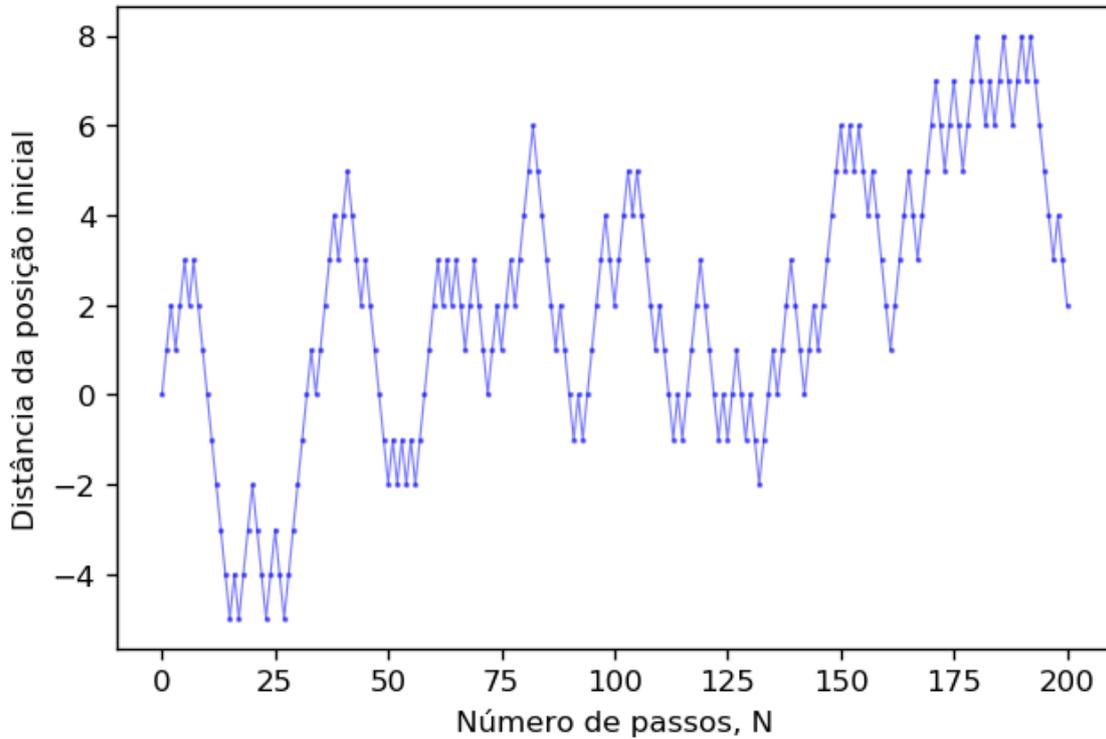


Figura 2.1: Simulação de uma caminhada aleatória com passos uniformes, em que as probabilidades deste caminhante dar um passo em sentido positivo ou negativo são iguais. Fonte: o autor, 2021.

indeterminado, qualquer posição deste sistema pode ser atingida de qualquer outra posição do mesmo, vide a Figura 2.2. Portanto, este processo estocástico fica definido até o instante de tempo $t = l$ pela distribuição de probabilidade conjunta

$$\mathcal{P}_l(n_0, n_1, \dots, n_l), \quad (2.1.1)$$

em que x_t assume os valores n_0 em $t = 0$, n_1 em $t = 1$ e assim sucessivamente, até n_l em $t = l$. Agora, se x_t assume o valor n_{l+1} em um instante $t = l + 1$, tal que em $t = i$ tinha-se n_i ($i = 0, 1, 2, \dots, l$), a probabilidade condicional desse processo é

$$\mathcal{P}_{l+1}(n_{l+1}|n_0, n_1, \dots, n_l). \quad (2.1.2)$$

Se esta probabilidade condicional for igual a

$$\mathcal{P}_{l+1}(n_{l+1}|n_l), \quad (2.1.3)$$

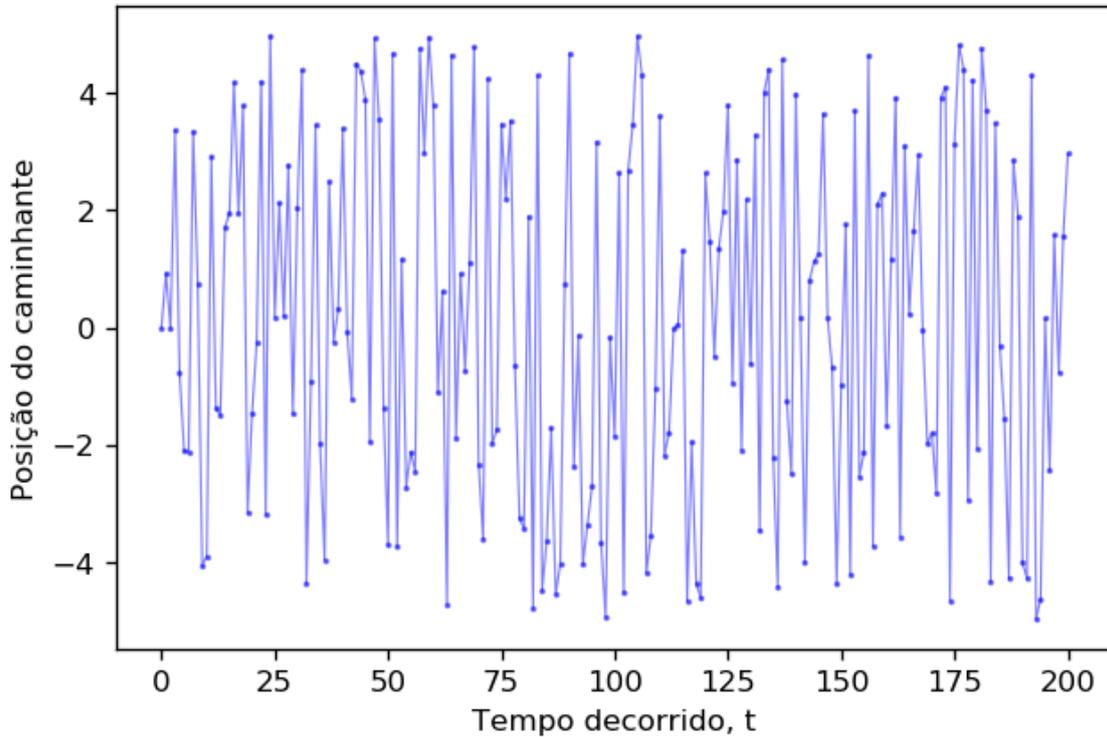


Figura 2.2: Simulação de uma caminhada aleatória com passos não uniformes, em que as probabilidades deste caminhante dar um passo em sentido positivo ou negativo são aleatórias, dependendo da posição atual do caminhante. Fonte: o autor, 2021.

então o processo estudado é chamado de markoviano [20]. Logo, um processo markoviano é aquele cuja probabilidade condicional de x_t tomar um determinado valor em certo instante de tempo que depende somente do valor adotado por ela no instante anterior. Desta forma, a partir da definição de probabilidade condicional, encontra-se

$$\mathcal{P}_l(n_0, n_1, \dots, n_l) = \mathcal{P}_l(n_l|n_{l-1})\mathcal{P}_{l-1}(n_{l-1}|n_{l-2}) \cdots \mathcal{P}_1(n_1|n_0)\mathcal{P}_0(n_0), \quad (2.1.4)$$

de forma que o processo fique definido pelas probabilidades condicionais $\mathcal{P}_l(n_l|n_{l-1})$ e por $\mathcal{P}_0(n_0)$. Definindo agora a probabilidade $P_l(n_l)$ da variável x_t assumir o valor n_l no instante de tempo $t = l$ independente dos valores tomados anteriormente, tem-se

$$P_l(n_l) = \sum \mathcal{P}_l(n_0, n_1, \dots, n_l). \quad (2.1.5)$$

Fazendo uso das equações (2.1.4) e (2.1.5), é obtida uma equação de recorrência para a probabilidade $P_l(n_l)$:

$$P_l(n_l) = \sum_{n_{l-1}} \mathcal{P}_l(n_l|n_{l-1})P_{l-1}(n_{l-1}), \quad (2.1.6)$$

em que a probabilidade condicional $\mathcal{P}_l(n_l|n_{l-1})$ é interpretada como a probabilidade de transição do estado n_{l-1} para o estado n_l . Deste modo, dado $P_0(n_0)$, é possível obter $P_l(n_l)$ para qualquer instante. Nesse sentido, é interessante investigar a probabilidade de encontrar uma partícula em determinado estado $P_{t+\Delta t}(n)$. Portanto, toma-se a soma de todas as possibilidades de passos possíveis com chegada no estado $P_{t+\Delta t}(n)$

$$\begin{aligned} P_{t+\Delta t}(n) &= p_0 P_t(n) \\ &+ p_1 P_t(n+1) + p_{-1} P_t(n-1) \\ &+ p_2 P_t(n+2) + p_{-2} P_t(n-2) + \dots \end{aligned} \quad (2.1.7)$$

$$P_{t+\Delta t}(n) = \sum_{m'} p_{m'} P_t(n+m'), \quad (2.1.8)$$

em que $p_{m'}$ é a respectiva probabilidade de partir do estado $P_t(m')$ e chegar ao estado $P_{t+\Delta t}(n)$. Seja $T(n, m)$ um elemento de uma matriz T e $p_{m'} = T(n, m')$, tem-se o seguinte

$$P_{t+\Delta t}(n) = \sum_m T(n, m) P_t(m). \quad (2.1.9)$$

Essa equação representa o somatório das probabilidades da partícula estar no estado m no tempo t e partir deste, chegando assim no estado n no tempo $t + \Delta t$. A probabilidade total do somatório de todas as possibilidades deve ser igual a 1, de maneira que

$$\sum_m T(n, m) = 1, \quad (2.1.10)$$

de forma que a soma dos elementos de uma coluna qualquer desta matriz T seja igual a 1. Este procedimento satisfaz uma cadeia de Markov, em que o estado seguinte do sistema depende exclusivamente de seu estado atual, que demonstra a falta de memória deste tipo de processo.

Como exemplo, toma-se um sistema genérico de 3 estados, com $n, m \in \{1, 2, 3\}$. Esse conjunto representa os estados acessíveis do sistema, enquanto n representa o estado futuro e m seu estado atual. Portanto, tem-se a seguinte forma matricial do conjunto de probabilidades de transição do sistema, chamada de matriz de transição

ou matriz estocástica. Cada coluna condiz com as possibilidades de transição:

$$T = \begin{pmatrix} T(1,1) & T(1,2) & T(1,3) \\ T(2,1) & T(2,2) & T(3,3) \\ T(3,1) & T(3,2) & T(3,3) \end{pmatrix}. \quad (2.1.11)$$

Devido a condição da equação (2.1.10), é obtido

$$T(1,1) + T(2,1) + T(3,1) = 1, \quad (2.1.12)$$

$$T(1,2) + T(2,2) + T(3,2) = 1, \quad (2.1.13)$$

$$T(1,3) + T(2,3) + T(3,3) = 1. \quad (2.1.14)$$

Chamando $a = T(2,1)$, $b = T(3,1)$, $c = T(1,2)$, $d = T(3,2)$, $e = T(1,3)$ e $f = T(2,3)$, encontra-se o seguinte processo markoviano de 3 estados

$$T = \begin{pmatrix} 1 - a - b & c & e \\ a & 1 - c - d & f \\ b & d & 1 - e - f \end{pmatrix}. \quad (2.1.15)$$

Ao definir a matriz P_t como a matriz coluna em que os elementos são $P_t(n)$, então pode-se reescrever a equação 2.1.9 como um produto de matrizes

$$P_{t+\Delta t} = TP_t. \quad (2.1.16)$$

Então, dada a matriz coluna inicial P_0 , é possível obter $P_{t+\Delta t}$ por meio de

$$P_{t+\Delta t} = T^{(t+\Delta t)}P_0, \quad (2.1.17)$$

cujos problema é reduzido ao cálculo da potência de ordem $(t + \Delta t)$ desta matriz estocástica T .

2.2 Modelo DTMC

O modelo epidemiológico DTMC (*discrete time Markov chains*¹) é caracterizado por adotar que o tempo é discreto, em conjunto com suas variáveis de estado. Traçando um paralelo com os modelos compartimentais descritos anteriormente, tais variáveis representam o tamanho dos compartimentos da população. Portanto, sejam $\mathcal{S}(t)$, $\mathcal{I}(t)$ e $\mathcal{R}(t)$ as variáveis aleatórias discretas representantes do tamanho da

¹Cadeias de Markov discretas no tempo.

população suscetível, infectada e recuperada em um tempo t , respectivamente, de modo que as mesmas satisfaçam

$$\mathcal{S}(t), \mathcal{I}(t), \mathcal{R}(t) \in \{0, 1, 2, \dots, N\}, \quad (2.2.1)$$

para os tempos $t \in \{0, \Delta t, 2\Delta t, \dots\}$. Tomando o modelo mais simples para este tipo de descrição matemática, o modelo SIS, nota-se a existência de uma única variável aleatória independente, $\mathcal{I}(t)$, uma vez que a população suscetível é descrita por $\mathcal{S}(t) = N - \mathcal{I}(t)$, com população total constante. Deste modo, tem-se que o processo estocástico $\{\mathcal{I}(t)\}_{t=0}^{\infty}$ possui $p_i(t)$ como distribuição de probabilidade, expressa por

$$p_i(t) = P\{\mathcal{I}(t) = i\}, \quad (2.2.2)$$

para $i = 0, 1, 2, \dots, N$, com $\sum_{i=0}^N p_i(t) = 1$ [21, 22]. Sabendo que o processo estocástico envolve cadeias de Markov, o mesmo possui a propriedade que o processo em um tempo $t + \Delta t$ depende somente do processo no intervalo de tempo anterior, i.e.,

$$\mathcal{P}\{\mathcal{I}(t + \Delta t) | \mathcal{I}(0), \mathcal{I}(\Delta t), \dots, \mathcal{I}(t)\} = \mathcal{P}\{\mathcal{I}(t + \Delta t) | \mathcal{I}(t)\}. \quad (2.2.3)$$

Para a completa formulação do modelo epidemiológico SIS DTMC, é necessário descrever a relação entre os estados $\mathcal{I}(t)$ e $\mathcal{I}(t + \Delta t)$, baseada nas premissas de sua contraparte compartimental e por meio de um objeto chamado matriz de transição, que agrega as possíveis mudanças de estado dada pelas probabilidades de transição. Considerando a transição do estado $\mathcal{I}(t) = i$ ao estado $\mathcal{I}(t + \Delta t) = j$, $i \rightarrow j$, tem-se a probabilidade de transição associada a este estado

$$p_{ji}(t + \Delta t, t) = \mathcal{P}\{\mathcal{I}(t + \Delta t) = j | \mathcal{I}(t) = i\}. \quad (2.2.4)$$

Como os modelos apresentados nesta seção são homogêneos com relação ao tempo, suas probabilidades de transição $p_{ji}(t + \Delta t, t)$ não dependem de t , sendo denotados por $p_{ji}(\Delta t)$. Em seguida, escolhe-se o valor do intervalo de tempo Δt suficientemente pequeno, para que as possíveis transições de estado descrevam uma mudança unitária de indivíduos por intervalo Δt ou que descrevam mudança alguma. Portanto, têm-se três casos possíveis para essa mudança: i) a infecção de um novo indivíduo $i \rightarrow i + 1$; ii) a recuperação de um indivíduo infectado $i \rightarrow i - 1$ ou iii) $i \rightarrow i$, em que o estado permanece inalterado em relação à infecções ou recuperações. As probabilidades de transição serão definidas a partir dos parâmetros usados nos outros modelos compartimentais em conjunto com $\mathcal{S}(t) = N - \mathcal{I}(t)$, multiplicados pelo intervalo de

tempo:

$$p_{ji}(\Delta t) = \begin{cases} \frac{\beta i(N-i)}{N} \Delta t, & j = i + 1 \\ \gamma i \Delta t, & j = i - 1 \\ 1 - \left[\frac{\beta i(N-i)}{N} + \gamma i \right] \Delta t, & j = i \\ 0 & j \neq i + 1, i, i - 1. \end{cases} \quad (2.2.5)$$

Portanto, a probabilidade de ocorrer uma nova infecção ($i \rightarrow i + 1$) é $\beta i(N - i)\Delta t/N$, a probabilidade de um indivíduo ser curado ($i \rightarrow i - 1$) é $\gamma i \Delta t$, e a probabilidade de não ocorrer nenhuma mudança no estado ($i \rightarrow i$) é $1 - [\beta i(N - i)/N + \gamma i] \Delta t$. Desse modo, o somatório de todas as possíveis alterações de estado é igual a 1. Por meio da propriedade de Markov, pode-se expressar $p_i(t + \Delta t)$ em termos do tempo t . Logo, ao tomar $\frac{\beta i(N-i)}{N} = a(i)$, $\gamma i = b(i)$ e $i = 1, 2, \dots, N$ por simplicidade, obtém-se

$$p_i(t + \Delta t) = p_{i-1}(t)a(i-1)\Delta t + p_{i+1}(t)b(i+1)\Delta t + p_i(t)(1 - [a(i) + b(i)]\Delta t). \quad (2.2.6)$$

A matriz de transição $P(\Delta t)$ é gerada de acordo com a ordenação dos estados de 0 a N , sendo esta uma matriz quadrada tridiagonal de ordem $(N + 1)$, cuja forma é:

$$P(\Delta t) = \begin{pmatrix} 1 & b(1)\Delta t & 0 & \dots & 0 & 0 \\ 0 & 1 - (a+b)(1)\Delta t & b(2)\Delta t & \dots & 0 & 0 \\ 0 & a(1)\Delta t & 1 - (a+b)(2)\Delta t & \dots & 0 & 0 \\ 0 & 0 & a(2)\Delta t & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & b(N-1)\Delta t & 0 \\ 0 & 0 & 0 & \dots & 1 - (a+b)(N-1)\Delta t & b(N)\Delta t \\ 0 & 0 & 0 & \dots & a(N-1)\Delta t & 1 - b(N)\Delta t \end{pmatrix}. \quad (2.2.7)$$

A matriz $P(\Delta t)$ é a matriz estocástica. O somatório dos elementos de cada coluna é igual a 1, representando a probabilidade total da soma das possíveis mudanças de estado. Sendo $i = 0, 1, \dots, N$, tem-se que os estados $\{0, 1, \dots, N\}$ podem ser divididos em dois conjuntos, o de estado recorrente, $\{0\}$, cuja probabilidade de voltar ao mesmo estado é 1, e o de estado transiente, $\{1, \dots, N\}$, cuja probabilidade de voltar ao mesmo estado é menor que 1, permitindo a mudança para outros estados além do estado atual. O primeiro é chamado de estado absorvente, pois, uma vez nele, não é possível atingir quaisquer outros estados do sistema. Portanto, dada a matriz estocástica (2.2.7), realizou-se uma simulação deste modelo SIS DTMC, levando em conta as diferentes probabilidades de transição de estado do sistema, mediante o número de infectados em seu estado atual, isto é, estas probabilidades de transição de estado se adaptam ao número de infectados na variável aleatória

$\mathcal{I}(t)$ para todos os instantes da simulação. A Figura 2.3 compreende esta simulação.

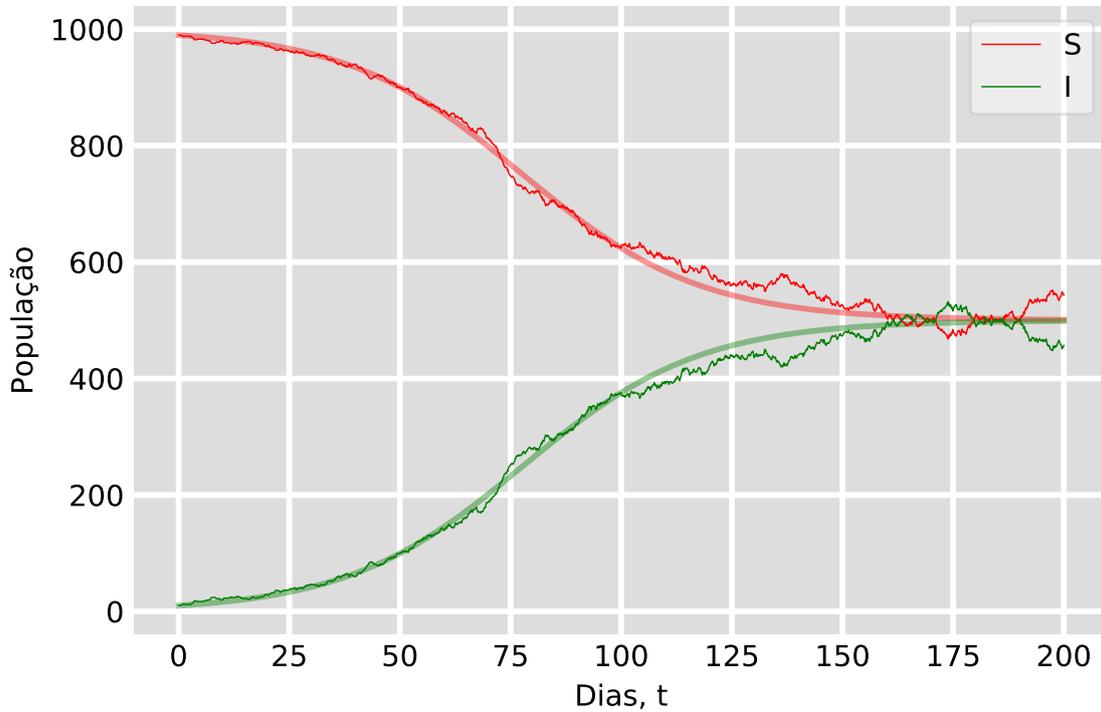


Figura 2.3: Gráfico do modelo SIS DTMC, em que são introduzidos 10 indivíduos infectados, com uma doença com $\beta = 0,1$ e $\gamma = 0,05$ em uma população com $N = 1.000$ indivíduos. Inicialmente, são introduzidos 10 infectados na população para garantir que o sistema apresente um comportamento parecido com sua contraparte determinista, uma vez que há a possibilidade destes indivíduos se recuperarem sem que estes possam difundir a doença, devido a possibilidade de cair no estado recorrente. Fonte: o autor, 2021.

Isto já estabeleceu uma base para trabalhar com os modelos estocásticos DTMC. Portanto, pode-se estender esta ideia ao modelo SIR. Sendo um processo com a presença de duas variáveis aleatórias independentes, $\mathcal{S}(t)$ e $\mathcal{I}(t)$, uma vez que $\mathcal{R}(t) = N - \mathcal{S}(t) - \mathcal{I}(t)$, tem-se a probabilidade conjunta associada ao processo $\{(\mathcal{S}(t), \mathcal{I}(t))\}_{t=0}^{\infty}$, descrita por

$$p_{(s,i)}(t) = P\{\mathcal{S}(t) = s, \mathcal{I}(t) = i\}. \quad (2.2.8)$$

Este processo bivariado é homogêneo no tempo e apresenta a propriedade markoviana, sendo as probabilidades de transição baseadas nas premissas do modelo SIR determinista. Análogo ao modelo SIS DTMC, assume-se Δt suficientemente pequeno de modo que haja, no máximo, a variação de um único indivíduo nos compartimentos

entre cada intervalo de tempo. Estas probabilidades são denotadas por

$$p_{(s+k,i+j),(s,i)}(\Delta t) = \mathcal{P}\{(\Delta \mathcal{S}, \Delta \mathcal{I}) = (k, j) | (\mathcal{S}(t), \mathcal{I}(t)) = (s, i)\}, \quad (2.2.9)$$

em que $\Delta \mathcal{S} = \mathcal{S}(t + \Delta t) - \mathcal{S}(t)$, e podem ser escritas como

$$p_{(s+k,i+j),(s,i)}(\Delta t) = \begin{cases} \frac{\beta si}{N} \Delta t, & (k, j) = (-1, 1) \\ \gamma i \Delta t, & (k, j) = (0, -1) \\ 1 - \left[\frac{\beta si}{N} + \gamma i \right] \Delta t, & (k, j) = (0, 0) \\ 0, & (k, j) \neq \begin{cases} (-1, 1) \\ (0, -1) \\ (0, 0) \end{cases} \end{cases} \quad (2.2.10)$$

Devido ao processo ser bivariado, os estados são pares ordenados, e a forma da matriz de transição depende de como os estados (s, i) são ordenados. Para este caso, é simples notar que os estados recorrentes do sistema são descritos por (s, i) com $i = \{0\} \forall s \in \{0, 1, 2, \dots, N-1, N\}$ e os transientes por todos os outros pares permitidos pelo mesmo. A Figura 2.4 descreve uma simulação do modelo SIR DTMC aplicado à uma população de 1000 indivíduos, em que novamente as probabilidades de transição de estado para todos os instantes de tempo t se adaptam às variáveis aleatórias $\mathcal{S}(t)$ e $\mathcal{I}(t)$.

Ao longo deste capítulo, foi introduzida uma nova abordagem no estudo da disseminação e evolução dos surtos epidêmicos em uma população com relação ao tempo, cujo comportamento das variáveis aleatórias $\mathcal{S}(t)$, $\mathcal{I}(t)$ e $\mathcal{R}(t)$ é muito similar ao comportamento da solução numérica obtida por meio do conjunto de equações da abordagem compartimental. Porém, apesar de similares, há pontos de divergência entre ambos os modelos, devido ao caráter probabilístico da dinâmica estocástica. Dentre os pontos apresentados, o de maior destaque é com respeito ao número de reprodução básico, uma vez que há a possibilidade de uma doença com $\mathcal{R}_0 \geq 1$ invadir uma população e rapidamente desaparecer da mesma, por meio da probabilidade de transição de um estado levar a outro que seja recorrente.

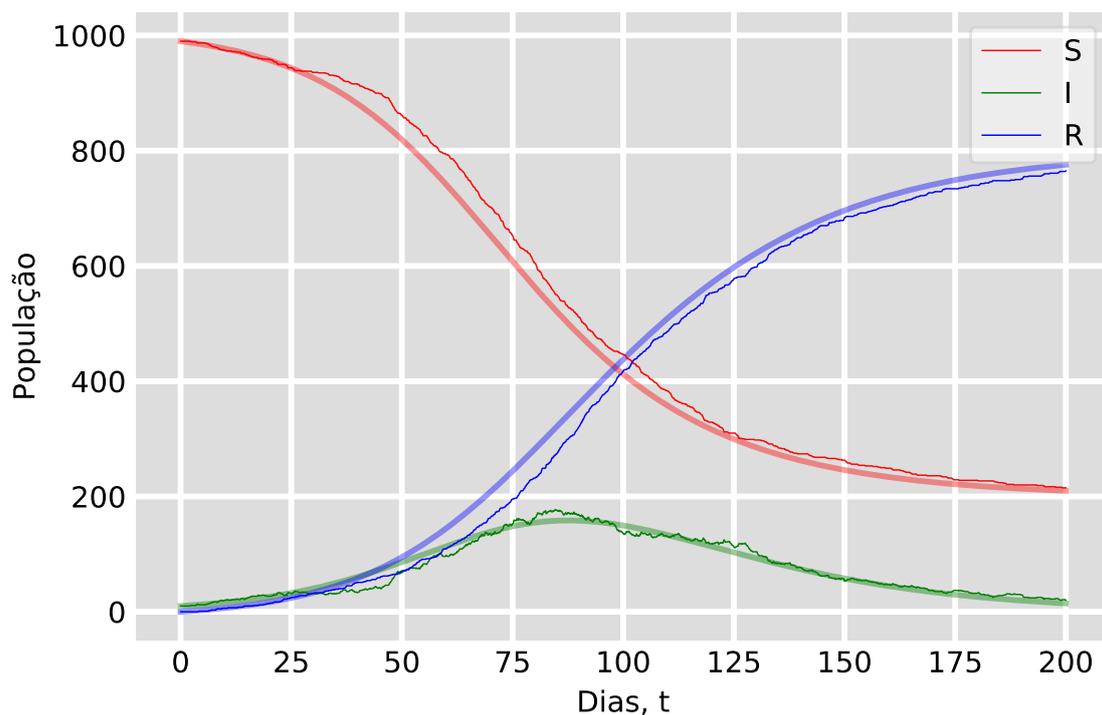


Figura 2.4: Gráfico do modelo SIR DTMC em comparação com o modelo SIR determinista, em que uma doença com $\beta = 0,1$ e $\gamma = 0,05$ é introduzida em uma população de $N = 1.000$ indivíduos cujas condições iniciais compreendem $I(0) = 10$, $R(0) = 0$ e $S(0) = N - I(0) - R(0) = 990$. São introduzidos 10 infectados inicialmente para diminuir a probabilidade do sistema adentrar o estado recorrente nos primeiros instantes da simulação. Fonte: o autor, 2021.

Os dados utilizados neste estudo, referentes à escala mundial¹, foram extraídos do repositório *Our World in Data* (OWID) [23–25]. Estes reportam o número de casos positivos e óbitos, fornecidos diariamente para cada país, assim como seus valores acumulados ao longo do tempo. Os dados referentes à escala nacional foram extraídos do repositório de dados públicos disponibilizados em formato acessível [26]. Para cada estado e cidade do Brasil, tem-se o número de novas infecções, casos recuperados e óbitos registrados por dia, concomitantemente com seus números acumulados ao longo do tempo.

Esses dados são apresentados para posteriormente serem estudados por meio do modelo epidemiológico SIR usual e com dinâmicas vitais.

3.1 Dados mundiais

A Figura 3.1 apresenta um conjunto de 4 gráficos com o total de ocorrências globais registradas, a partir do dia 1º de janeiro de 2020, com respeito à COVID-19: a) novos casos, b) casos acumulados, c) novas mortes e d) mortes acumuladas. Observa-se que o número de casos acumulados ultrapassa 140 milhões e o de óbitos 3 milhões. A Figura 3.2 mostra esses dados para os três países com o maior número de casos registrados. Os Estados Unidos acumulam mais de 30 milhões de casos e mais de 500 mil óbitos. A Índia e o Brasil possuem mais de 10 milhões de casos, com número de mortes passando dos dos 150 mil e 350 mil, respectivamente. Para

¹Inclui o Brasil.

se ter uma ideia do cenário atual dos países, mapas da distribuição de casos e mortes reportadas são demonstrados nas Figuras 3.3 e 3.4, em que o número total de casos é representado pela cor respectiva a cada país.

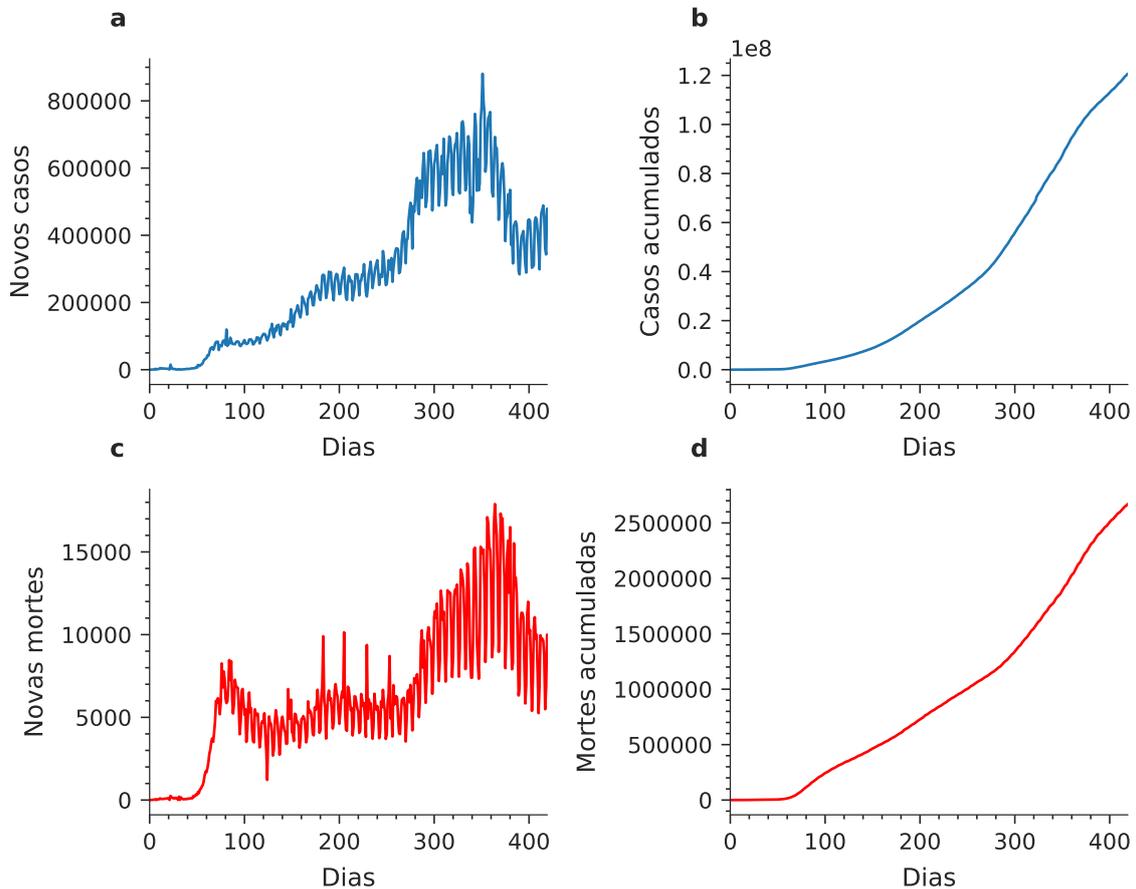


Figura 3.1: Dados de ocorrências globais da COVID-19 (OWID). A primeira coluna (painéis **a** e **c**) de gráficos apresenta as atualizações diárias no número de casos e mortes desde seu registro inicial (01-01-2020), respectivamente. Por outro lado, a segunda coluna (painéis **b** e **d**) de gráficos apresenta o número de casos e mortes acumuladas, compreendendo o mesmo intervalo de tempo da anterior. É notável a evolução no número de casos com o passar do tempo, refletindo na acentuação da curva presente no primeiro gráfico. Fonte: o autor, 2021.

Esses dados também fornecem o número de reprodução \mathcal{R} dependente do tempo para cada país, que será usado no Capítulo 4 para determinar a qualidade do ajuste para os países analisados, considerando o modelo SIR usual e com dinâmicas vitais.

3.2 Brasil

Para os dados referentes à escala nacional, são expostos o número de casos e mortes acumuladas para todos os 26 estados brasileiros e para o Distrito Federal.

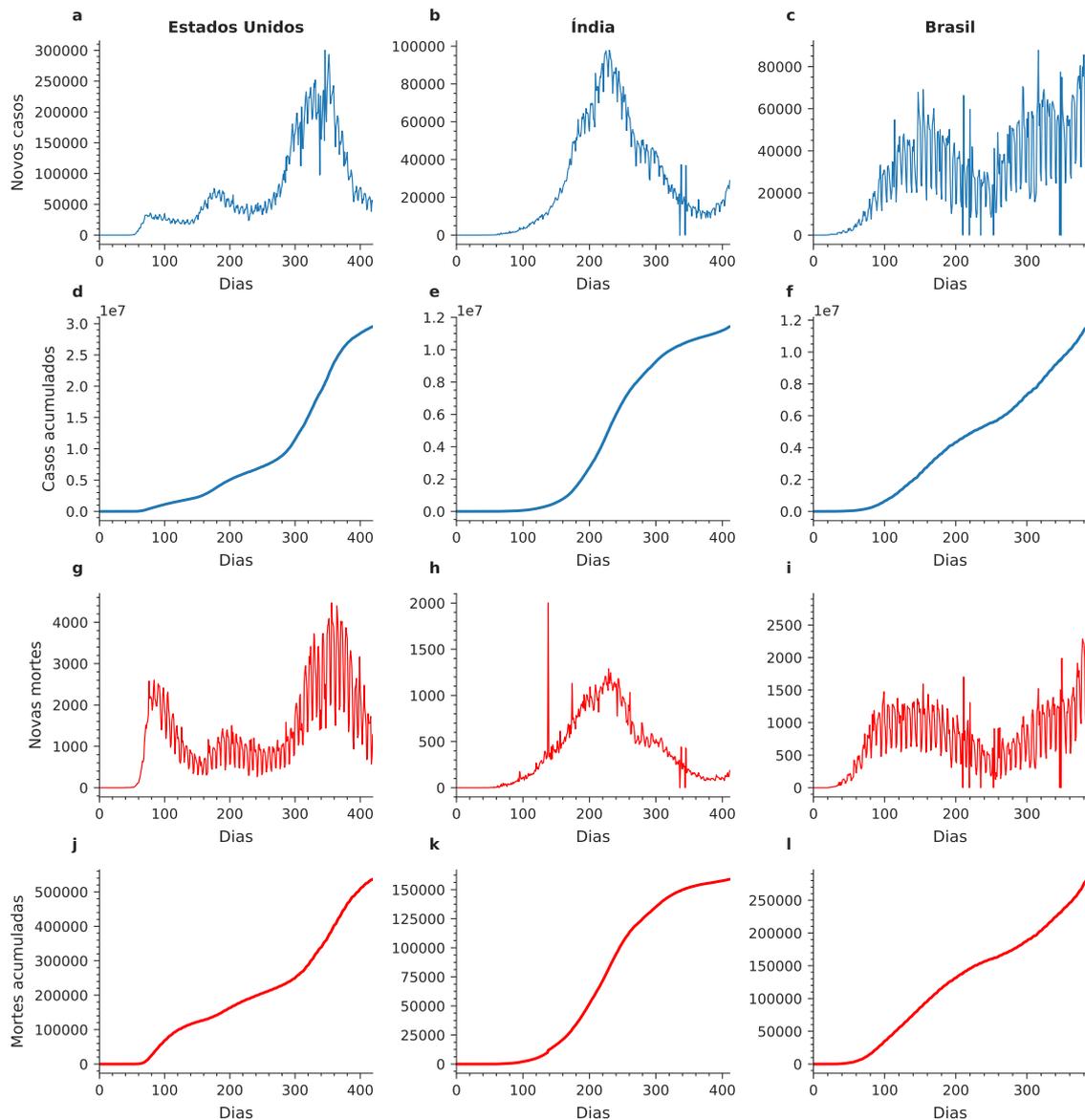


Figura 3.2: Top três países com maior número de casos (OWID). Fez-se um apanhado dos três países com maior número de casos acumulados. As colunas descrevem as ocorrências presentes na Figura 3.1 para os Estados Unidos, Índia e Brasil, nesta ordem. A primeira linha representa a atualização diária no número de casos registrados, a segunda descreve a evolução do número de casos cumulativos, a terceira caracteriza a atualização diária no número de mortes registradas e, por fim, a evolução temporal do número de mortes cumulativas dos países. As datas iniciais variam de acordo com o primeiro registro feito em cada país, todas no mês de março de 2020. Fonte: o autor, 2021.

A Figura 3.5 descreve a evolução temporal no número de casos acumulados nos estados brasileiros, desde seu registro inicial, em meados de fevereiro, até a data de coleta dos dados. Concomitantemente, a Figura 3.6 descreve a evolução temporal no número de mortes acumuladas nos estados brasileiros em detrimento da doença,

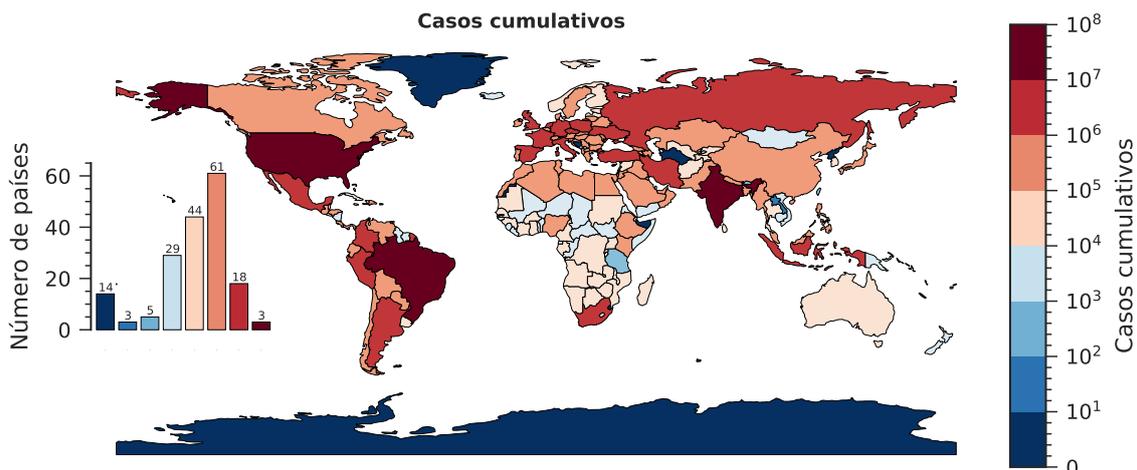


Figura 3.3: Mapa referente à distribuição de casos acumulados da COVID-19 (OWID). A barra de cores mostra o número de casos acumulados em cada um dos países e o inset mostra o histograma para essa quantidade.

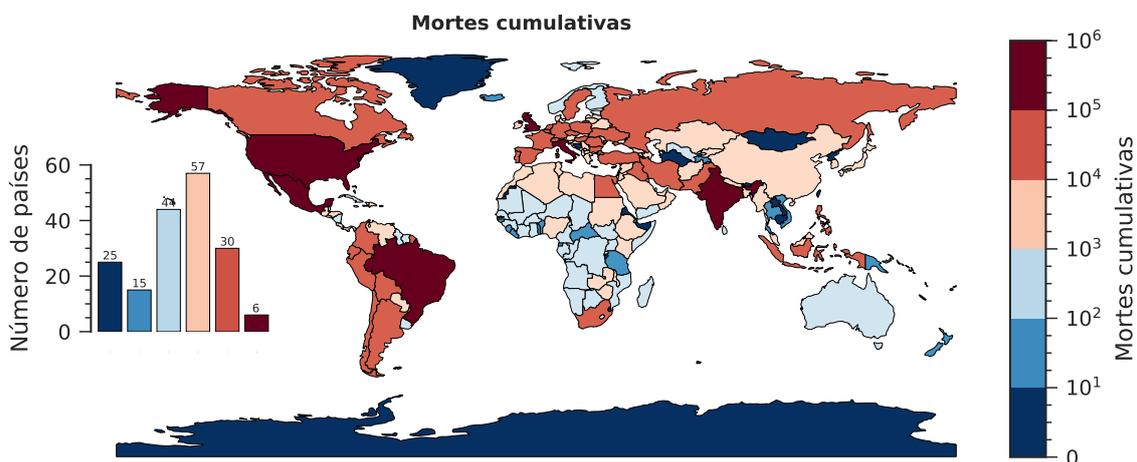


Figura 3.4: Mapa referente à distribuição de mortes acumuladas pela COVID-19 (OWID). A barra de cores mostra o número de mortes acumuladas em cada um dos países e o inset mostra o histograma para essa quantidade. Fonte: o autor, 2021.

com seu registro sendo realizado no mesmo período.

O número de reprodução \mathcal{R} não encontra-se disponível para essa base de dados. Desta forma, será considerado somente o modelo SIR usual para o ajuste do número de casos acumulados no Capítulo 4.

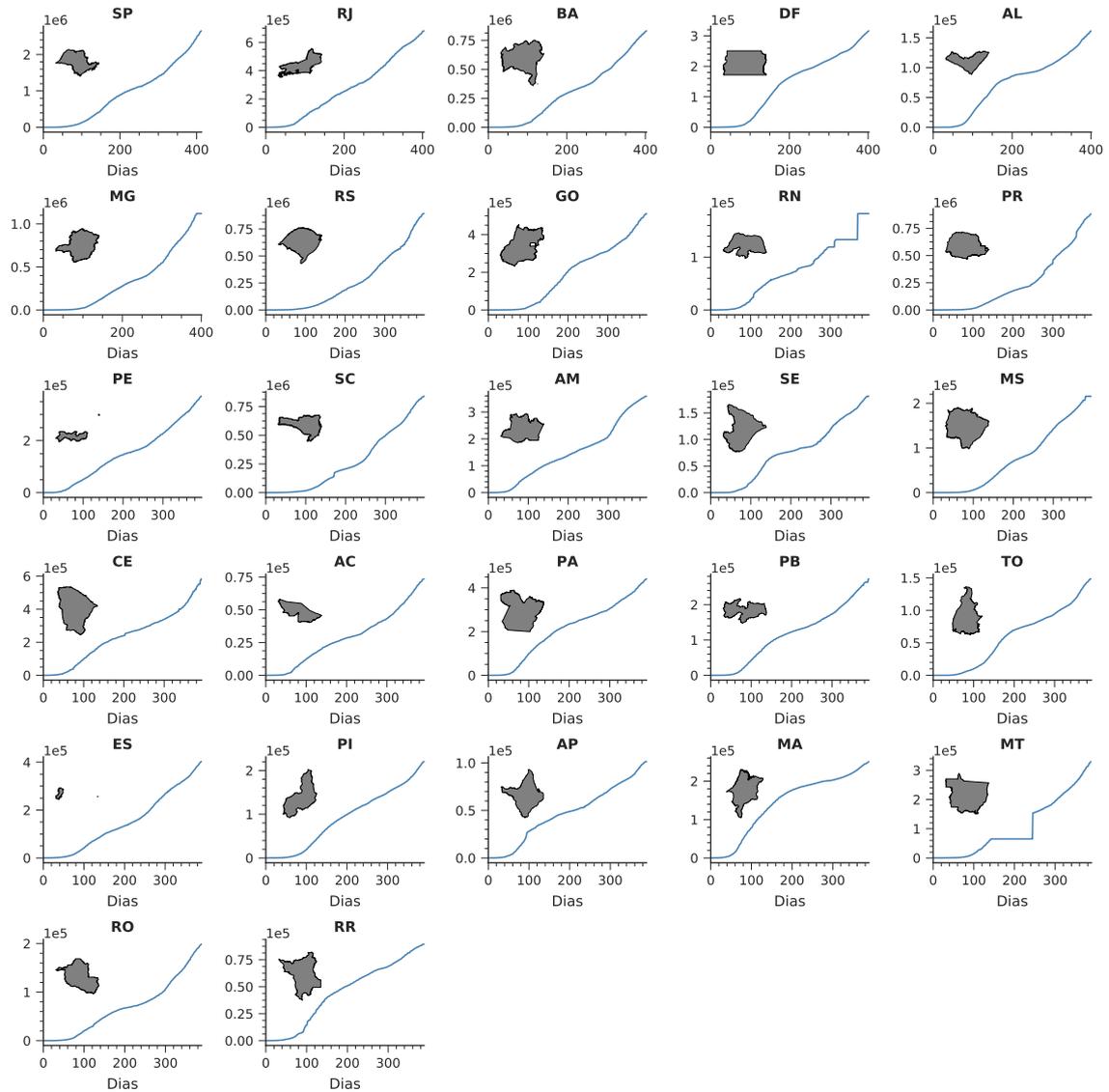


Figura 3.5: Distribuições de casos acumulados da COVID-19 nos estados brasileiros, demonstrando o crescimento dos mesmos ao longo do tempo. A data inicial depende de quando foi registrado o primeiro caso no estado. Fonte: o autor, 2021.

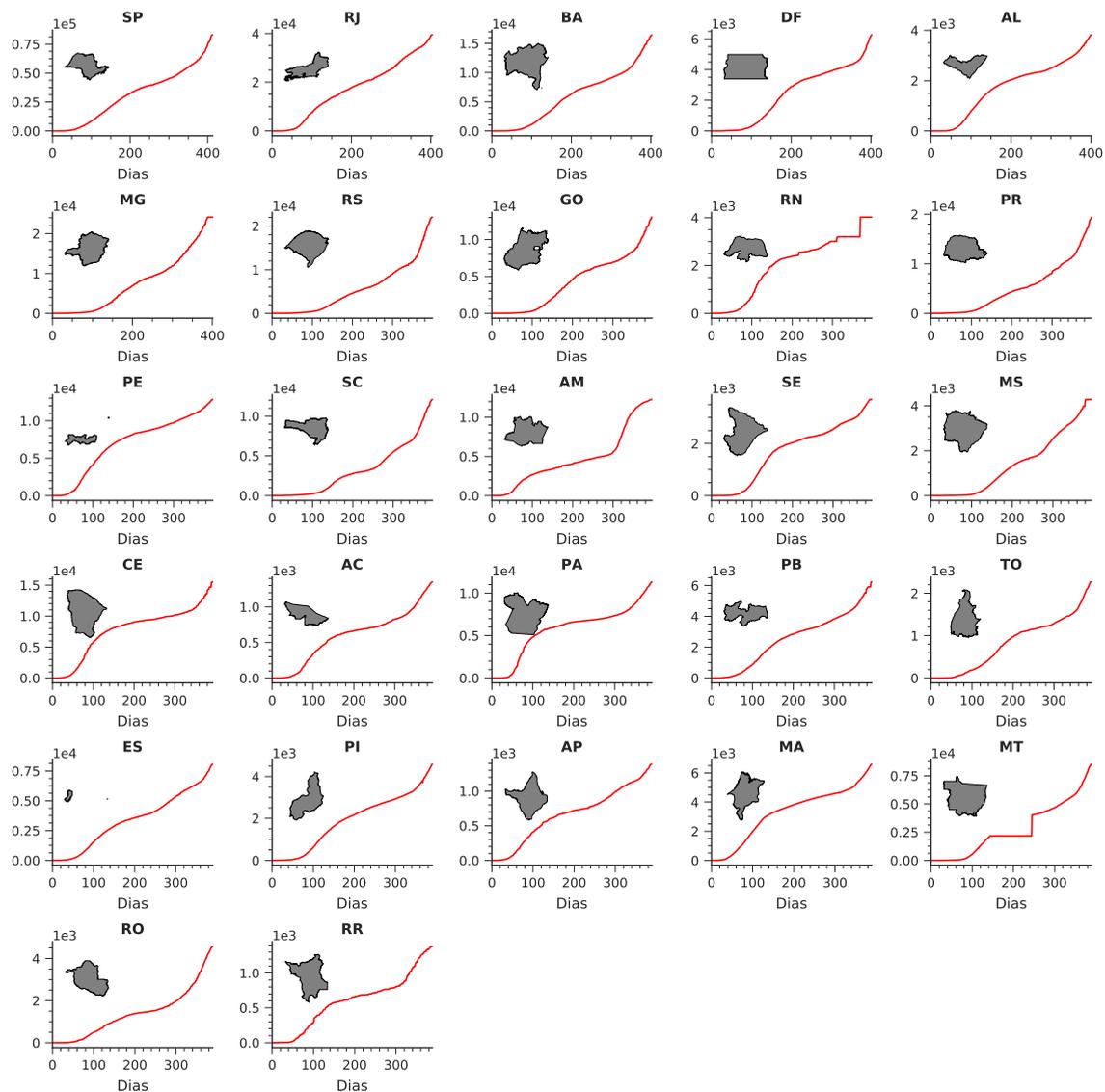


Figura 3.6: Distribuições de mortes da COVID-19 nos estados brasileiros para o mesmo intervalo de tempo da Figura 3.5. Fonte: o autor, 2021.

Modelagem dos dados

Nos capítulos 1 e 2, foram apresentados uma série de modelos que explicam o mecanismo de difusão de uma doença em uma população, por meio de um surto epidêmico. Foi apresentado também, no capítulo 3, o conjunto de dados relativo à atual pandemia da COVID-19, desde a escala nacional à mundial. Uma vez em posse das equações que descrevem a dinâmica do sistema concomitantemente com os dados, recorre-se ao trabalho de modelagem matemática para unir estas duas informações. Por conseguinte, têm-se como objetivo a determinação dos respectivos parâmetros referentes aos modelos epidemiológicos descritos anteriormente, de modo a ajustar as equações para estas condizerem de fato com o comportamento observado em nossa sociedade. Para esta tarefa, será utilizada a linguagem de programação *Python*. O método consiste em resolver numericamente um conjunto de equações diferenciais representadas pelo modelo escolhido, com tempo t aproximadamente igual ao dos dados.

4.1 Ajustando o modelo SIR

Devido as informações contidas no conjunto de dados, será utilizado o modelo SIR, descrito pelo compartimento dos suscetíveis, infectados e removidos. Com respeito ao modelo SIR, a tarefa de ajuste recai na determinação dos parâmetros presentes no sistema de equações do mesmo, sendo estes a taxa de transmissão β e a taxa de recuperação γ . Inicialmente, gera-se listas de valores para os parâmetros β e γ , isto é, $(\beta_1, \beta_2, \dots, \beta_n)$ e $(\gamma_1, \gamma_2, \dots, \gamma_n)$. Em seguida, escolhe-se um país ou

estado a ser estudado, de maneira que o número acumulado de ocorrências (casos confirmados, casos recuperados e óbitos) seja conhecido para o tempo t , em dias. Tal número será designado por $D(t)$. A partir dos dados, é possível determinar o valor de cada compartimento para o tempo $t = 0$ e, conseqüentemente, o tamanho da população N . Em posse desses, usa-se os valores de β e γ para criar diversas soluções que fornecem o número acumulado de casos teórico até o tempo t . Esse consiste no número de casos ativos (I) e de removidos (R), que incluem os óbitos, e será designado por $C(t) \equiv I(t) + R(t)$. Esse modelo simplificado não leva em consideração o número de mortes, portanto esses entram também no compartimento dos removidos.

Para a descrição da solução aproximada destas equações, é utilizado um método computacional chamado ODEINT [16], presente na biblioteca de análise de dados *SciPy* [27], em que resolve-se numericamente o conjunto de equações diferenciais para uma gama de valores dos parâmetros β e γ , pertencentes as respectivas listas. Posteriormente, são escolhidas somente as soluções cujo desvio percentual em relação aos dados seja menor ou igual a 10%:

$$\left| \frac{D(t) - C(t)}{D(t)} \right| 100\% \leq 10\%. \quad (4.1.1)$$

Feito isso, calcula-se a média e o desvio padrão de β e γ para as soluções adotadas, com seu respectivo número de reprodução para cada instante de tempo.

Com relação à escala global, realizou-se esse procedimento para os três países com maior número de casos registrados da COVID-19, tomando como base a fonte de dados apresentada no Capítulo 3. Como mencionado, esses dados fornecem o valor da taxa de reprodução \mathcal{R} . Dessa forma, nossa comparação ainda possibilita o cálculo dessa quantidade para uma comparação direta.

A lista de valores do parâmetro β usada no ajuste dos dados compreende o intervalo $[0,050, 0,600]$ [28], em que a diferença entre um valor β_n e seu consecutivo β_{n+1} é 0,001.

Com relação à lista de valores do parâmetro $\gamma = 1/T$, em que T é o período infeccioso médio da doença, será usada a lista de valores do período infeccioso, em dias, proposta pela OMS [29] [8, 18], cuja variação entre T_n e T_{n+1} é unitária. Tomando a inversa dos valores desta lista e aproximando-os para um intervalo com maior abrangência, obtém-se a lista de valores do parâmetro γ , compreendida pelo intervalo $[0,020, 0,200]$, em que a diferença entre um valor γ_n e γ_{n+1} é 0,001.

Devido às soluções que atendem aos critérios descritos previamente, originam-se os parâmetros generalizados $\beta(t)$ e $\gamma(t)$, que correspondem ao ajuste do modelo. Por

meio destes, é possível determinar o número de reprodução generalizado [28, 30] da doença para a série temporal em questão:

$$\mathcal{R}(t) = \frac{S(t) \beta(t)}{N(t) \gamma(t)}. \quad (4.1.2)$$

Os resultados deste método são mostrados, para os três países com maior número de casos registrados, na Figura 4.1. Por este procedimento, nota-se a proximidade dos valores do número de ocorrências acumuladas do modelo para com os dados, assim como é descrito visualmente pela primeira linha de gráficos da Figura 4.1. Apesar dessa concordância, observou-se diferenças em $\mathcal{R}(t)$ apresentadas na linha inferior, mas é possível observar que a tendência seguida é a mesma. Na Tabela 4.1, apresenta-se uma comparação numérica entre o número de reprodução dependente do tempo do modelo e dos dados.

	\mathcal{R} do modelo	\mathcal{R} do dado
Estados Unidos	(1,174, 6,132)	(0,730, 3,610)
Índia	(1,298, 3,488)	(0,860, 2,110)
Brasil	(1,188, 4,384)	(0,930, 2,380)

Tabela 4.1: Resultados do modelo SIR à nível mundial, apresentando os valores mínimos e máximos atingidos pelo número de reprodução para cada país.

Reduzindo a escala em nível nacional, emprega-se o método descrito aos cinco estados brasileiros, cujos resultados são apresentados na Figura 4.2. Não foi possível realizar tal análise mais aprofundada em nível nacional, uma vez que não encontram-se disponíveis os valores de \mathcal{R} para os estados em questão. Porém, levando em consideração os resultados referentes à escala mundial, é crível que a curva obtida represente um limite superior para os valores reais.

Quando uma doença invade uma população majoritariamente suscetível, seu número de reprodução tende a ser expressivo nos instantes iniciais devido a diversos fatores, principalmente para uma doença até então desconhecida. Conforme o avançar do tempo, nota-se uma redução monotônica em \mathcal{R} , oriunda da diminuição em $\beta(t)$ e no número de indivíduos suscetíveis na população. Essas são as explicações comportamentais do número de reprodução fornecidas pelo modelo SIR.

Todavia, este modelo é uma simplificação dos processos dinâmicos presentes em uma população, que leva em conta apenas três compartimentos e dois parâmetros para a descrição de sua dinâmica. Tais simplificações geram certos erros nos resultados com relação aos dados, como pode ser observado ao comparar o número \mathcal{R} dos dados e do modelo na segunda linha da Figura 4.1.

Esta diferença pode ser explicada em partes por fatores sociais impostos à popu-

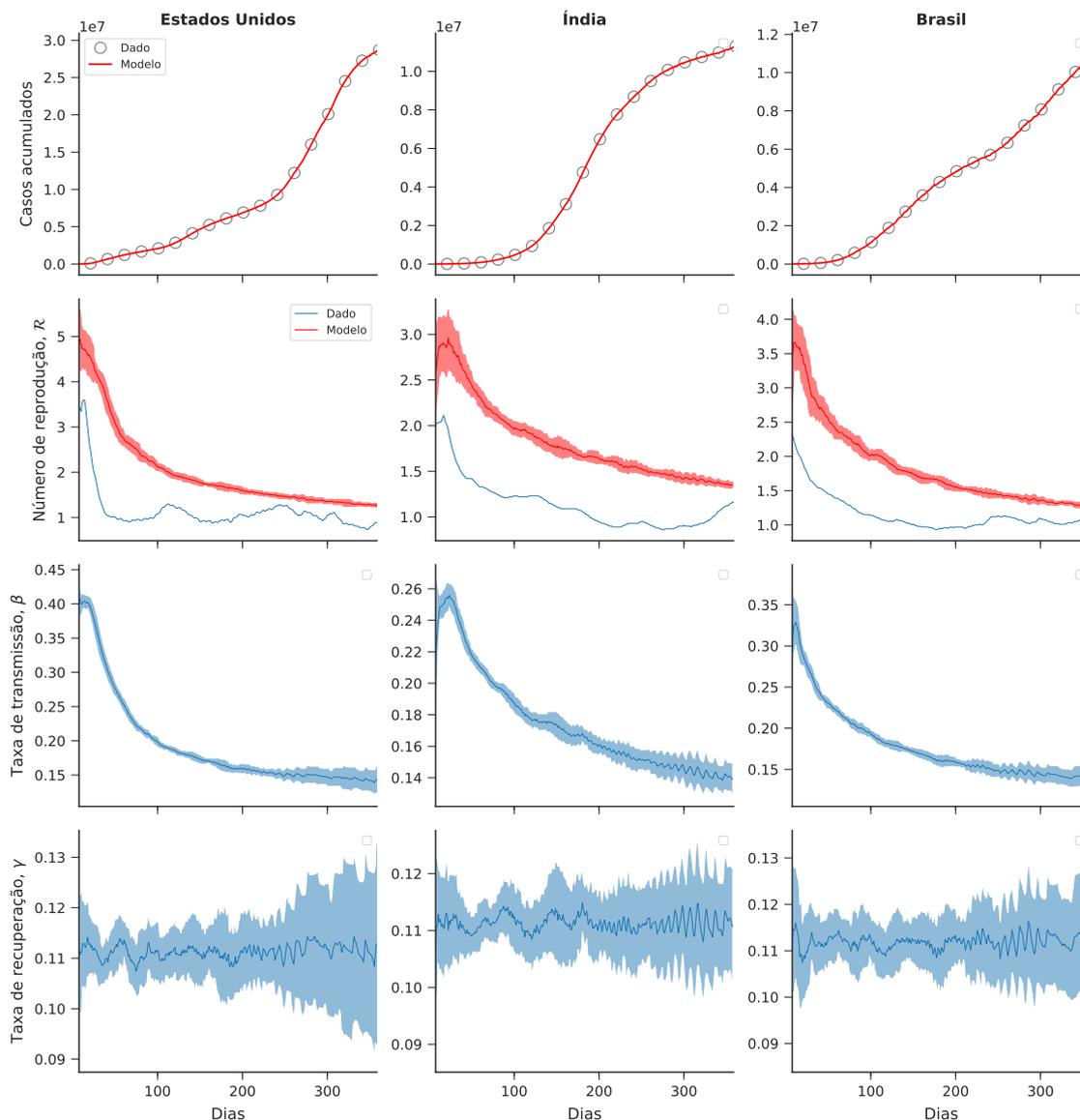


Figura 4.1: Ajuste do modelo SIR aos dados globais. Resolveu-se o conjunto de equações diferenciais do modelo SIR para um conjunto de valores de β e γ e foram escolhidas somente as soluções com desvio percentual relativo aos dados menor ou igual a 10%. Os painéis mostram os resultados para os Estados Unidos, Índia, e Brasil que são denotados, respectivamente, pela ordem em que os gráficos aparecem na figura. Na primeira e segunda linha são exibidos os ajustes do modelo SIR em comparação com o número de casos acumulados e a taxa de reprodução, respectivamente. As duas últimas linhas condizem com os valores de β e γ respectivamente, determinados pelo ajuste ao longo do tempo. As curvas destacadas em relação à área sombreada que as cerca representam a média móvel dos valores num intervalo de tempo semanal, enquanto esta área descreve o desvio padrão desta média em relação aos valores supracitados. Fonte: o autor, 2021.

lação, tal qual a adoção de medidas sanitárias e de isolamento social, que inevitavelmente ocasionam uma redução considerável em \mathcal{R} , até mesmo nos instantes iniciais

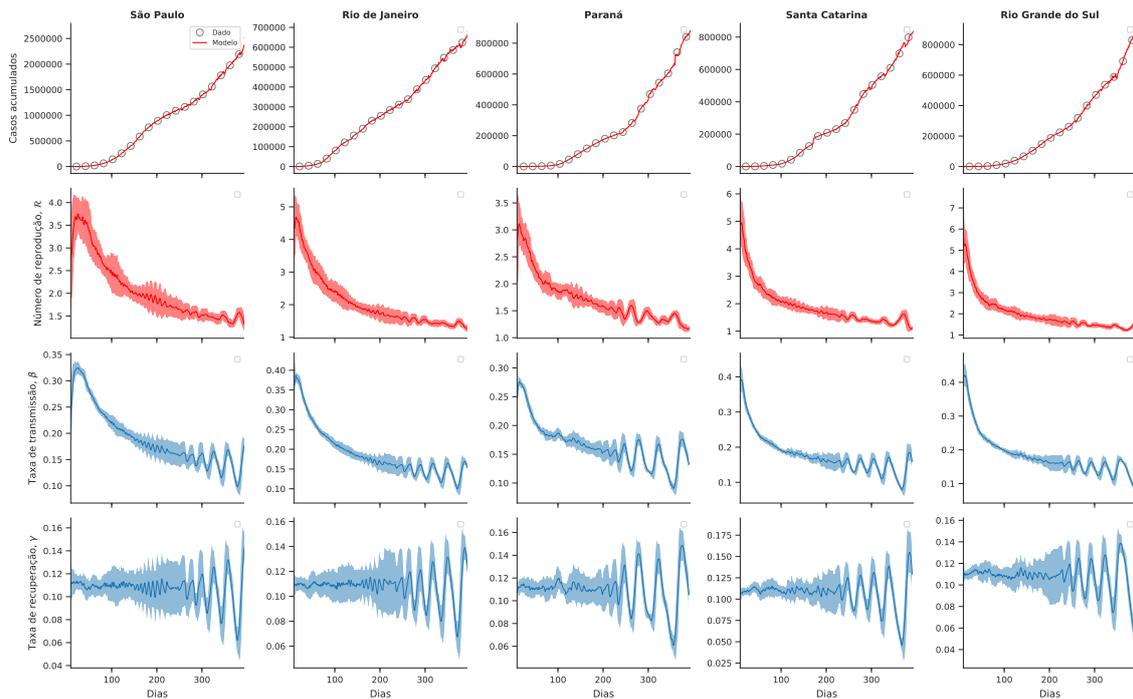


Figura 4.2: Ajuste do modelo SIR aos estados brasileiros. Resolveu-se o conjunto de equações diferenciais do modelo SIR para um conjunto de valores de β e γ e foram escolhidas somente as soluções com desvio percentual relativo aos dados menor ou igual a 10%. Os painéis mostram os resultados para cinco estados brasileiros: São Paulo, Rio de Janeiro, Paraná, Santa Catarina e Rio Grande do Sul. Na primeira e segunda linha são exibidos os ajustes do modelo SIR em comparação com o número de casos acumulados e a taxa de reprodução. As duas últimas linhas condizem com os valores de β e γ , respectivamente, determinados pelo ajuste ao longo do tempo. As curvas destacadas em relação à área sombreada que as cerca representam a média móvel dos valores num intervalo de tempo semanal, enquanto esta área descreve o desvio padrão desta média em relação aos valores supracitados. É notável a diferença entre o comportamento de β e γ ao comparar esta figura com a Figura 4.1. As oscilações desses valores tornam-se muito mais perceptíveis ao final do intervalo de tempo fornecido. Fonte: o autor, 2021.

do surto epidêmico.

Logo, para uma melhor descrição do número de reprodução dependente do tempo, será necessário adicionar fenômenos ligados à doença ou à população, tal qual a presença de efeitos demográficos (natalidade e mortalidade natural) como exemplo deste último, uma vez que a acreção ou redução de indivíduos no sistema produz um impacto direto no número de reprodução.

4.2 Ajustando o modelo SIR com dinâmicas vitais

Visando reduzir a discrepância entre \mathcal{R} do modelo SIR em comparação com os dados, foi realizado o ajuste deste para o modelo SIR com a adição de dinâmicas vitais, isto é, com a presença de uma taxa de natalidade B e uma taxa de mortalidade μ . Seguiu-se os mesmos procedimentos para o ajuste do modelo SIR, com a inclusão de uma lista de valores para estes parâmetros, isto é, (B_1, B_2, \dots, B_n) e $(\mu_1, \mu_2, \dots, \mu_n)$. Sob este novo ponto de vista, foram analisados novamente os três países com maior número de ocorrências da COVID-19. A lista de valores de B e μ usada no ajuste desses dados depende das taxas de natalidade e mortalidade de cada país. Usando as taxas referentes ao ano de 2020 [31], foi considerado intervalos de $\pm 0,00300$ para a taxa de natalidade e de $\pm 0,00200$ para a taxa de mortalidade natural, de modo que:

$$\begin{aligned} \text{Estados Unidos: } B &= [0,01100, 0,01400], & \mu &= [0,00500, 0,01200] \\ \text{Índia: } B &= [0,01600, 0,02000], & \mu &= [0,00400, 0,01000] \\ \text{Brasil: } B &= [0,01100, 0,01500], & \mu &= [0,00300, 0,00900]. \end{aligned} \quad (4.2.1)$$

Para determinar o número de reprodução $\mathcal{R}(t)$, é válido considerar a equação 1.1.36 concomitantemente com os parâmetros dependentes do tempo, como $\beta(t)$, $\gamma(t)$ e $\mu(t)$, sendo este valor determinado por

$$\mathcal{R}(t) = \frac{S(t)\beta(t)}{N(t) [\gamma(t) + \mu(t)]}. \quad (4.2.2)$$

A Figura 4.3 mostra os resultados obtidos por meio desse método. Observa-se que os valores do número de reprodução ajustados pelo modelo SIR com dinâmicas vitais assemelham-se mais com os dados, em comparação com o modelo SIR básico. Quanto aos parâmetros B e μ , mesmo considerando um intervalo de valores para estes, o ajuste satisfaz as condições para os valores mínimos considerados ou valores muito próximos deles, motivo de não estarem presentes na figura.

Referente ao número de reprodução obtido pelo ajuste, os resultados dos três países para o modelo SIR com dinâmicas vitais estão presentes na Tabela 4.2.

Destaca-se na Tabela 4.2 o oposto do presenciado na Tabela 4.1, isto é, a proximidade nos valores mínimos e máximos do modelo com relação aos dados, valores estes esperados dadas as curvas da Figura 4.3.

Em suma, observou-se que o modelo SIR com dinâmicas vitais se ajustou melhor aos dados, visto que este representa uma versão mais fidedigna do cenário real de nossa sociedade que o modelo SIR usual, apesar de não ser descartada a hipótese

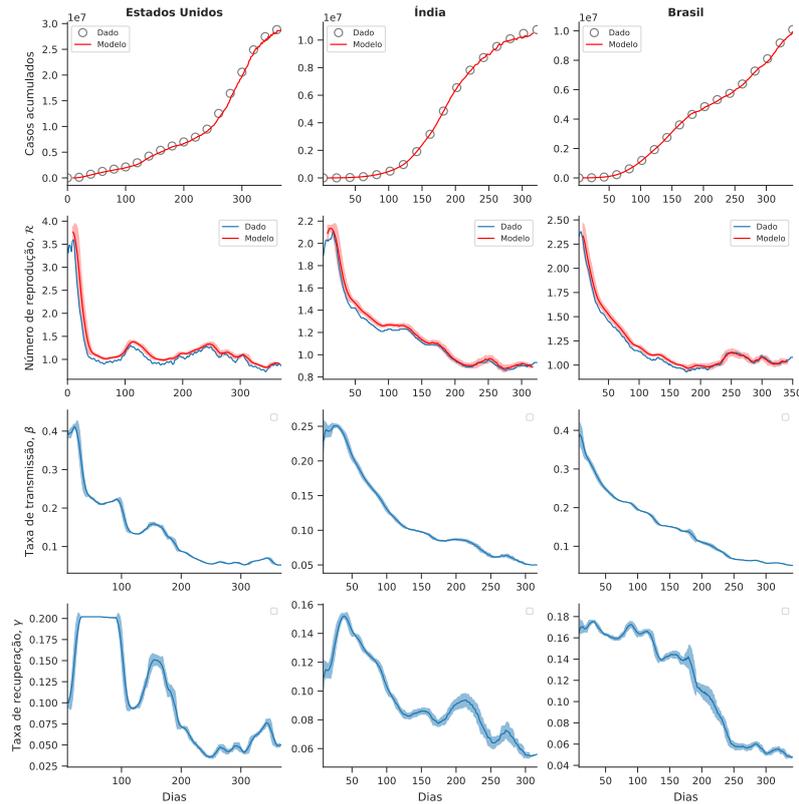


Figura 4.3: Ajuste do modelo SIR com dinâmicas vitais aos dados globais.

Resolveu-se o conjunto de equações diferenciais do modelo SIR com dinâmicas vitais e foram escolhidas somente as soluções com desvio percentual relativo aos dados menor ou igual a 10% para os três países com maior incidência de casos: Estados Unidos, Índia e Brasil, apresentados, respectivamente, pela ordem em que os gráficos aparecem na figura. O primeiro elemento de cada coluna representa o ajuste do modelo SIR com dinâmicas vitais em comparação com os dados do país em questão. O segundo elemento de cada coluna condiz com o número de reprodução \mathcal{R} dado pela equação 4.2.2. Uma comparação imediatamente emergente é a diferença acerca do número de reprodução entre o modelo SIR com dinâmicas vitais e os dados, contrastando com os resultados obtidos na Figura 4.1, devido à proximidade do valor teórico e do empírico. Os dois últimos elementos de cada coluna condizem com os valores de β e γ determinados pelo ajuste ao longo do tempo, representados pela taxa de transmissão generalizada $\beta(t)$ e pela taxa de recuperação generalizada $\gamma(t)$. As curvas destacadas em relação à área sombreada que as cerca representam a média móvel dos valores num intervalo de tempo semanal, enquanto esta área descreve o desvio padrão desta média em relação aos valores supracitados. Fonte: o autor, 2021.

de que este resultado mais acurado possa ser proveniente de um *overfitting*¹.

¹Refere-se a uma análise que corresponda muito próxima ou exatamente a um determinado conjunto de dados e pode, portanto, não conseguir ajustar dados adicionais ou prever observações futuras de forma confiável

	\mathcal{R} do modelo	\mathcal{R} do dado
Estados Unidos	(0,904, 4,123)	(0,730, 3,610)
Índia	(0,876, 2,135)	(0,860, 2,110)
Brasil	(0,967, 2,335)	(0,930, 2,380)

Tabela 4.2: Resultados do modelo SIR com dinâmicas vitais à nível mundial, apresentando os valores mínimos e máximos atingidos pelo número de reprodução para cada país.

Conclusões

Neste trabalho, foi realizado o ajuste do modelo epidemiológico SIR aos dados da atual pandemia da COVID-19, para cinco estados brasileiros (São Paulo, Rio de Janeiro, Paraná, Santa Catarina e Rio Grande do Sul) e para os três países com maior número de casos reportados (Estados Unidos da América, Índia e Brasil).

Foi observado que, para a escala mundial, o ajuste do modelo foi preciso com relação ao número de casos acumulados. Entretanto, tomando como medida de qualidade do ajuste o número de reprodução \mathcal{R} , percebeu-se a disparidade entre o modelo e os dados, apesar do comportamento entre ambos ser similar.

Para a escala nacional, observou-se que o ajuste empregado para os cinco estados brasileiros analisados foi satisfatório. Embora não houvesse \mathcal{R} nos dados para comparar com o modelo analisado, é razoável supor, considerando os resultados anteriores, que a curva do número de reprodução seja considerada um limite superior para seus valores reais. Nessa mesma análise, foram percebidas oscilações cada vez mais aparentes para β e γ conforme a evolução do sistema no tempo. Apesar de não haver indícios concretos para tal, é possível que estas oscilações derivem das medidas sanitárias adotadas pelos governos locais, como a adoção de distanciamento social e, em último caso, o *lockdown*², amplamente empregadas de modo recorrente durante a pandemia.

Devido à presença de um número de reprodução empírico nos dados globais, submeteu-se os mesmos ao ajuste do modelo SIR com a presença de dinâmicas vitais, como a natalidade B e a mortalidade natural μ destas populações. Nesse cenário, houve um incremento na precisão do ajuste do número \mathcal{R} , com valores

²Bloqueio que, imposto pelo Estado ou por uma ação judicial, restringe a circulação de pessoas em áreas e vias públicas, incluindo fechamento de fronteiras. Ocorre em situações de pandemia com o intuito de evitar a disseminação do vírus.

mínimos e máximos mais próximos aos valores reais reportados.

Em suma, verificou-se que a teoria proposta por Kermack e McKendrick em 1927 consegue modelar surtos epidêmicos reais a partir de um conjunto de equações diferenciais aliado a uma série de premissas iniciais. Mais ainda, quase um século depois de sua concepção, essa teoria se mostra capaz de extraír informações imprescindíveis sobre a maneira que um surto epidêmico age em nossa sociedade. Isso facilita a compreensão quanto às medidas cabíveis para o controle de um surto, uma vez que permite estimar a extensão de tais eventos e planejar os próximos passos rumo ao combate de epidemias.

Referências

- [1] Statement on the second meeting of the international health regulations (2005) emergency committee regarding the outbreak of novel coronavirus (2019-ncov). URL <https://bit.ly/3kZslWx>. Acessado em 23 de fevereiro de 2021.
- [2] COVID-19 coronavirus pandemic. URL <https://www.worldometers.info/coronavirus/>. Acessado em 30 de outubro de 2020.
- [3] Coronavirus: Overview, prevention and symptoms. URL <https://www.who.int/health-topics/coronavirus>. Acessado em 21 de abril de 2021.
- [4] Symptoms of covid-19. URL <https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/symptoms.html>. Acessado em 21 de abril de 2021.
- [5] Ministério da saúde: Sintomas. URL <https://www.gov.br/saude/pt-br/coronavirus/sintomas>. Acessado em 21 de abril de 2021.
- [6] Kermack, W. O. & McKendrick, A. 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* **115**, 700–721 (1927).
- [7] Kermack, W. O. & McKendrick, A. G. Contributions to the mathematical theory of epidemics. ii.—the problem of endemicity. *Proceedings of the Royal Society of London. Series A, containing papers of a mathematical and physical character* **138**, 55–83 (1932).
- [8] Kermack, W. O. & McKendrick, A. G. Contributions to the mathematical theory of epidemics. iii.—further studies of the problem of endemicity. *Proceedings*

- of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **141**, 94–122 (1933).
- [9] Sigler, L. *Fibonacci's Liber Abaci: a translation into modern English of Leonardo Pisano's book of calculation* (Springer Science & Business Media, 2003).
- [10] Graunt, J. *Natural and political observations made upon the bills of mortality*. 2 (Johns Hopkins Press, 1939).
- [11] Daley, D. J. & Gani, J. *Epidemic modelling: an introduction*. 15 (Cambridge University Press, 2001).
- [12] Smallpox. URL <https://web.archive.org/web/20070921235036/http://www.who.int/mediacentre/factsheets/smallpox/en/>. Acessado em 23 de janeiro de 2021.
- [13] Hamer, W. H. *The milroy lectures on epidemic diseases in england: The evidence of variability and of persistency of type; delivered before the royal college of physicians of london, march 1st, 6th, and 8th, 1906* (Bedford Press, 1906).
- [14] Bohner, M., Streipert, S. & Torres, D. F. Exact solution to a dynamic sir model. *Nonlinear Analysis: Hybrid Systems* **32**, 228–238 (2019).
- [15] Shabbir, G., Khan, H. & Sadiq, M. A note on exact solution of sir and sis epidemic models. *arXiv preprint arXiv:1012.5035* (2010).
- [16] Ahnert, K. & Mulansky, M. Odeint–solving ordinary differential equations in c++. In *AIP Conference Proceedings*, vol. 1389, 1586–1589 (American Institute of Physics, 2011).
- [17] Van den Driessche, P. Reproduction numbers of infectious disease models. *Infectious Disease Modelling* **2**, 288–303 (2017).
- [18] Luiz, M. H. R. Modelos matemáticos em epidemiologia (2012).
- [19] Salinas, S. R. A. *Introdução à Física Estatística* (Editora da Universidade de São Paulo, 2018).
- [20] Tomé, T. & De Oliveira, M. J. *Dinâmica Estocástica e Irreversibilidade* (Editora da Universidade de São Paulo, 2014).
- [21] Brauer, F., Driessche, P. d. & Wu, J. *Lecture notes in mathematical epidemiology* (Berlin, Germany: Springer, 2008).

-
- [22] Daley, D. J. & Gani, J. *Epidemic modelling: an introduction*. 15 (Cambridge University Press, 2001).
- [23] Coronavirus pandemic (covid-19) – the data. URL <https://ourworldindata.org/coronavirus-data>. Acessado em 11 de novembro de 2020.
- [24] Coronavirus (covid-19). URL <https://www.globaldata.com/covid-19/>. Acessado em 11 de novembro de 2020.
- [25] Who coronavirus disease (covid-19) dashboard. URL <https://www.who.int/#>. Acessado em 23 de fevereiro de 2021.
- [26] O Brasil em dados libertos. URL <https://brasil.io/home/>. Acessado em 11 de novembro de 2020.
- [27] Jones, E., Oliphant, T., Peterson, P. *et al.* SciPy: Open source scientific tools for Python (2001–). URL <http://www.scipy.org/>.
- [28] Dos Santos, I., Almeida, G. & de Moura, F. Adaptive sir model for propagation of sars-cov-2 in brazil. *Physica A: Statistical Mechanics and its Applications* **569**, 125773 (2021).
- [29] Report of the who-china joint mission on coronavirus disease 2019 (covid-19). URL <https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf>. Acessado em 09 de abril de 2021.
- [30] Jo, H., Son, H., Jung, S. Y. & Hwang, H. J. Analysis of covid-19 spread in south korea using the sir model with time-dependent parameters and deep learning. *medRxiv* (2020).
- [31] Index mundi country comparison - birth rate. URL <https://www.indexmundi.com/g/r.aspx?t=0&v=25&l=en>. Acessado em 19 de abril de 2021.

APÊNDICE A

Tabelas de dados

Devido ao grande volume de informações que os dados utilizados neste estudo carregam para estarem no Capítulo 3, estas foram separadas neste capítulo à parte. Portanto, o conjunto de tabelas A.1 a A.8 condizem com os dados globais, reportados individualmente pelo ministério da saúde dos países e territórios reconhecidos pela organização das nações unidas (ONU), agrupados e disponibilizados em formato online pela organização mundial da saúde (OMS) [25]. Essas tabelas contêm informações sobre os 237 países e territórios informados, tal o número de casos e óbitos acumulados ao longo da pandemia, estruturadas em ordem decrescente de casos cumulativos.

Com relação às ocorrências nacionais [26], a tabela A.9 dispõe os dados reportados pelos mais de 5000 municípios existentes no Brasil, agregados por estado e pelo Distrito Federal, apresentando a população estimada do ano 2019 e o número de casos e óbitos acumulados ao longo da pandemia até a data de extração destes dados.

Tabela A.1: Dados das ocorrências globais disponibilizadas pela OMS, cujos países são ordenados pelo número de casos cumulativos apresentados. São listados os 30 primeiros países neste quesito.

	País	Casos cumulativos	Mortes cumulativas
1	United States of America	27560643	486970
2	India	10963394	156111
3	Brazil	9978747	242090
4	Russian Federation	4139031	82396
5	The United Kingdom	4083246	119387
6	France	3476135	82959
7	Spain	3121687	66704
8	Italy	2765412	94887
9	Turkey	2616600	27821
10	Germany	2369719	67206
11	Colombia	2207701	58134
12	Argentina	2039124	50616
13	Mexico	2013563	177061
14	Poland	1623218	41823
15	Iran (Islamic Republic of)	1550142	59264
16	South Africa	1498766	48708
17	Ukraine	1293672	24972
18	Indonesia	1252685	33969
19	Peru	1252137	44308
20	Czechia	1134957	18913
21	Netherlands	1042674	15089
22	Canada	834182	21435
23	Portugal	792829	15754
24	Chile	788142	19798
25	Romania	771843	19659
26	Belgium	746302	21821
27	Israel	731037	5417
28	Iraq	657453	13220
29	Sweden	627022	12598
30	Pakistan	567261	12488

Tabela A.2: Dados das ocorrências globais disponibilizadas pela OMS, cujos países são ordenados pelo número de casos cumulativos apresentados. São listados os países que ocupam a colocação de 31° a 60°.

País	Casos cumulativos	Mortes cumulativas
31 Philippines	555163	11673
32 Switzerland	544137	9139
33 Bangladesh	542268	8329
34 Morocco	480056	8524
35 Austria	436100	8210
36 Serbia	429048	4292
37 Japan	421967	7274
38 Hungary	397116	14145
39 Saudi Arabia	374029	6450
40 United Arab Emirates	361877	1073
41 Jordan	356744	4512
42 Lebanon	348793	4206
43 Panama	333755	5672
44 Slovakia	287752	6350
45 Malaysia	274875	1030
46 Belarus	273659	1885
47 Nepal	273166	2058
48 Ecuador	269860	15444
49 Georgia	267701	3406
50 Kazakhstan	254712	3246
51 Croatia	239254	5407
52 Bolivia (Plurinational State of)	238495	11303
53 Bulgaria	234428	9786
54 Dominican Republic	232780	3007
55 Azerbaijan	232636	3190
56 Tunisia	226015	7684
57 Ireland	212647	4082
58 Denmark	206065	2319
59 Costa Rica	200888	2745
60 occupied Palestinian territory, including east...	193418	2155

Tabela A.3: Dados das ocorrências globais disponibilizadas pela OMS, cujos países são ordenados pelo número de casos cumulativos apresentados. São listados os países que ocupam a colocação de 61° a 90°.

País	Casos cumulativos	Mortes cumulativas
61 Lithuania	192986	3145
62 Slovenia	182484	4008
63 Kuwait	181484	1027
64 Egypt	176333	10201
65 Greece	176059	6221
66 Republic of Moldova	174155	3717
67 Armenia	170011	3158
68 Guatemala	168880	6164
69 Honduras	162584	3933
70 Qatar	159053	256
71 Nigeria	150246	1803
72 Ethiopia	150179	2249
73 Paraguay	147253	2991
74 Myanmar	141709	3192
75 Oman	138494	1549
76 Venezuela (Bolivarian Republic of)	134319	1297
77 Libya	128740	2061
78 Bosnia and Herzegovina	127135	4962
79 Bahrain	115705	414
80 Algeria	111418	2950
81 Kenya	103615	1807
82 China	101624	4842
83 Puerto Rico	98506	1940
84 North Macedonia	98174	3018
85 Albania	96838	1617
86 Republic of Korea	86128	1550
87 Kyrgyzstan	85772	1450
88 Uzbekistan	79598	622
89 Latvia	79351	1521
90 Ghana	78689	568

Tabela A.4: Dados das ocorrências globais disponibilizadas pela OMS, cujos países são ordenados pelo número de casos cumulativos apresentados. São listados os países que ocupam a colocação de 91° a 120°.

	País	Casos cumulativos	Mortes cumulativas
91	Sri Lanka	78420	430
92	Zambia	72467	991
93	Montenegro	70656	916
94	Norway	67497	607
95	Kosovo	65307	1548
96	Singapore	59832	29
97	El Salvador	58023	1776
98	Estonia	56246	518
99	Afghanistan	55575	2430
100	Luxembourg	53562	619
101	Mozambique	53527	571
102	Finland	52209	725
103	Uruguay	50208	553
104	Cuba	41688	282
105	Uganda	40154	331
106	Namibia	36680	394
107	Zimbabwe	35543	1420
108	Cyprus	32935	228
109	Cameroon	32681	500
110	Senegal	32099	781
111	Côte d'Ivoire	31612	183
112	Sudan	30128	1864
113	Malawi	29634	984
114	Australia	28912	909
115	Botswana	25812	226
116	Thailand	25241	83
117	Democratic Republic of the Congo	24793	694
118	Angola	20452	498
119	Malta	20409	301
120	Jamaica	20310	384

Tabela A.5: Dados das ocorrências globais disponibilizadas pela OMS, cujos países são ordenados pelo número de casos cumulativos apresentados. São listados os países que ocupam a colocação de 121° a 150°.

País	Casos cumulativos	Mortes cumulativas
121 Madagascar	19598	292
122 Maldives	18340	58
123 French Polynesia	18335	136
124 Rwanda	17835	243
125 Mauritania	17058	433
126 Eswatini	16709	640
127 French Guiana	16456	80
128 Guinea	15154	85
129 Syrian Arab Republic	15045	990
130 Cabo Verde	14885	142
131 Mayotte	14492	89
132 Tajikistan	13714	91
133 Gabon	13107	75
134 Haiti	12206	247
135 Belize	12195	314
136 Burkina Faso	11703	139
137 Réunion	11562	48
138 Andorra	10610	107
139 Lesotho	10455	285
140 Guadeloupe	9455	160
141 Suriname	8832	168
142 Congo	8419	123
143 Bahamas	8403	179
144 Guyana	8313	189
145 Mali	8256	342
146 Trinidad and Tobago	7663	138
147 Aruba	7513	69
148 Guam	7497	130
149 Martinique	6593	45
150 Togo	6085	81

Tabela A.6: Dados das ocorrências globais disponibilizadas pela OMS, cujos países são ordenados pelo número de casos cumulativos apresentados. São listados os países que ocupam a colocação de 151° a 180°.

País	Casos cumulativos	Mortes cumulativas
151 South Sudan	6084	83
152 Iceland	6044	29
153 Djibouti	6008	63
154 Equatorial Guinea	5747	88
155 Somalia	5632	179
156 Benin	5143	65
157 Nicaragua	5106	172
158 Central African Republic	4996	63
159 Niger	4718	170
160 Curaçao	4658	22
161 Gambia	4510	139
162 Gibraltar	4226	87
163 Sierra Leone	3829	79
164 Chad	3751	132
165 Comoros	3458	138
166 San Marino	3420	72
167 Jersey	3207	68
168 Guinea-Bissau	3025	46
169 Eritrea	2675	7
170 Saint Lucia	2620	25
171 Liechtenstein	2617	51
172 Barbados	2608	29
173 United States Virgin Islands	2558	25
174 Mongolia	2493	2
175 Viet Nam	2347	35
176 Seychelles	2264	10
177 Yemen	2158	619
178 Sint Maarten	2013	27
179 New Zealand	1992	26
180 Liberia	1985	85

Tabela A.7: Dados das ocorrências globais disponibilizadas pela OMS, cujos países são ordenados pelo número de casos cumulativos apresentados. São listados os países que ocupam a colocação de 181° a 210°.

País	Casos cumulativos	Mortes cumulativas
181 Burundi	1955	3
182 Turks and Caicos Islands	1928	13
183 Monaco	1835	22
184 Sao Tome and Principe	1548	20
185 Saint Vincent and the Grenadines	1474	6
186 Saint Martin	1456	12
187 Papua New Guinea	961	10
188 Bhutan	866	1
189 Guernsey	806	14
190 Other	745	13
191 Bermuda	697	12
192 Faroe Islands	658	1
193 Mauritius	603	10
194 Antigua and Barbuda	525	10
195 United Republic of Tanzania	509	21
196 Cambodia	484	0
197 Saint Barthélemy	475	0
198 Isle of Man	437	25
199 Cayman Islands	419	2
200 Bonaire	381	4
201 Brunei Darussalam	185	3
202 British Virgin Islands	153	1
203 Grenada	148	1
204 Northern Mariana Islands (Commonwealth of the)	135	2
205 Dominica	134	0
206 Timor-Leste	103	0
207 Fiji	56	2
208 New Caledonia	55	0
209 Falkland Islands (Malvinas)	49	0
210 Lao People's Democratic Republic	45	0

Tabela A.8: Dados das ocorrências globais disponibilizadas pela OMS, cujos países são ordenados pelo número de casos cumulativos apresentados. São listados os países que ocupam a colocação de 211° a 237°.

País	Casos cumulativos	Mortes cumulativas
211 Saint Kitts and Nevis	41	0
212 Greenland	30	0
213 Holy See	26	0
214 Saint Pierre and Miquelon	24	0
215 Montserrat	20	1
216 Sint Eustatius	20	0
217 Anguilla	18	0
218 Solomon Islands	18	0
219 Wallis and Futuna	9	0
220 Saba	6	0
221 Marshall Islands	4	0
222 Samoa	4	0
223 Vanuatu	1	0
224 American Samoa	0	0
225 Cook Islands	0	0
226 Democratic People's Republic of Korea	0	0
227 Kiribati	0	0
228 Micronesia (Federated States of)	0	0
229 Nauru	0	0
230 Niue	0	0
231 Palau	0	0
232 Pitcairn Islands	0	0
233 Saint Helena	0	0
234 Tokelau	0	0
235 Tonga	0	0
236 Turkmenistan	0	0
237 Tuvalu	0	0

Tabela A.9: Dados das ocorrências nacionais, reportados pelos municípios por meio de boletins epidemiológicos. São listados os estados brasileiros e o Distrito Federal, em ordem alfabética, contendo a população estimada de 2019, os casos cumulativos e os óbitos registrados pelos municípios.

Estado	População Estimada de 2019	Casos cumulativos	Mortes cumulativas
AC	881935	32515	707
AL	3337357	92166	2285
AM	4144597	168089	4681
AP	845731	54056	771
BA	14873064	369259	7882
CE	9132078	282471	9430
DF	3015268	218479	3797
ES	4018650	166999	3982
GO	7018354	266553	5936
MA	7075181	189048	4154
MG	21168791	376537	9259
MS	2778986	86428	1674
MT	3484466	151391	4026
PA	8602865	261131	6808
PB	4018127	137474	3187
PE	9557071	168880	8794
PI	3273227	119456	2516
PR	11433957	227823	5610
RJ	17264943	322383	21090
RN	3506853	83463	2622
RO	1777225	74248	1492
RR	605761	59946	705
RS	11377239	271340	6125
SC	7164788	284982	3282
SE	2298696	85804	2250
SP	45919049	1156652	40202
TO	1572866	77528	1125