

Modelo epidemiológico com compartimentos SEIQHRF em redes: implementação e inferência via MCMC

Modelo epidemiológico com compartimentos SEIQHRF em redes: implementação e inferência via MCMC

Modelo canônico de trabalho monográfico acadêmico em conformidade com as normas ABNT apresentado à comunidade de usuários \LaTeX .

Universidade Federal de Minas Gerais
Instituto de Ciências Exatas
Departamento de Estatística

Orientador: Adrian Luna

Brasil

Modelo epidemiológico com 20 compartimentos SEIQHRF em redes: implementação e inferência via MCMC/ . – Brasil, -
Orientador: Adrian Luna

Monografia – Universidade Federal de Minas Gerais
Instituto de Ciências Exatas
Departamento de Estatística

Palavra-chave2. 2. Palavra-chave3. I. Orientador. II. Universidade xxx. III. Faculdade de xxx. IV. Título

Errata

Elemento opcional da ??, 4.2.1.2). Exemplo:

FERRIGNO, C. R. A. **Tratamento de neoplasias ósseas apendiculares com reimplantação de enxerto ósseo autólogo autoclavado associado ao plasma rico em plaquetas:** estudo crítico na cirurgia de preservação de membro em cães. 2011. 128 f. Tese (Livre-Docência) - Faculdade de Medicina Veterinária e Zootecnia, Universidade de São Paulo, São Paulo, 2011.

Folha	Linha	Onde se lê	Leia-se
1	10	auto-conclavo	autoconclavo

Modelo epidemiológico com compartimentos SEIQHRF em redes: implementação e inferência via MCMC

Modelo canônico de trabalho monográfico acadêmico em conformidade com as normas ABNT apresentado à comunidade de usuários L^AT_EX.

Trabalho aprovado. Brasil, 24 de novembro de 2012:

Adrian Luna
Orientador

Professor
Convidado 1

Professor
Convidado 2

Brasil

*Este trabalho é dedicado às crianças adultas que,
quando pequenas, sonharam em se tornar cientistas.*

Agradecimentos

Os agradecimentos principais são direcionados à Gerald Weber, Miguel Frasson, Leslie H. Watter, Bruno Parente Lima, Flávio de Vasconcellos Corrêa, Otavio Real Salvador, Renato Machniewscz¹ e todos aqueles que contribuíram para que a produção de trabalhos acadêmicos conforme as normas ABNT com L^AT_EX fosse possível.

Agradecimentos especiais são direcionados ao Centro de Pesquisa em Arquitetura da Informação² da Universidade de Brasília (CPAI), ao grupo de usuários *latex-br*³ e aos novos voluntários do grupo *abnT_EX2*⁴ que contribuíram e que ainda contribuirão para a evolução do abnT_EX2.

¹ Os nomes dos integrantes do primeiro projeto abnT_EX foram extraídos de <<http://codigolivres.org.br/projects/abntex/>>

² <<http://www.cpai.unb.br/>>

³ <<http://groups.google.com/group/latex-br>>

⁴ <<http://groups.google.com/group/abntex2>> e <<http://www.abntex.net.br/>>

*Someday there'll be a cure for pain
That's the day I throw my drugs away
(Mark Sandman)*

Resumo

Segundo a ??, 3.1-3.2), o resumo deve ressaltar o objetivo, o método, os resultados e as conclusões do documento. A ordem e a extensão destes itens dependem do tipo de resumo (informativo ou indicativo) e do tratamento que cada item recebe no documento original. O resumo deve ser precedido da referência do documento, com exceção do resumo inserido no próprio documento. (...) As palavras-chave devem figurar logo abaixo do resumo, antecedidas da expressão Palavras-chave:, separadas entre si por ponto e finalizadas também por ponto.

Palavras-chave: latex. abntex. editoração de texto.

Abstract

This is the english abstract.

Keywords: latex. abntex. text editoration.

Lista de ilustrações

Figura 1 – SIR. Linhas pontilhadas são infecções e linhas sólidas são progressões.	27
Figura 2 – Comparação entre ajustes determinístico (DCM, linhas sólidas) e estocástico (ICM, linhas tracejadas).	28
Figura 3 – SEIQHRF. Linhas pontilhadas são infecções e linhas sólidas são progressões.	29
Figura 4 – Faux Magnolia High. À esquerda, apenas as arestas do grafo. À direita, os vértices coloridos de acordo com a turma do aluno.	33
Figura 5 – 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.	34

Lista de tabelas

Tabela 1 – tabela de...	34
Tabela 2 – Coeficientes do ajuste Magnolia.	35

Sumário

	Sumário	23
1	INTRODUÇÃO	25
2	METODOLOGIA	27
2.1	Epidemiologia	27
2.1.1	Modelos compartimentais	27
2.1.2	Modelos em redes	28
2.2	SEIQRD	28
2.3	Grafos aleatórios	29
2.3.1	Exponential random graph models	29
3	IMPLEMENTAÇÃO	31
4	SIMULAÇÃO	33
4.1	Síntese dos dados	33
4.1.1	Faux Magnolia High	33
4.1.2	Ajuste	34
5	RESULTADOS	37
6	CONCLUSÃO	39
	REFERÊNCIAS	41
	APÊNDICES	43
	APÊNDICE A – CÓDIGO	45

Sumário

	Sumário	23
1	INTRODUÇÃO	25
2	METODOLOGIA	27
2.1	Epidemiologia	27
2.1.1	Modelos compartimentais	27
2.1.2	Modelos em redes	28
2.2	SEIQRD	28
2.3	Grafos aleatórios	29
2.3.1	Exponential random graph models	29
3	IMPLEMENTAÇÃO	31
4	SIMULAÇÃO	33
4.1	Síntese dos dados	33
4.1.1	Faux Magnolia High	33
4.1.2	Ajuste	34
5	RESULTADOS	37
6	CONCLUSÃO	39
	REFERÊNCIAS	41
	APÊNDICES	43
	APÊNDICE A – CÓDIGO	45

1 Introdução

O avanço de epidemias em populações fechadas é um problema que pode ser modelado como um processo Markoviano temporal com um número finito de estados, também chamados de compartimentos ([KENDALL, 1956](#)). As estimativas fornecidas por esses modelos sob o efeito de diferentes intervenções informam a construção de políticas públicas eficientes no combate às doenças estudadas.

No último ano, muitos modelos foram propostos para estimar a propagação e os impactos da doença COVID-19, causada pelo vírus Sars-CoV-2. Alguns conjuntos de estados foram criados com o propósito de estimar os efeitos sociais e estruturais da pandemia (??). Um desses conjuntos é o SEIQHRF (Susceptible-Exposed-Quarantined-Hospitalized-Recovered-Fatality) ([CHURCHES, 2020](#)).

Na UFMG, o projeto *COVID-19: proposta de um modelo epidemiológico que incorpora estruturas sociais de contágio* tem mapeado o grafo de relações pessoais do Aglomerado da Serra em Belo Horizonte, com o objetivo de modelar a progressão da COVID-19 em contextos de aglomerações urbanas. Esse projeto motivou a implementação de modelos de redes utilizando os estados SEIQHRF nesta monografia.

2 Metodologia

Neste capítulo, descreveremos modelos epidemiológicos, ...

2.1 Epidemiologia

2.1.1 Modelos compartimentais

A maior parte dos modelos epidemiológicos é baseado na compartimentalização de indivíduos de acordo com seus respectivos estados da doença (KEELING; EAMES, 2005). Esses modelos descrevem a proporção de indivíduos de uma população em cada estado ao longo do tempo. O modelo *SIR*, ilustrado na Figura 1, tem três estados:

- **Susceptible**: suscetíveis. Esses indivíduos podem entrar em contato com infectados e prosseguir para o estado infectado;
- **Infected**: infectados. Podem entrar em contato com suscetíveis e infectá-los, ou prosseguir para o estado recuperado;
- **Recovered**: recuperados.

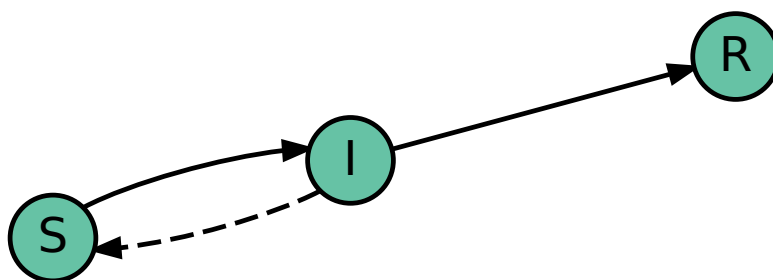


Figura 1 – SIR. Linhas pontilhadas são infecções e linhas sólidas são progressões.

O SIR é adequado para doenças infecciosas que conferem imunidade vitalícia, como caxumba. Sem dinâmicas vitais (processos naturais de nascimento e morte) e com uma população com mistura aleatória (cada indivíduo tem uma probabilidade pequena e idêntica de se encontrar com qualquer outro indivíduo), o *SIR* pode ser expressado pelo conjunto de equações diferenciais (KEELING; EAMES, 2005):

$$\begin{aligned}\frac{dS}{dt} &= -\beta SI \\ \frac{dI}{dt} &= \beta SI - \gamma I \\ \frac{dR}{dt} &= \gamma I\end{aligned}$$

Onde β é... pode ser um vetor para expressar diferentes subpopulações... γ é...

Esses modelos podem não possuem soluções analíticas exatas, e podem ser ajustados de forma determinística (aproximada) ou estocástica. (??). A Figura 2... usando o pacote `EpiModel` (JENNESS; GOODREAU; MORRIS, 2018)... DCM e ICM.

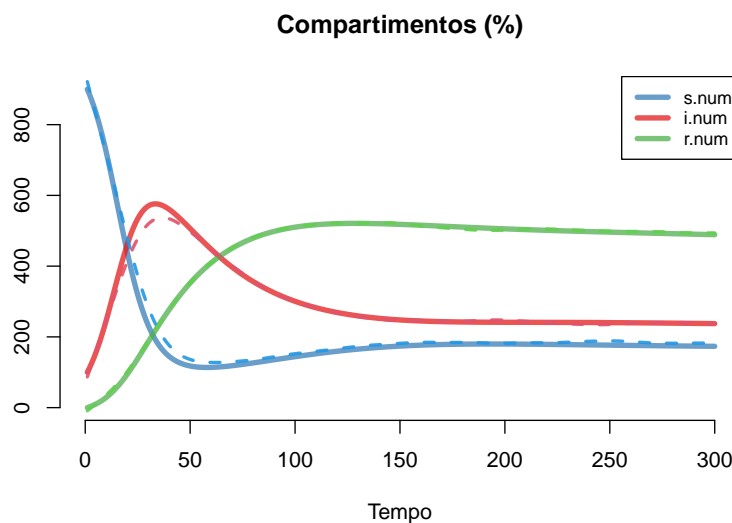


Figura 2 – Comparação entre ajustes determinístico (DCM, linhas sólidas) e estocástico (ICM, linhas tracejadas).

2.1.2 Modelos em redes

Nem sempre é razoável presumir que uma população ou algumas subpopulações se misturam aleatoriamente. Muitas vezes, temos conhecimento de relações determinísticas entre indivíduos, que condicionam as probabilidades de contato entre os pares. Modelos em redes...

Não há soluções analíticas para essa classe de modelos, que são ajustados numericamente por meio de algoritmos de Markov Chain Monte Carlo (??)...

Um exemplo de simulação na Figura...

... (ANDERSSON; BRITTON, 2000)

2.2 SEIQRHD

Muitos conjuntos de estados foram propostos SIDARTHE (GIORDANO et al., 2020)...

O SEIQHRF (CHURCHES, 2020), ilustrado na Figura 3, adiciona quatro estados ao SIR:

- **Exposed:** infectado e assintomático. Pode infectar o suscetível;
- **Quarantined:** infectado e sintomático. Pode infectar o suscetível, com menor probabilidade de encontro que os demais estados infectados;
- **Hospitalized:** infectado e sintomático. Não encontra suscetíveis e tem maior probabilidade de morte;
- **Fatality:** morte.

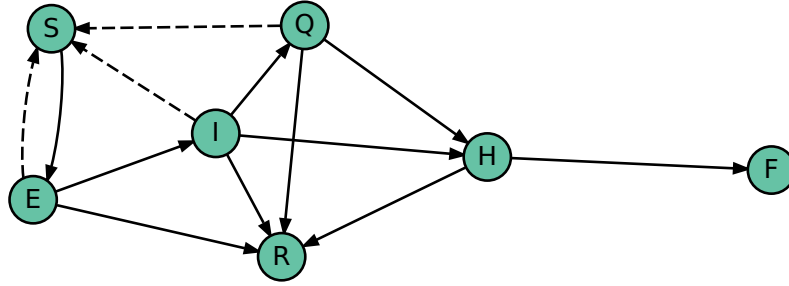


Figura 3 – SEIQRHF. Linhas pontilhadas são infecções e linhas sólidas são progressões.

2.3 Grafos aleatórios

Formalmente, um grafo aleatório $Y \in \mathcal{Y}$ consiste em um conjunto de n vértices e m arestas $\{Y_{ij} : i = 1, \dots, n; j = 1, \dots, n\}$, no qual $Y_{ij} = 1$ se os vértices (i, j) são conectados e $Y_{ij} = 0$ caso contrário.

2.3.1 Exponential random graph models

Exponential random graph models (ERGMs) são modelos da família exponencial que descrevem grafos aleatórios. Um ERGM é definido como:

$$P(Y = y|\theta) = \frac{\exp(\theta^T s(y))}{c(\theta)}, \quad \forall y \in \mathcal{Y}$$

onde θ é um vetor de parâmetros, $s(y)$ é um vetor de estatísticas suficientes e $c(\theta)$ é uma constante normalizadora. As estatísticas suficientes podem ser funções da rede, ou atributos dos vértices.

ERGMs são adequados para a modelagem de redes sociais (EXPONENTIAL... , 2012).

3 Implementação

O objetivo principal deste trabalho é implementar o conjunto de estados em um network model, utilizando a linguagem de programação R ([R Core Team, 2021](#)) e os pacotes `ergm` ([HANDCOCK et al., 2020](#)) para a simulação de grafos e `EpiModel` ([JENNESS; GOODREAU; MORRIS, 2018](#)) para o ajuste dos modelos via MCMC.

4 Simulação

Neste capítulo, primeiro sintetizaremos um grafo.

Depois, simularemos cenários de progressão da COVID-19 em nosso grafo. Fixaremos estimativas calculadas por estudos recentes como parâmetros de infecção e progressão da doença. Para os parâmetros de progressão para o estado *Quarantined*, bem como...

Compararemos inferências...

4.1 Síntese dos dados

O pacote `ergm` (HANDCOCK et al., 2020)... O pacote `EpiModel` (JENNESS; GOODREAU; MORRIS, 2018)...

4.1.1 Faux Magnolia High

O `ergm` contém diversos exemplos de grafo, entre eles o `faux.magnolia.high`. Este grafo é uma simulação de uma rede de amigos de uma escola do sul dos Estados Unidos. A figura 4 ilustra o grafo em questão.

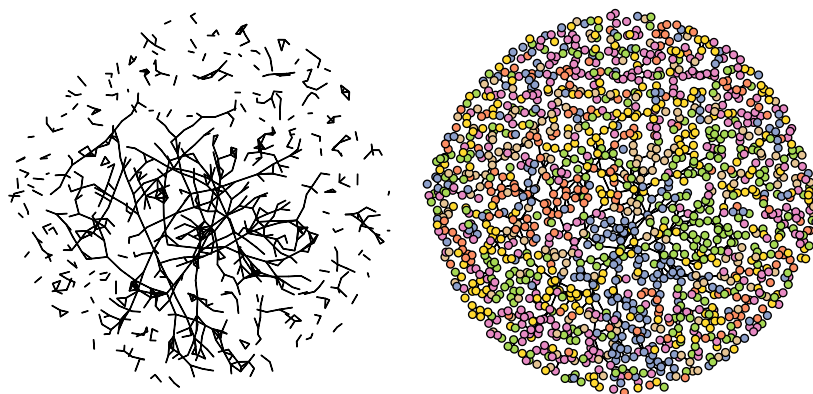


Figura 4 – Faux Magnolia High. À esquerda, apenas as arestas do grafo. À direita, os vértices coloridos de acordo com a turma do aluno.

Cada vértice do grafo representa um aluno, e as arestas as amizades entre alunos. Cada aluno tem três atributos: **Grade** (turma), **Race** (raça) e **Sex** (sexo).

A documentação do `ergm` contém detalhes sobre o modelo utilizado na simulação do `faux.magnolia.high`. O modelo em questão é um ERGM ajustado com o grafo `magnolia` original e o vetor de estatísticas suficientes $s(y)$ descrito na Tabela...

Tabela 1 – tabela de...

Estatística	Função	Atributo	Detalhes
Número de arestas	edges		
Homofilia diferencial	nodematch(attr, diff = T)	Grade, Race	
Homofilia uniforme	nodematch(attr, diff = F)	Sex	
Diferença absoluta	absdiff(attr)	Grade	
Geometrically weighted edgewise shared partner distribution	gwesp(0.25, fixed = T)		

Como nossa especificação do ERGM descreve perfeitamente o processo gerador dos dados, não é necessário que verifiquemos o ajuste do modelo de forma muito criteriosa.

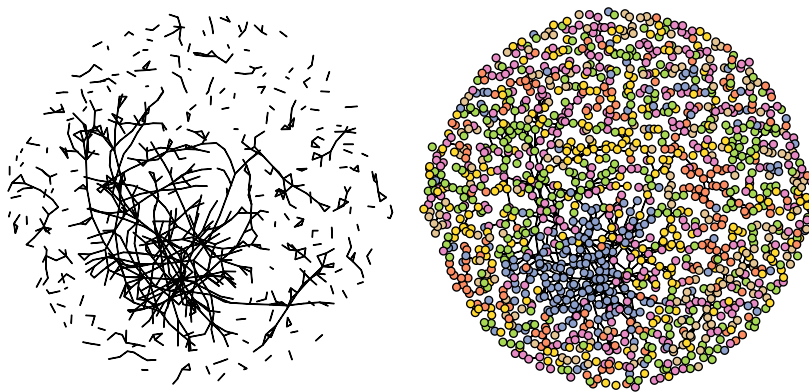


Figura 5 – Simulação gerada com nosso ajuste do Faux Magnolia High. À esquerda, apenas as arestas do grafo. À direita, os vértices coloridos de acordo com a turma do aluno.

4.1.2 Ajuste

O pacote `EpiModel` ([JENNESS; GOODREAU; MORRIS, 2018](#))...

Tabela 2 – Coeficientes do ajuste Magnolia.

	x
edges	-7.5328726
nodematch.Grade.7	1.3911537
nodematch.Grade.8	1.4275003
nodematch.Grade.9	0.9261669
nodematch.Grade.10	0.9713186
nodematch.Grade.11	1.2118756
nodematch.Grade.12	1.1530003
nodematch.Sex	0.7362824
absdiff.Grade	-0.9960848
gwesp.fixed.0.25	1.8757341

5 Resultados

TODO...

6 Conclusão

TODO...

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 28.

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 2 vezes nas páginas 25 e 28.

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 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 28.

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 33.

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 4 vezes nas páginas 28, 31, 33 e 34.

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 na página 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 25.

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.

Apêndices

APÊNDICE A – Código

```

library(knitr)

opts_chunk$set(
  echo = FALSE,
  fig.align = "center",
  out.width = "67%",
  out.height = "67%",
  message = FALSE,
  warning = FALSE,
  cache = TRUE
)

COLORS <- RColorBrewer::brewer.pal(8, "Set2")
USE_CACHE <- TRUE
library(DiagrammerR)

plot_states <- function(states, edge_df) {
  set.seed(0)
  set <- paste(states, collapse = "")
  path <- file.path("images", paste0(set, ".png"))
  create_graph() |>
    add_node(
      label = states,
      node_aes = node_aes(
        color = "black",
        fillcolor = COLORS[1],
        fontcolor = "black",
        height = 0.25,
        width = 0.25
      )
    ) |>
    add_edge_df(edge_df) |>
    render_graph(layout = "fr")
}

states <- c("S", "I", "R")
ids <- seq_along(states)
names(ids) <- states

```

```

edge_df <- create_edge_df(
  from = ids[c("S", "I", "I")],
  to = ids[c("I", "R", "S")],
  style = c("solid", "solid", "dashed"),
  color = "black"
)
plot_states(states, edge_df)
library(EpiModel)
# set.seed(1995)

# Example from: http://statnet.org/tut/BasicICMs.html

param <- param.dcm(inf.prob = 0.2, act.rate = 0.8, rec.rate = 1/50,
  a.rate = 1/100, ds.rate = 1/100, di.rate = 1/90,
  dr.rate = 1/100)
init <- init.dcm(s.num = 900, i.num = 100, r.num = 0)
control <- control.dcm(type = "SIR", nsteps = 300)
det <- dcm(param, init, control)

param <- param.icm(inf.prob = 0.2, act.rate = 0.8, rec.rate = 1/50,
  a.rate = 1/100, ds.rate = 1/100, di.rate = 1/90,
  dr.rate = 1/100)
init <- init.icm(s.num = 900, i.num = 100, r.num = 0)
control <- control.icm(type = "SIR", nsteps = 300, nsims = 10)
sim <- icm(param, init, control)

plot(det, alpha = 0.75, lwd = 4, main = "Compartimentos (%)", ylab = "", xlab = "Tempo")
plot(sim, qnts = FALSE, sim.lines = FALSE, add = TRUE, mean.lty = 2, legend = FALSE)
states <- c("S", "E", "I", "Q", "H", "R", "F")
ids <- seq_along(states)
names(ids) <- states
from <- c('E', 'E', 'I', 'I', 'I', 'Q', 'Q', 'H', 'H', 'S', 'E', 'I', 'Q')
to <- c('I', 'R', 'Q', 'H', 'R', 'H', 'R', 'R', 'F', 'E', 'S', 'S', 'S')
edge_df <- create_edge_df(
  from = ids[from],
  to = ids[to],
  style = rep(c("solid", "dashed"), c(10, 3)),
  color = "black"
)
plot_states(states, edge_df)
library(ergm)

plot_magnolia <- function(network, vertex_size = 0) {

```



```

set.seed(2)
plot(
  network,
  vertex.cex = vertex_size,
  vertex.col = COLORS[get.node.attr(faux.magnolia.high, "Grade") - 5]
)
}

data(faux.magnolia.high)
par(mar = rep(0, 4), mfrow = c(1, 2))
plot_magnolia(faux.magnolia.high, 0)
plot_magnolia(faux.magnolia.high, 1)
library(kableExtra)

statistics_table <- data.frame(
  Estatística = c(
    "Número de arestas",
    "Homofilia diferencial",
    "Homofilia uniforme",
    "Diferença absoluta",
    "Geometrically weighted edgewise shared partner distribution"
  ),
  Função = c(
    "edges",
    "nodematch(attr, diff = T)",
    "nodematch(attr, diff = F)",
    "absdiff(attr)",
    "gwesp(0.25, fixed = T)"
  ),
  Atributo = c("", "Grade, Race", "Sex", "Grade", ""),
  Detalhes = ""
)

kbl(statistics_table, caption = "tabela de...", booktabs = TRUE)
library(EpiModel)

magnolia_path <- file.path("models", "magnolia.rds")
if (USE_CACHE && file.exists(magnolia_path)) {
  magnolia <- readRDS(magnolia_path)
} else {
  formation <-
    ~
    edges +

```

```

    nodematch("Grade", diff = T) +
    # nodematch("Race", diff = T) +
    nodematch("Sex", diff = F) +
    absdiff("Grade") +
    gwesp(0.25, fixed = T)
# formation <- as.formula(paste(as.character(form)[-2], collapse = " "))
coef_diss <- dissolution_coefs(dissolution = ~ offset(edges), duration = 50)
coef_diss$coef.crude <- -Inf
magnolia <- netest(
  faux.magnolia.high,
  formation = formation,
  target.stats = NULL,
  coef.diss = coef_diss,
  set.control.ergm = control.ergm(
    MCMC.burnin = 1e5, #
    MCMC.interval = 1000, #
    MCMLE.maxit = 4000,
    MCMLE.steplength = 1 #
  )
)
saveRDS(magnolia, magnolia_path)
}

magnolia_sim <- simulate(
  magnolia$fit,
  nsim = 1,
  burnin = 100000000,
  constraint = "ConstantEdges"
)

par(mar = rep(0, 4), mfrow = c(1, 2))
plot_magnolia(magnolia_sim, 0)
plot_magnolia(magnolia_sim, 1)

coefs <- coef(magnolia$fit)
# strsplit(names(coefs), "\\.")
kbl(
  coef(magnolia$fit),
  caption = "Coeficientes do ajuste Magnolia.",
  booktabs = TRUE
)
# library(parallel)

```

```
#
# n_cores <- detectCores()
# eff_cores <- if (n_cores > 4) 4 else n_cores
# dx <- netdx(
#   est,
#   nsims = eff_cores,
#   nsteps = 365,
#   # set.control.ergm = control.simulate.ergm(
#   #   MCMC.burnin = 1e5,
#   #   MCMC.interval = 1000,
#   #   MCMLE.maxit = 1000,
#   #   MCMLE.steplength = 1
#   # ),
#   ncores = eff_cores
# )
# plot(dx, plots.joined = FALSE, qnts.alpha = 0.8)
```