

Modelo SEIQHRF em Redes

Departamento de Estatística da UFMG

Aluno: Luís Assunção, Orientador: Adrian Luna

Introdução

- ▶ Epidemias em populações fechadas podem ser modeladas como processos markovianos com um número finito de estados
- ▶ Na UFMG, há o projeto *COVID-19: proposta de um modelo epidemiológico que incorpora estruturas sociais de contágio*
- ▶ *SEIQHRF* (Susceptible-Exposed-Quarantined-Hospitalized-Recovered-Fatality)

Metodologia

Conjuntos de estados

- ▶ Compartimentalização de indivíduos de uma população

SIR

- ▶ *SIR* (Susceptible-Exposed-Recovered)
- ▶ Conjunto utilizado para doenças infecciosas que conferem imunidade vitalícia (caxumba, por exemplo)

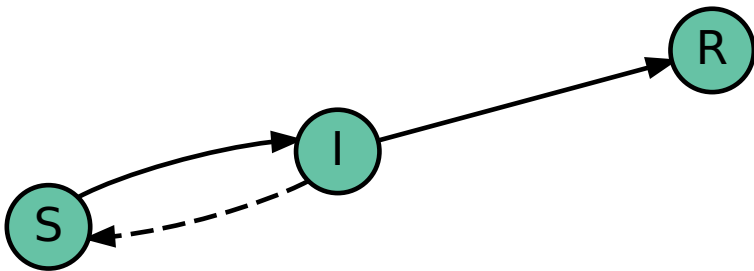


Figure 1: SIR. Linhas pontilhadas são infecções e linhas sólidas são progressões.

SEIQHRF

- ▶ Susceptible-Exposed-Quarantined-Hospitalized-Recovered-Fatality
- ▶ Criado para modelar as consequências sociais e estruturais da pandemia do COVID-19

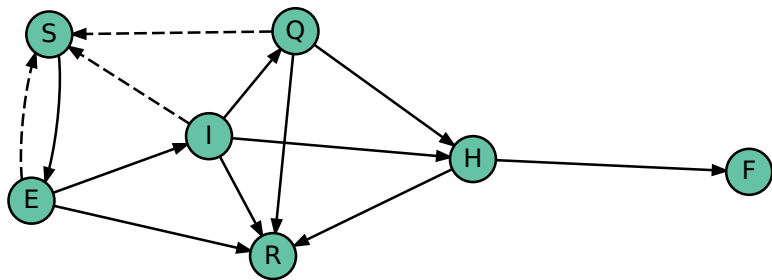


Figure 2: SEIQHRF. Linhas pontilhadas são infecções e linhas sólidas são progressões.

Modelos epidemiológicos

Tipos

- ▶ Determinísticos (equações diferenciais),
- ▶ Estocásticos (processos estocásticos),
- ▶ Em redes (simulações numéricas)

Redes

- ▶ Vantagem: expressam estruturas sociais complexas
- ▶ Desvantagem: tempo computacional

- ▶ Implementamos simulações numéricas do *SEIQHRF* em redes, mas para simplificar a apresentação da metodologia, mostraremos as definições do *SIR*
- ▶ Essas definições podem ser generalizadas para o *SEIQHRF*

SIR em redes

- ▶ Seja o grafo $Y = (E, V)$, $Y \in \mathcal{Y}$
- ▶ Vértices V são indivíduos e arestas E são relações

- Seja o processo $X_i(t) = S, I, R, \quad i \in V$

$$\Pr(X_i(t + \delta t) = x'_i | X_i(t) = x_i) = \begin{cases} \beta M_i(x) \delta t & \text{se } x_i = S \text{ e } x'_i = I \\ \gamma \delta t & \text{se } x_i = I \text{ e } x'_i = R \\ 1 - [\beta M_i(x) + \gamma] \delta t & \text{c.c.,} \end{cases}$$

- $M_i(x)$ é o número de vizinhos infectados do vértice i

Grafos aleatórios

$$\{\Pr_{\theta}(Y), Y \in \mathcal{Y}_n, \theta \in \Theta\},$$

$$\mathcal{Y}_n = \{Y = (E, V), |V| = n, Y \text{ simples}\}$$

Exponential random graph models

Sejam

- ▶ $g : \mathcal{Y} \rightarrow \mathbb{R}^p$ estatísticas suficientes
- ▶ $\theta \in \mathbb{R}^q$ um vetor de q parâmetros
- ▶ $\eta(\theta) : \mathbb{R}^q \rightarrow \mathbb{R}^p$ parâmetros naturais
- ▶ O modelo exponencial para grafos aleatórios é definido como:

$$\Pr_{\eta, g}(Y = y | \theta) = \frac{\exp(\eta(\theta) \cdot g(y))}{c_{\eta, g}(\theta)}, \quad y \in \mathcal{Y}$$

- ▶ No caso mais comum, tratado neste trabalho, $\eta(\theta) = \theta$.

Temporal ERGM

- ▶ ERGMs temporais
- ▶ Sejam $g : \mathcal{Y}^2 \rightarrow \mathbb{R}^p$
- ▶ A probabilidade de transição entre y^{t-1} e y^t é definida como:

$$\Pr_{\eta, g}(Y^t = y^t | Y^{t-1} = y^{t-1}; \theta) = \frac{\exp(\eta(\theta) \cdot g(y^t, y^{t-1}))}{c_{\eta, g}(\theta, y^{t-1})}, \quad y^t, y^{t-1} \in \mathcal{Y}$$

- ▶ Ajuste via Metropolis-Hastings MCMCMLE
- ▶ São Separable TERGMs (STERGMs) quando os coeficientes de formação e dissolução de arestas são independentes

Implementação

- ▶ Objetivo: implementar o *SEIQHRF* em redes
- ▶ Ferramentas: R, *ergm*, *EpiModel*

ergm

Ajuste

- ▶ `stergm(formula)`: ajuste do STERGM

Estatísticas

- ▶ `edges`: número de arestas - tamanho do conjunto $\{(i,j)\}$
- ▶ `nodematch(attr, diff = FALSE)`: homofilia uniforme - tamanho do conjunto $\{(i,j) | \text{atributo}_i = \text{atributo}_j\}$
- ▶ `nodematch(attr, diff = TRUE)` - homofilia diferencial. p estatísticas da rede, onde p é o tamanho do conjunto $\{\text{atributos}\}$. $\{(i,j) | \text{atributo}_i = \text{atributo}_j = k\}$, onde o valor de k é o menor valor único do atributo em questão.
- ▶ `absdiff(attr)` diferença absoluta - $|\text{atributo}_i - \text{atributo}_j|$.
- ▶ `gwesp(decay, fixed)`: ...

EpiModel

- ▶ API para ajustar modelos epidemiológicos
- ▶ Modelos complexos são programados de forma modular (nosso caso)
- ▶ Modelos comuns são triviais de se ajustar

Exemplo SIR Determinístico vs. Estocástico

- ▶ $N = 1000$, $\beta = 0.16$ e $\gamma = 0.02$
- ▶ Quando $t = 0$, $S = 90$ e $I = 10$

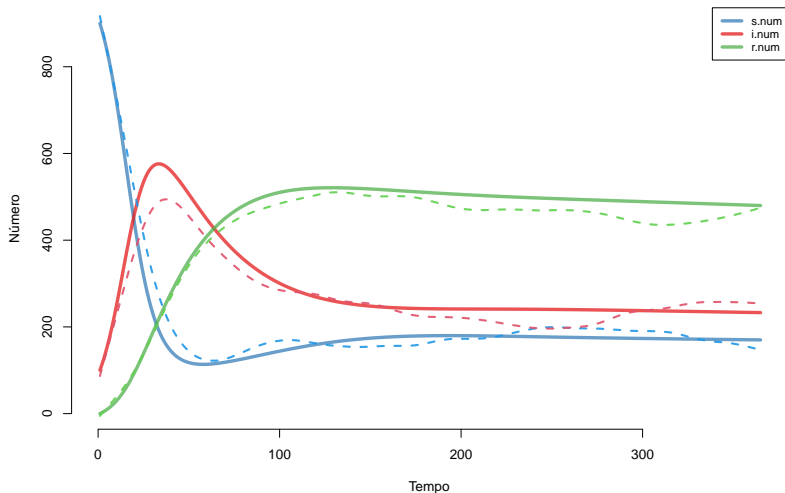


Figure 3: Comparação entre ajustes DCM (linhas sólidas) e ICM (linhas tracejadas).

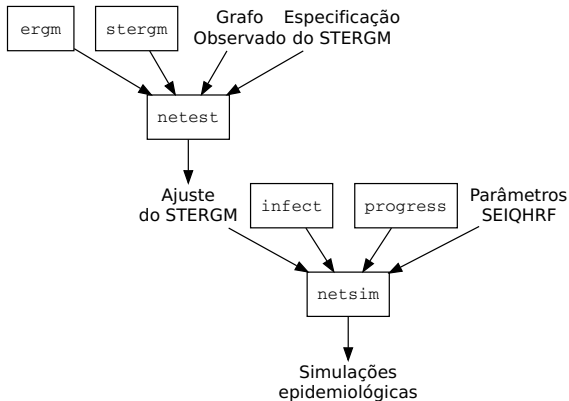


Figure 4: Fluxograma da implementação.

Algorithm 1: netsim

Result: Simulações epidemiológicas

Input : Ajuste do STERGM, Número de simulações, Tempo das simulações, Função de infecção (infect), Função de progressão (progress), Parâmetros das funções de infecção e progressão

Output: Lista de matrizes, cada uma representando uma simulação, com o número de indivíduos em cada estado em cada tempo.

```
1 for simulação = 1 até simulação = número de simulações do
2   for simulação = 1 até simulação = número de simulações do
3     simular um STERGM a partir do ajuste
4     for tempo = 1 até tempo = tempo da simulação do
5       formar ou dissolver arestas do STERGM
6       rodar o processo de infecção
7       rodar o processo de progressão
8       salvar os resultados
9     end
10  end
11 end
```

Simulações e resultados

Faux Magnolia High

- ▶ Disponível no `ergm`
- ▶ Simulação de rede de rede de amigos de uma escola do sul dos EUA
- ▶ Existem 1461 alunos, e cada aluno tem três atributos: Grade (turma), Race (raça) e Sex (sexo)

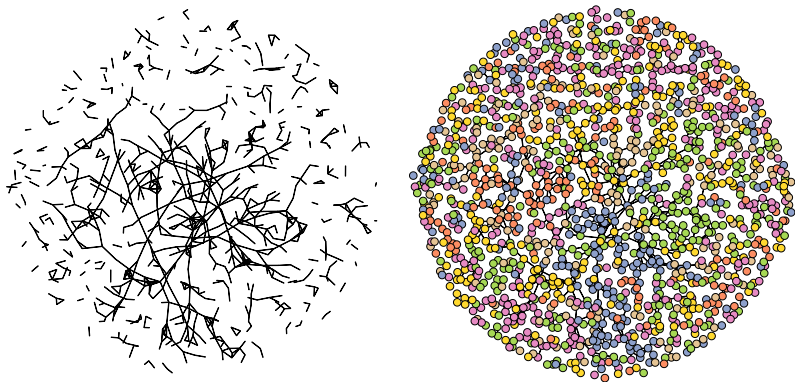


Figure 5: Faux Magnolia High. À esquerda, apenas as arestas do grafo. À direita, os vértices coloridos de acordo com a turma do aluno.

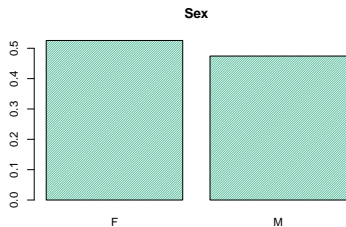
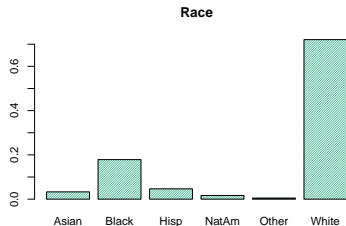
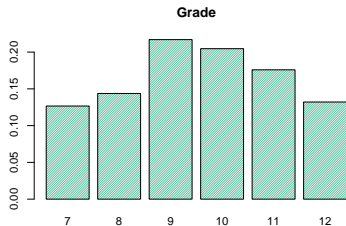


Figure 6: Proporções dos níveis de cada atributo do Faux Magnolia High.

Ajuste do ERGM

- Especificado na documentação do `faux.magnolia.high`

```
magnolia ~  
  edges +  
  nodematch("Grade", diff = TRUE) +  
  nodematch("Race", diff = TRUE) +  
  nodematch("Sex", diff = FALSE) +  
  absdiff("Grade") +  
  gwesp(0.25, fixed = TRUE)
```


Table 1: Coeficientes do ajuste Magnolia.

Estatística	Fator	Nível	Coeficiente	OR
edges	NA	NA	-8.27	0.00
nodematch	Grade	7	1.50	4.50
nodematch	Grade	8	1.42	4.13
nodematch	Grade	9	0.96	2.60
nodematch	Grade	10	1.06	2.88
nodematch	Grade	11	1.20	3.32
nodematch	Grade	12	1.29	3.63
nodematch	Race	Asian	1.96	7.07
nodematch	Race	Black	1.36	3.89
nodematch	Race	Hisp	0.36	1.43
nodematch	Race	NatAm	3.46	31.86
nodematch	Race	Other	-Inf	0.00
nodematch	Race	White	0.88	2.42
nodematch	Sex	NA	0.73	2.07
absdiff	Grade	NA	-0.91	0.40
gwesp	NA	NA	1.86	6.43

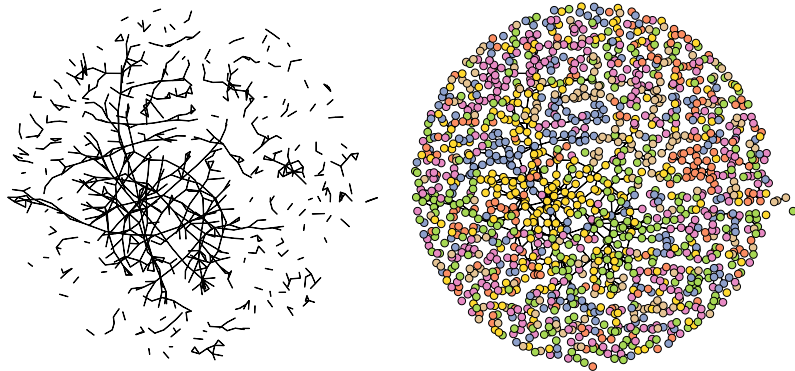


Figure 7: Simulação gerada com nosso ajuste do Faux Magnolia High. À esquerda, apenas as arestas do grafo. À direita, os vértices coloridos de acordo com a turma do aluno.

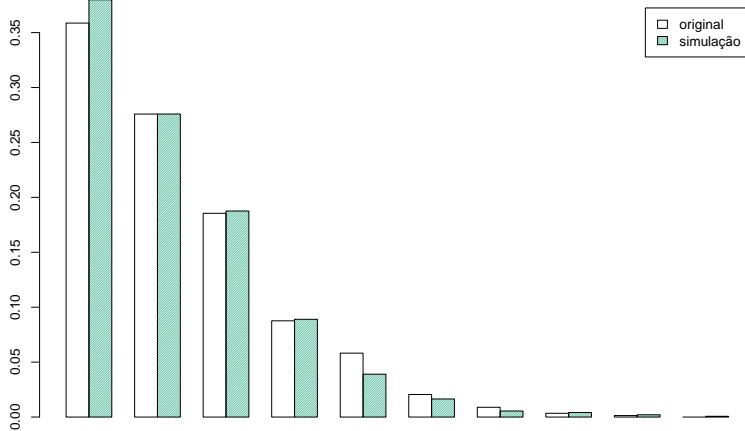


Figure 8: Comparação entre as distribuições de grau do grafo original e da simulação.

Simulações epidemiológicas

- ▶ $t_{max} = 170$ dias
- ▶ 30 simulações por cenário
- ▶ No $t = 0$, 20 indivíduos expostos e 10 infectados,
- ▶ 5 vagas no estado de hospitalização
- ▶ Probabilidade de transição $H \rightarrow F$ é 1.5x maior em caso de superlotação de hospitais

Cenário A

Table 2: Parâmetros de infecção do Cenário A.

De	Para	Encontros/dia	P(Infecção)	P(Encontro e infecção)
E	S	10.0	0.02	0.18
I	S	10.0	0.05	0.40
Q	S	2.5	0.02	0.05

Table 3: Parâmetros de progressão do Cenário A.

De	Para	Transições/dia
E	I	0.10
E	R	0.05
I	Q	0.03
I	H	0.03
I	R	0.05
Q	H	0.03
Q	R	0.05
H	R	0.07
H	F	0.02

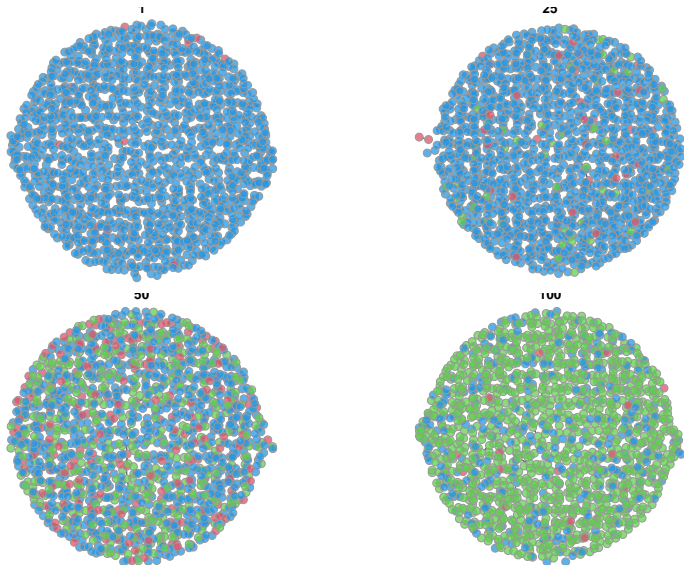


Figure 9: Rede ao longo de uma das simulações do Cenário A. O número acima da rede é o número de dias percorridos, e as cores indicam vértices suscetíveis (azul), em estados de infecção (vermelho) e recuperados

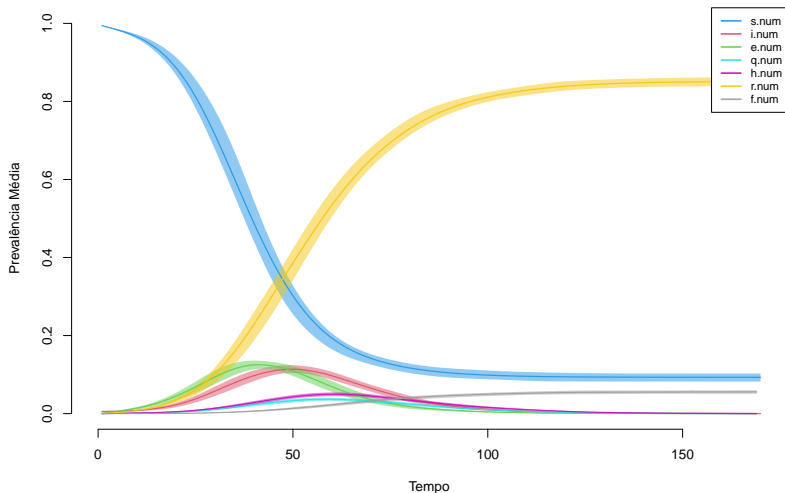


Figure 10: Prevalência ao longo do tempo no Cenário A. Linhas são prevalências médias, e áreas são intervalos de prevalência.

Podemos ver a distribuição por turma...

Cenário B

- ▶ Reduzimos as taxas de encontro em 70%, simulando isolamento social

Table 4: Parâmetros de infecção do Cenário B.

De	Para	Encontros/dia	P(Infecção)	P(Encontro e infecção)
E	S	3.00	0.02	0.06
I	S	3.00	0.05	0.14
Q	S	0.75	0.02	0.02

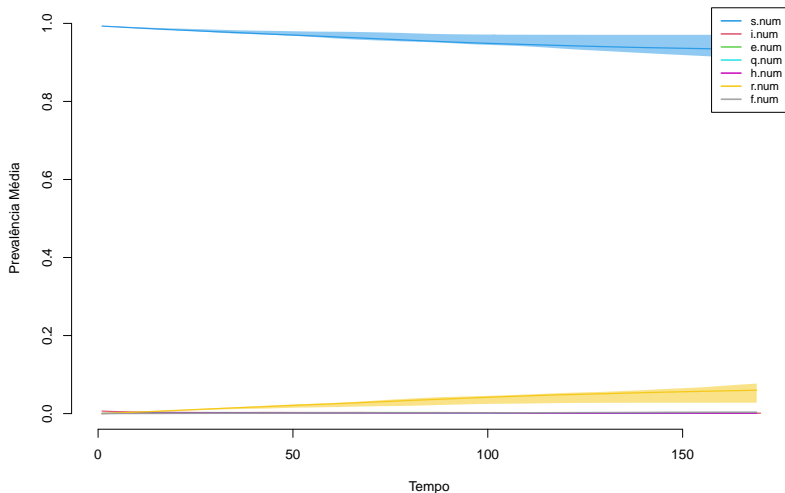


Figure 11: Prevalência ao longo do tempo no Cenário A. Linhas são prevalências médias, e áreas são intervalos de prevalência.

Considerações

- ▶ A implementação foi bem sucedida
- ▶ O modelo está no Github para qualquer pesquisador ou pesquisadora utilizar

Trabalhos futuros

- ▶ Simular o *SEIQHRF* em redes no contexto da rede do Aglomerado da Serra

Obrigado!

Referências

- ▶ ANDERSSON, H.; BRITTON, T. Stochastic epidemic models and their statistical analysis, volume 151 of lecture notes in statistics. In: . [S.l.: s.n.], 2000. v. 151. Citado na página 27.
- ▶ CARNEGIE, N. et al. An approximation method for improving dynamic network model fitting. Journal of Computational and Graphical Statistics, v. 24, p. 00–00, 05 2014. Citado na página 32.

- ▶ CHURCHES, T. Tim Churches Health Data Science Blog: Modelling the effects of public health interventions on COVID-19 transmission using R - part 2. 2020. Disponível em: <<<https://timchurches.github.io/blog/posts/2020-03-18-modelling-the-effects-of-public-health-interventions-on-covid-19-transmission-part-2/>>>. Citado 3 vezes nas páginas 23, 26 e 43.
- ▶ EXPONENTIAL Random Graph Models for Social Networks: Theory, Methods, and Applications. [S.l.]: Cambridge University Press, 2012. (Structural Analysis in the Social Sciences). Citado na página 28.

- ▶ GEYER, C. J.; THOMPSON, E. A. Constrained monte carlo maximum likelihood for dependent data. Journal of the Royal Statistical Society. Series B (Methodological), [Royal Statistical Society, Wiley], v. 54,
n. 3, p. 657–699, 1992. ISSN 00359246. Disponível em: <<http://www.jstor.org/stable/2345852>>. Citado na página 29.
- ▶ GIORDANO, G. et al. Modelling the covid-19 epidemic and implementation of population-wide interventions in italy. Nature Medicine, Springer Science and Business Media LLC, v. 26, n. 6, p. 855–860, Apr 2020. ISSN 1546-170X. Disponível em: <<http://dx.doi.org/10.1038/s41591-020-0883-7>>. Citado na página 25.

- ▶ HANDCOCK, M. S. et al. ergm: Fit, Simulate and Diagnose Exponential-Family Models for Networks. [S.l.],

2020. R package version 3.11.0. Disponível em:
<<https://CRAN.R-project.org/package=ergm>>. Citado 2
vezes nas páginas 31 e 39.

- ▶ HUNTER, D.; HANDCOCK, M. Inference in curved exponential family models for networks. Journal of Computational and Graphical Statistics, American Statistical Association, v. 15, n. 3, p. 565–583, set. 2006. ISSN 1061-8600. Funding Information: The authors are grateful to Steven Goodreau and Tom Snijders for very helpful comments and discussions. This research is supported by Grant DA012831 from NIDA and Grant HD041877 from NICHD. Citado 2 vezes nas páginas 28 e 29.

- ▶ JENNESS, S. M.; GOODREAU, S. M.; MORRIS, M. Epimodel: An r package for mathematical modeling of infectious disease over networks. *Journal of Statistical Software, Articles*, v. 84, n. 8, p. 1–47, 2018. ISSN 1548-7660. Disponível em: <<https://www.jstatsoft.org/v084/i08>>. Citado 2 vezes nas páginas 31 e 32.
- ▶ KEELING, M. J.; EAMES, K. T. Networks and epidemic models. *Journal of The Royal Society Interface*, v. 2, n. 4, p. 295–307, 2005. Disponível em: <<https://royalsocietypublishing.org/doi/abs/10.1098/rsif.2005.0051>>. Citado 3 vezes nas páginas 25, 26 e 27.

- ▶ KENDALL, D. G. Deterministic and stochastic epidemics in closed populations. In: . Volume 4 Contributions to Biology and Problems of Health. University of California Press, 1956. p. 149–166. Disponível em:
<<https://doi.org/10.1525/9780520350717-011>>. Citado na página 23.
- ▶ KRIVITSKY, P. N.; HANDCOCK, M. S. A separable model for dynamic networks. Journal of the Royal Statistical Society: Series B (Statistical Methodology), Wiley, v. 76, n. 1, p. 29–46, Mar 2013. ISSN 1369-7412. Disponível em:
<<http://dx.doi.org/10.1111/rssb.12014>>. Citado na página 28.

- ▶ R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria, 2021. Disponível em: <<https://www.R-project.org/>>. Citado na página 31.
- ▶ RESNICK, M. et al. Protecting adolescents from harmfindings from the national longitudinal study on adolescent health. JAMA : the journal of the American Medical Association, v. 278, p. 823–32, 09 1997. Citado na página 39.