



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTATÍSTICA

LUCIANO SOARES MENDES JÚNIOR

**ESTUDO DA CORRELAÇÃO ENTRE VIZINHOS ADJACENTES EM MODELOS
DE TRANSMISSÃO DEFINIDOS EM GRAFOS FINITOS**

Recife

2022

LUCIANO SOARES MENDES JÚNIOR

**ESTUDO DA CORRELAÇÃO ENTRE VIZINHOS ADJACENTES EM MODELOS
DE TRANSMISSÃO DEFINIDOS EM GRAFOS FINITOS**

Dissertação apresentada ao Programa de Pós-Graduação em Estatística do Centro de Ciências Exatas e da Natureza da Universidade Federal de Pernambuco, como requisito parcial à obtenção do título de mestre em Estatística.

Área de Concentração: Probabilidade

Orientador (a): Prof. Dr. Pablo Martín Rodríguez

Coorientador (a): Dr. Guilherme Ferraz de Arruda

Recife

2022

Catálogo na fonte
Bibliotecária Monick Raquel Silvestre da S. Portes, CRB4-1217

M538e Mendes Júnior, Luciano Soares
Estudo da correlação entre vizinhos adjacentes em modelos de transmissão definidos em grafos finitos / Luciano Soares Mendes Júnior. – 2022.
87 f.: il., fig.

Orientador: Pablo Martín Rodríguez.
Dissertação (Mestrado) – Universidade Federal de Pernambuco. CCEN, Estatística, Recife, 2022.
Inclui referências e apêndices.

1. Probabilidade. 2. Processo estocástico. 3. Correlações. I. Martín Rodríguez, Pablo (orientador). II. Título.

519.2 CDD (23. ed.) UFPE - CCEN 2023-14

LUCIANO SOARES MENDES JUNIOR

**"Estudo da correlação entre vizinhos adjacentes em
modelos de transmissão definidos em grafos finitos"**

Dissertação apresentada ao Programa de Pós-Graduação em Estatística da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Mestre em Estatística.

Aprovada em 29 de agosto de 2022.

BANCA EXAMINADORA

PROFA. ANGÉLICA YOHANA PACHÓN PINZÓN
University of South Wales

PROF. LEON ALEXANDER VALENCIA HENAO
Universidad de Antioquia

PROF. ÉLCIO LEBENSZTAYN
Universidade Estadual de Campinas - UNICAMP

PROF. PABLO MARTIN RODRIGUEZ
Universidade Federal de Pernambuco - UFPE

AGRADECIMENTOS

Primeiramente, agradeço a Deus, que me deu o dom da vida e forças para superar cada obstáculo.

Aos meus pais, Maria das Dores e Luciano Soares, que sempre me incentivaram e batalharam por mim. E, também minha irmã, Júnia Talita, que sempre acreditou em mim, até quando eu mesmo duvidei.

À minha esposa, Bianca Freitas, pela compreensão e paciência nos momentos em que precisei ficar ausente, pelo amor, carinho e cuidado.

Aos amigos que estiveram presentes ao longo desta jornada, e que, direta ou indiretamente, contribuíram positivamente para a conclusão dela, em especial, César Diogo, pelo incentivo a entrar no PPGE, Ithallo Pontes e Guilherme Bernoti, pela solidariedade em fornecer seus computadores para que eu pudesse executar minhas simulações.

Aos Professores Dr. Élcio Lebensztayn, Dr^a. Angélica Yohana Pachón Pinzón e Dr. León Alexander Valencia Henao, que se dispuseram a compor a Banca Examinadora desta dissertação.

Ao meu orientador, Prof. Dr. Pablo Martín Rodríguez, pela paciência, disposição em ajudar e pela oportunidade de desenvolver esta pesquisa. Também ao meu coorientador, Dr. Guilherme Ferraz de Arruda, por todas as conversas e conhecimentos transmitidos sobre a parte computacional.

Aos professores do Programa de Pós-Graduação em Estatística, pelos ensinamentos.

À FACEPE, pelo apoio financeiro.

RESUMO

Este trabalho é um estudo teórico e numérico sobre o comportamento das correlações envolvidas nos modelos SIS, SIR e MT. Inicialmente, foram desenvolvidos estudos de argumentos apresentados recentemente na literatura. Essa etapa envolveu a aplicação da desigualdade FKG para mostrar que as variáveis de interesse têm correlação não-negativa. Posteriormente, foi analisado um outro argumento, em que é exibido um contraexemplo mostrando que os argumentos estudados anteriormente não são válidos para o modelo SIR. Então, foram desenvolvidos métodos de simulação de Monte Carlo para verificar o comportamento das correlações nos modelos mencionados. O primeiro método desenvolvido estima a correlação média ao longo do processo em um par de vértices vizinhos; o segundo método, por sua vez, estima a correlação média em cada unidade de tempo do processo. Os métodos desenvolvidos foram aplicados em dois tipos de grafos: um grafo k -regular e um outro grafo formado por duas estrelas com n vértices conectadas pelos seus respectivos vértices centrais. Verificamos que, para os modelos SIR e MT definidos no grafo k -regular, as correlações apresentam um comportamento não-negativo. Por outro lado, esse comportamento não se repete para os mesmos modelos definidos no grafo formado por duas estrelas ligadas pelos seus respectivos vértices centrais.

Palavras-chave: modelo epidêmico; modelo de rumor; processo estocástico; correlações; transmissão de informação.

ABSTRACT

This paper is a theoretical and numerical study of the correlation behavior involved in the SIS, SIR, and MT models. Initially, were developed studies about recent arguments in the literature. In this step was applied the FKG inequality to show that the variables of interest have a non-negative correlation. Subsequently, another argument was analyzed with a counterexample and consequently showed that the arguments previously studied are not valid for the SIR model. Then, Monte Carlo simulation methods were developed to verify the behavior of the correlations in the mentioned models. The first method developed estimates the average correlation along the process in a pair of neighboring vertices. The second method, in turn, estimates the average correlation in each unit of process time. The developed methods were applied in two types of graphs: a k -regular graph and another graph formed by two stars with n vertices connected by their respective central vertices. We verified that in the SIR and MT models defined in the k -regular graph, the correlations present a non-negative behavior. On the other hand, this behavior is not repeated by the same models defined in the graph formed by two stars connected by their respective central vertices.

Keywords: epidemic model; rumor model; stochastic process; correlations; information transmission.

LISTA DE FIGURAS

Figura 1 – Representação gráfica de diversos tipos de grafos.	15
Figura 2 – Grafo completo com $n = 10$ vértices e sua respectiva matriz de adjacências.	16
Figura 3 – Grafo estrela com $n = 10$ vértices e sua respectiva matriz de adjacências.	17
Figura 4 – Grafo completo com $n = 10$ vértices e sua respectiva lista de adjacências.	17
Figura 5 – Grafo estrela com $n = 10$ vértices e sua respectiva lista de adjacências.	18
Figura 6 – Representação visual dos grafos escolhidos nos estudos deste trabalho.	18
Figura 7 – Modelo compartimental SIS.	19
Figura 8 – Modelo compartimental SIR.	20
Figura 9 – Modelo compartimental MT.	22
Figura 10 – Representação gráfica de Harris.	25
Figura 11 – Representação visual da função Z	25
Figura 12 – Representação visual da coleção de todos os instantes de tempo que ocorreram eventos em $[0, t]$	26
Figura 13 – Representação visual da discretização do intervalo $[0, t]$	27
Figura 14 – Representação visual de uma realização do processo e a discretização de $[0, t]$	29
Figura 15 – Realização do processo a partir de $Z^{(1)}(\cdot)$ e $Z^{(2)}(\cdot)$	31
Figura 16 – Comparação entre duas realizações do processo SIR de acordo com $Z^{(1)}(\cdot)$ e $Z^{(2)}(\cdot)$ de modo que $Z^{(1)} \leq Z^{(2)}$	36
Figura 17 – Comportamento das marcas favorável ao evento $\{X_1(1) = 1, X_2(1) = 0\}$	37
Figura 18 – Comportamento das marcas favorável ao evento $\{X_1(1) = 1, X_2(1) = 1\}$	38
Figura 19 – Comportamento das marcas favorável ao evento $\{X_1(1) = 1, X_2(1) = 2\}$	39
Figura 20 – Comportamento das marcas favorável ao evento $\{X_1(1) = 2, X_2(1) = 0\}$	40
Figura 21 – Comportamento das marcas favorável ao evento $\{X_1(1) = 2, X_2(1) = 1\}$	41
Figura 22 – Simulações para o modelo SIS no grafo completo de 500 vértices.	47
Figura 23 – Simulações para o modelo SIR no grafo completo de 500 vértices.	49
Figura 24 – Simulações para o modelo MT no grafo completo e k -regular.	51
Figura 25 – Gráfico Q-Q das densidades finais de indivíduos ignorantes do modelo Maki-Thompson.	53

Figura 26 – Comparação entre os métodos discretos e contínuos na aferição das covariâncias no modelo SIR.	63
Figura 27 – Resultados das estimativas dos tempos de extinção do processo SIR e MT.	64
Figura 28 – Estimativas das covariâncias e correlações médias entre os vértices 0 e 1 no grafo k-regular para o modelo SIR com $\alpha = 1.0$ e $\lambda = 20/k$	65
Figura 29 – Estimativas das covariâncias e correlações médias entre os vértices 0 e 1 no grafo estrela $n \times n$ para o modelo SIR com $\delta = 1$ e $\lambda = 20.0/n$	65
Figura 30 – Estimativas das covariâncias e correlações em função do tempo entre os vértices 0 e 1 no grafo k-regular com $k = 30$, $\delta = 1$, $\lambda = 20/30$ para o modelo SIR.	66
Figura 31 – Estimativas das covariâncias e correlações em função do tempo entre os vértices 0 e 1 no grafo k-regular com $k = 30$, $\delta = 1$, $\lambda = 20/610$ para o modelo SIR.	66
Figura 32 – Estimativas das covariâncias e correlações médias entre os vértices 0 e 1 no grafo k-regular para o modelo MT com $\alpha = \lambda = 1.0$	68
Figura 33 – Estimativas das covariâncias e correlações médias entre os vértices 0 e 1 no grafo estrela $n \times n$ para o modelo MT com $\alpha = \lambda = 1$	68
Figura 34 – Estimativas das covariâncias e correlações em função do tempo entre os vértices 0 e 1 no grafo k-regular com $k = 30$, $\delta = \lambda = 1,0$ para o modelo MT.	69
Figura 35 – Estimativas das covariâncias e correlações médias entre os vértices 0 e 1 no grafo estrela 610×610 com $\alpha = \lambda = 1,0$ para o modelo MT.	69
Figura 36 – Estados do vértice do grafo G em t_0	75
Figura 37 – Representação visual de uma realização do processo no tempo t	75
Figura 38 – Representação visual de uma realização do processo no tempo t	80
Figura 39 – Representação visual de uma realização do processo no tempo t	80

LISTA DE SÍMBOLOS

$\hat{\mu}_{X^i}$	Estimador do primeiro momento amostral da variável X^i
$\hat{\mu}''_{X^i}$	Estimador do segundo momento amostral da variável X^i
$\hat{\mu}_{X^i X^j}$	Estimador do primeiro momento amostral de $X^i X^j$
$\hat{\sigma}_{X^i}^2$	Estimador da variância amostral da variável X^i
$\hat{\sigma}_{X^i X^j}$	Estimador da covariância amostral entre as variáveis X^i e X^j
$\hat{\rho}_{X^i X^j}$	Estimador da correlação amostral entre as variáveis X^i e X^j

SUMÁRIO

1	INTRODUÇÃO	12
1.1	GRAFOS	14
1.2	MODELO SIS EM GRAFOS	19
1.3	MODELO SIR EM GRAFOS	20
1.4	MODELO DE MAKI-THOMPSON EM GRAFOS	21
1.5	CORRELAÇÕES EM MODELOS EPIDÊMICOS	22
1.6	ORGANIZAÇÃO DO TRABALHO	23
2	ANÁLISE TEÓRICA	24
2.1	CONSTRUÇÃO GRÁFICA DE CATOR-VAN MIEGHEM PARA O MO- DELO SIS	24
2.2	DISCRETIZAÇÃO DO MODELO SIS	27
2.2.1	Monotonicidade da função $Z_n(\cdot)$	31
2.2.2	Monotonicidade dos estados	32
2.2.3	Correlação entre vizinhos infectados	34
2.3	A NÃO MONOTONICIDADE DO PROCESSO SIR	35
3	SIMULAÇÃO DE PROCESSOS DE TRANSMISSÃO	46
3.1	ALGORITMOS	46
3.1.1	Algoritmo para simulação do processo SIS	46
3.1.2	Algoritmo para simulação do processo SIR	48
3.1.3	Algoritmo para simulação do processo de Maki-Thompson	50
<i>3.1.3.1</i>	<i>Validação do Algoritmo para simulação do Modelo Maki-Thompson</i>	<i>52</i>
3.2	CÁLCULO DAS CORRELAÇÕES E COVARIÂNCIAS	54
3.2.1	Métodos Discretos	54
3.2.2	Cálculo da Correlação e Covariância: Modelo SIS	55
3.2.3	Cálculo da Correlação e Covariância: Modelo SIR	56
3.2.4	Cálculo da Correlação e Covariância: Modelo MT	57
3.2.5	Método Instantâneo	58
3.2.6	Método Instantâneo — Cálculo da Covariância: Modelo SIS	59
3.2.7	Método Instantâneo — Cálculo da Covariância: Modelo SIR	60
3.2.8	Método Instantâneo — Cálculo da Covariância: Modelo MT	61

3.2.9	Comparação Entre os Métodos Discretos e Instantâneos	62
4	RESULTADOS E DISCUSSÕES	64
4.1	CORRELAÇÕES E COVARIÂNCIAS	65
4.1.1	Modelo SIR	65
4.1.2	Modelo MT	68
5	CONCLUSÕES	71
	REFERÊNCIAS	73
	APÊNDICE A – DEDUÇÕES DAS EXPRESSÕES 3.1 E 3.3	75
	APÊNDICE B – DEDUÇÕES DAS EXPRESSÕES 3.7 E 3.10	80
	APÊNDICE C – CORRIDA DAS EXPONENCIAIS	83
	APÊNDICE D – CÓDIGO FONTE	84

1 INTRODUÇÃO

A epidemiologia, no seu sentido etimológico, significa o estudo do que acontece em uma população. A palavra epidemiologia é derivada do grego, *epi* = sobre, *demos* = população ou povo e *logos* = estudo (FARIAS, 2017). A epidemiologia matemática é uma área da Biomatemática que tem como objetivo descrever de forma quantitativa e qualitativa o comportamento das epidemias usando modelos matemáticos. Seu principal interesse é descrever o alastramento do agente infeccioso na população e fornecer informações sobre os parâmetros envolvidos para que haja possibilidades de previsão e controle do avanço da doença (JUNIOR, 2021).

É importante definir, antes de entrarmos na formulação matemática, alguns conceitos básicos que desempenham um papel relevante no entendimento dos modelos que serão tratados ao longo do texto.

- **Epidemia:** É o acontecimento de uma doença que se alastra em uma população, ultrapassando a expectativa considerada normal;
- **Transmissão:** Definida como o processo em que um patógeno é disseminado entre os indivíduos da população;
- **Indivíduos suscetíveis (S):** São aqueles que estão saudáveis ou que estão vulneráveis a contrair a doença por meio de algum tipo de contato potencialmente infeccioso. É assumido que todos os sujeitos suscetíveis estão possibilitados a contrair a doença;
- **Indivíduos infectados (I):** São aqueles que já contraíram a doença. Esses sujeitos são o meio de propagação do patógeno;
- **Indivíduos recuperados ou resistentes (R):** São caracterizados pelos sujeitos que já foram infectados e, após um período foram, totalmente curados da doença. Os sujeitos que se encontram neste estado estão imunes a doença;
- **Taxa de infecção:** Medida da proporção de indivíduos que são infectados por unidade de tempo;
- **Taxa de cura:** Medida da proporção de indivíduos que passam do estado infectado para suscetível ou recuperado por unidade de tempo.

Em 1927, com o intuito de estudar a propagação de uma doença infecciosa, Kermack e McKendrick propuseram o modelo SIR, que é considerado um dos modelos de maior importância, com uma influência notável no desenvolvimento de outros modelos epidêmicos (JUNIOR, 2021). A partir desse estudo, Kermack e McKendrick concluíram que um número pequeno em relação ao tamanho total da população de indivíduos infectados, mesmo em contato com os suscetíveis, não causavam uma epidemia.

Com base nos estudos de Kermack e McKendrick, inúmeras variações do modelo SIR foram desenvolvidas e estudadas, dando origem aos chamados modelos compartimentais, que foram nomeados dessa forma devido ao fato da população ser separada em compartimentos ou classes que denotam em qual estado se encontra um determinado grupo de indivíduos (FARIAS, 2017). Existem diversos outros modelos compartimentais. Caso o leitor tenha curiosidade, poderá consultar as referências (JUNIOR, 2021; EMERITO, 1989).

Os modelos epidemiológicos podem ser classificados quanto a variável tempo como contínuos ou discretos. Em um modelo a tempo discreto, as mudanças ocorrem em intervalos constantes e bem definidos. Já a tempo contínuo, essas mudanças ocorrem em intervalos variados. Em termos dos estados, também podemos classificar os modelos como estocásticos ou determinísticos. Um modelo determinístico é aquele em que o conjunto de estados das variáveis de interesse é totalmente caracterizado por um sistema de equações diferenciais e suas condições iniciais. O estocástico é dado por um processo estocástico.

Um exemplo de modelo estocástico simples para propagação de epidemias pode ser construído considerando uma população de tamanho fixo $N + 1$, em que a propagação da doença pode ser modelada por meio de um processo *Markoviano* a tempo contínuo (GRIMMETT, 2018). Seja $X(t)$ a variável aleatória que conta o número de indivíduos no estado suscetível (saudáveis) no instante de tempo t , vamos supor que $X(0) = N$ e assumir que: se $X(t) = n$, a probabilidade de ocorrer uma nova infecção durante um intervalo muito pequeno de tempo $[t, t + h]$ é proporcional à quantidade de encontros possíveis entre indivíduos doentes e saudáveis, em que a infecção é transmitida a uma taxa λ , que é uma constante arbitraria. Então temos que a probabilidade de $X(t + h) = n - 1$ dado $X(t) = n$ é:

$$\mathbb{P}(X(t + h) = n - 1 | X(t) = n) = \lambda n(N + 1 - n)h + o(h), \text{ para } h \rightarrow 0. \quad (1.1)$$

Note que esse modelo é simplista, pois consideramos que temos apenas dois estados possíveis, nos quais não há recuperação total do indivíduo. Também foi considerada uma população

totalmente misturada e homogênea, em que a taxa de infecção é proporcional à quantidade de encontros possíveis.

Contudo, uma pessoa pode infectar apenas as outras que são parte de sua vizinhança imediata, isto é, existe uma relação espacial no alastramento da doença. A incorporação das relações espaciais em tal modelo nos leva ao *Processo de Contato de Harris* (HARRIS, 1977a). Nesse modelo a população é representada por um grafo, em que os indivíduos são o conjunto de vértices e as arestas indicam seus respectivos vizinhos.

1.1 GRAFOS

Como visto na seção anterior, ao introduzir as relações espaciais em modelos epidêmicos, temos a necessidade de representar a população através de um grafo. Grafos são objetos matemáticos que representam relações binárias entre elementos de um conjunto finito de interesse. Em geral, um grafo é formado por um conjunto de vértices que podem estar ligados dois a dois por arestas. Dizemos que dois vértices são vizinhos se os mesmos são ligados por uma aresta.

Definição: Seja $G = (V, A)$ um grafo, em que V é um conjunto de vértices e A é um conjunto de arestas.

Definição: Seja k_i o grau do vértice i , como sendo o número de arestas ligadas a i .

Lema 1. Para todo grafo $G = (V, A)$,

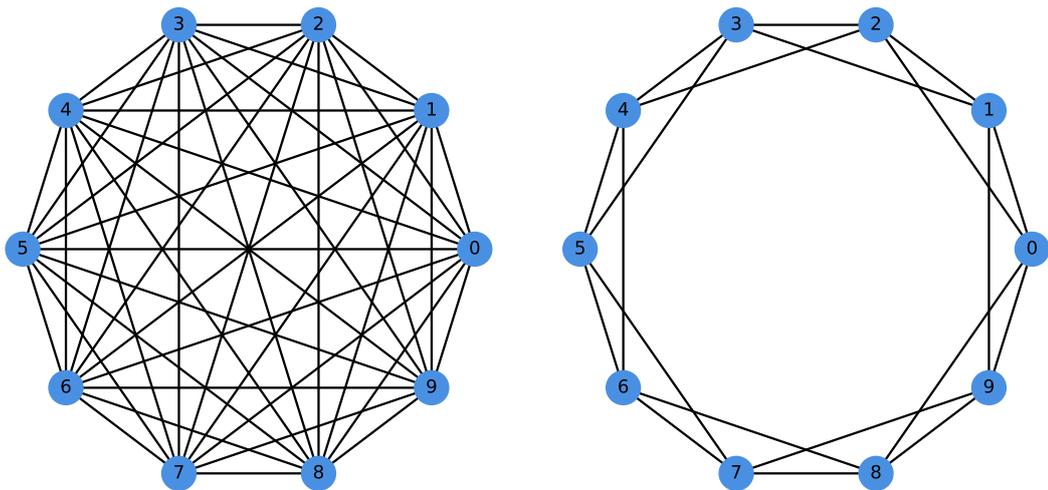
$$\sum_{i \in V} k_i = 2m. \quad (1.2)$$

Isto é, a soma de todos os graus dos vértices de um grafo sempre é o dobro do número de arestas (JURKIEWICZ, 2009).

Existem grafos com diversas topologias, podemos citar alguns deles. O *grafo completo*, é definido como sendo um grafo em que todos os vértices são vizinhos entre si. O grafo completo de n vértices é usualmente denotado por K_n , e o número de arestas é dado por $n(n-1)/2$, correspondendo a todos os possíveis pares de vértices. Temos também o *grafo k -regular*, que é definido como sendo um grafo em que seus vértices têm o mesmo número de vizinhos, isto é, cada vértice tem o mesmo grau. Um outro exemplo de grafo é o do tipo *estrela*, que usualmente é representado por S_k . Em um *grafo estrela*, apenas um único vértice tem grau maior que um.

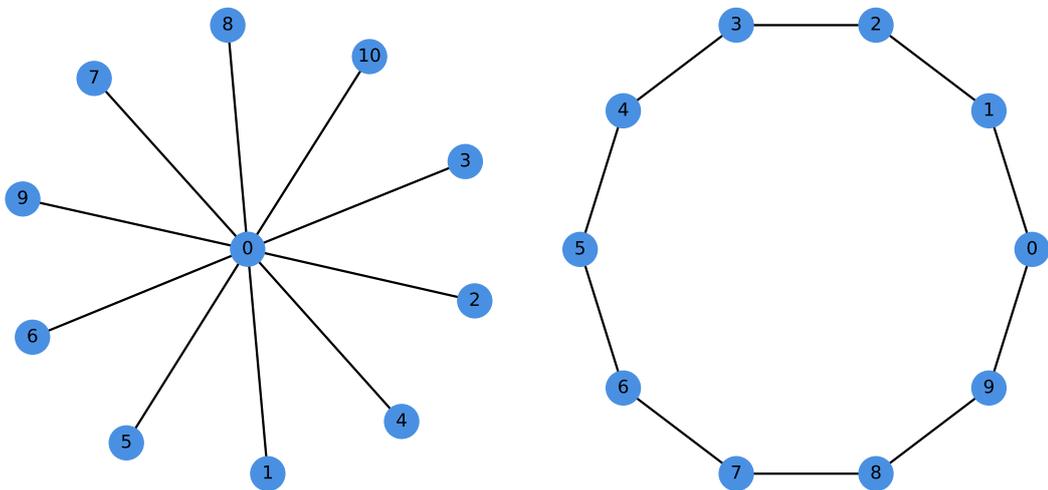
Temos também o *grafo circular*, que consiste em um anel ou, em outras palavras, vértices conectados a uma rede fechada. Um *grafo circular* com N vértices é usualmente representado por C_n . O número de vértices em C_n é igual a quantidade de arestas e cada vértice tem grau 2. Para mais informações, o leitor poderá consultar as referências (JURKIEWICZ, 2009; NETTO; JURKIEWICZ, 2017). Na Figura 1, podemos encontrar imagens dos grafos citados.

Figura 1 – Representação gráfica de diversos tipos de grafos.



(a) Grafo completo com $n = 10$ vértices.

(b) Grafo k -regular com $n = 10$ vértices e $k = 4$.



(c) Grafo estrela com $n = 10$ vértices.

(d) Grafo circular com $n = 10$ vértices.

Fonte: Produzido pelo autor.

Para o *grafo k -regular*, um resultado que mais a frente será de nosso interesse é que a quantidade total de arestas é $\frac{Nk}{2}$, sendo N o número total de vértices. Esse resultado pode ser demonstrado sabendo que todos os vértices possuem o mesmo grau, portanto, pelo Lema

1, o número de arestas de um *grafo k-regular* é dado por:

$$\begin{aligned} \sum_{i \in V} k_i &= 2m, \\ 2m &= Nk, \\ m &= \frac{Nk}{2}. \end{aligned} \quad (1.3)$$

Se tratando da representação computacional de um grafo, a forma mais comum se dá por meio de matrizes. A representação usando matriz é feita utilizando um tipo especial de matriz chamada de matriz de *adjacências*.

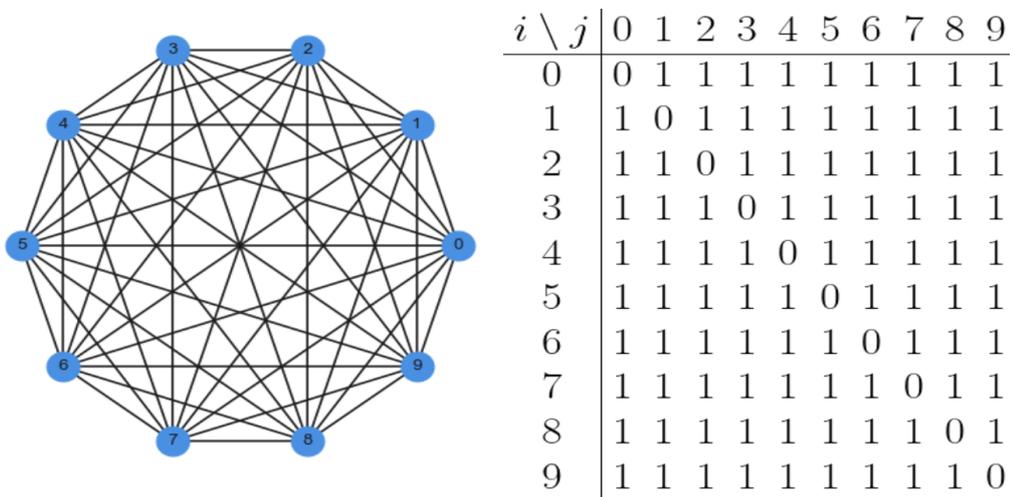
Definição: Seja A uma matriz quadrada $n \times n$ definida como:

$$a_{ij} = \begin{cases} 1, & \text{se } ij \text{ vizinhos;} \\ 0 & \text{caso contrário.} \end{cases} \quad (1.4)$$

Para i e $j \in V$, dizemos que A é uma matriz de adjacências.

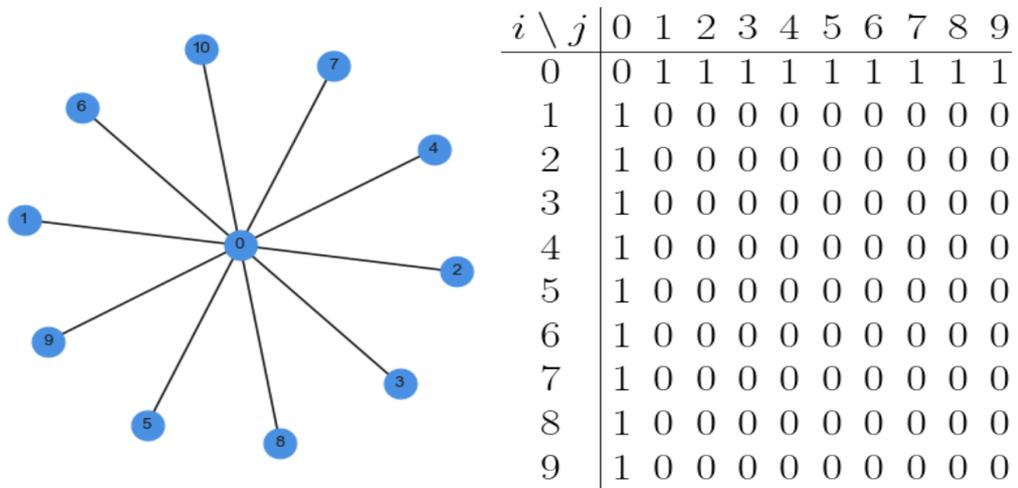
Nas Figuras 2 e 3, apresentamos a matriz de adjacência do *grafo completo* e *estrela*, respectivamente.

Figura 2 – Grafo completo com $n = 10$ vértices e sua respectiva matriz de adjacências.



Fonte: Produzido pelo autor.

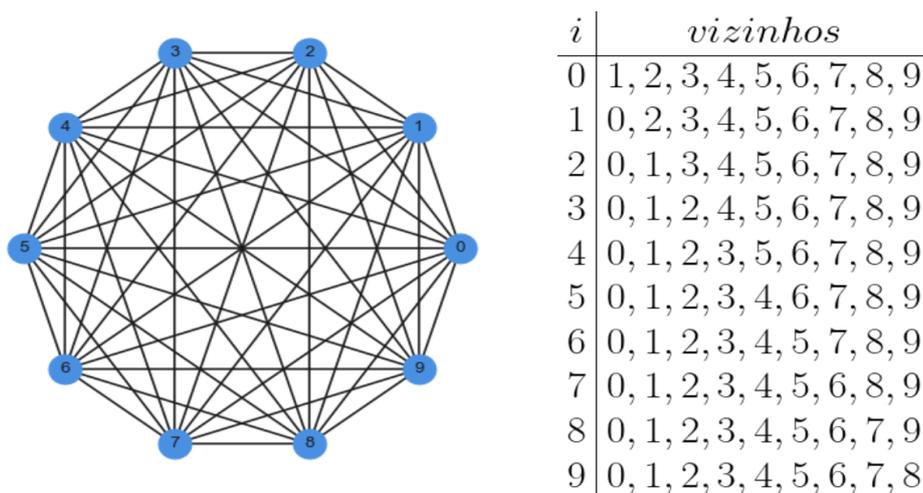
Figura 3 – Grafo estrela com $n = 10$ vértices e sua respectiva matriz de adjacências.



Fonte: Produzido pelo autor.

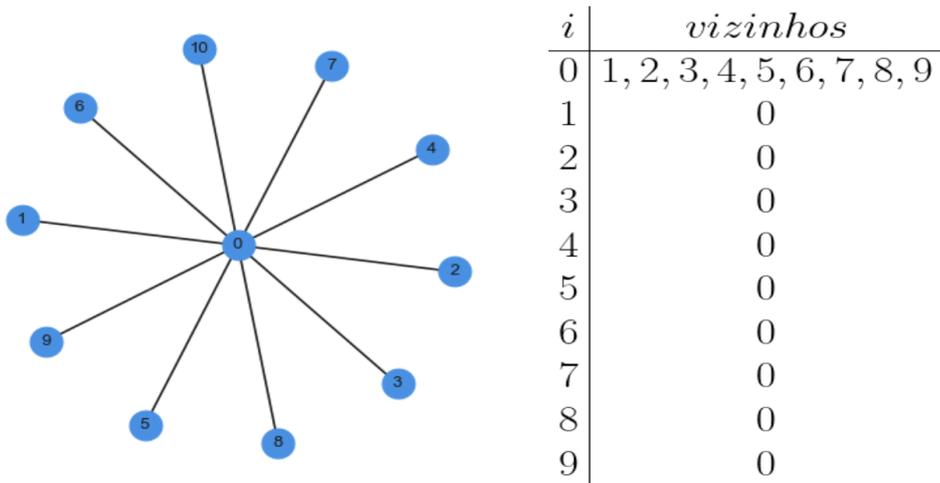
Uma outra forma de representação de grafos é por meio da lista de adjacências. A lista de adjacências é uma estrutura de dados em que podemos manter, para cada vértice do grafo, uma lista contendo todos os outros vértices com os quais se compartilha uma aresta. Esse tipo de representação demonstra ser mais eficiente em termos de memória quando comparada à representação por matriz de adjacências. Nas Figuras 4 e 5 apresentamos as listas de adjacências para o grafo completo e estrela, respectivamente. Note que a quantidade de elementos é bem menor em comparação com as matrizes apresentadas nas Figuras 2 e 3.

Figura 4 – Grafo completo com $n = 10$ vértices e sua respectiva lista de adjacências.



Fonte: Produzido pelo autor.

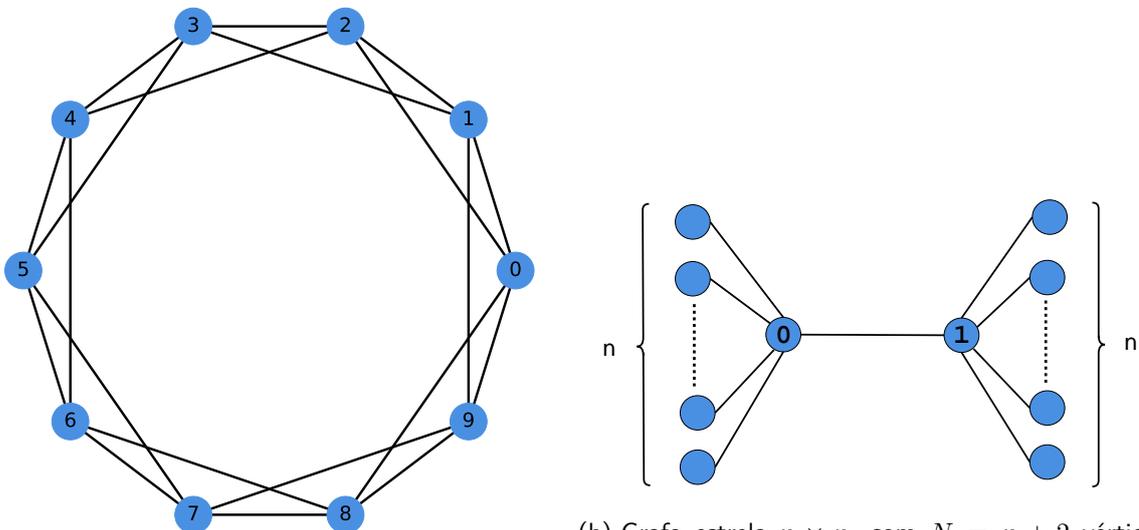
Figura 5 – Grafo estrela com $n = 10$ vértices e sua respectiva lista de adjacências.



Fonte: Produzido pelo autor.

Para os estudos desenvolvidos neste trabalho, foram selecionadas a família dos *grafos k -regulares* e um outro tipo de grafo formado por dois grafos do tipo *estrela* de tamanho n , conectados pelos seus respectivos vértices centrais, para esse grafo, demos o nome de *grafo estrela $n \times n$* . Veja, na Figura 6, a representação visual dos grafos mencionados.

Figura 6 – Representação visual dos grafos escolhidos nos estudos deste trabalho.



(a) Grafo k -regular, com $N = 10$ vértices e $k = 4$.

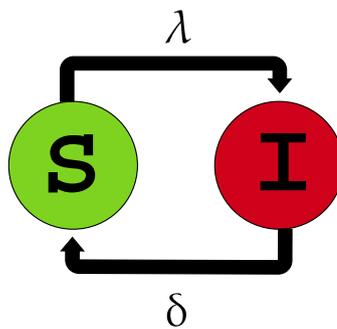
(b) Grafo estrela $n \times n$, com $N = n + 2$ vértices, sendo n a quantidade de vizinhos dos vértices 0 e 1 respectivamente.

Fonte: Produzido pelo autor.

1.2 MODELO SIS EM GRAFOS

No modelo SIS, os indivíduos da população podem se encontrar em dois estados: suscetível ou infectado. Esse modelo é específico para doenças das quais os infectados não desenvolvem imunidade. Os indivíduos que estão doentes, ao se recuperarem, voltam para a classe dos suscetíveis, podendo, assim, contrair novamente o patógeno. Por exemplo, podemos citar a meningite, que possui esse tipo de comportamento (JUNIOR, 2021). Para tornar o modelo mais simples, as taxas de natalidade e mortalidade podem ser desconsideradas se elas não representarem uma variação notável no tamanho da população. Nesse caso, dizemos que o modelo é sem dinâmica vital. A Figura 7 apresenta a diagramação do modelo SIS.

Figura 7 – Modelo compartimental SIS.



Fonte: Produzido pelo autor.

Para introduzir as relações espaciais no modelo, vamos considerar que a população seja representada por um grafo. Seja $G = (V, A)$ um grafo qualquer contendo N vértices, tal que a quantidade de conexões para cada um dos vértices seja finita. Suponha que tal modelo seja um processo *Markoviano* a tempo contínuo e que seu espaço de estados seja $\mathcal{S} : V \mapsto \{0, 1\}^N$. Podemos computar os estados dos vértices por meio de um vetor *zero-um*, $\mathbf{X} = \{X_i : i \in V\}$, em que, para cada vértice, 0 representa o estado suscetível e 1 representa o estado infectado. Fixam-se taxas de cura, δ , para todo $i \in V$, e de infecção, λ , para todo $(i, j) \in A$. Supondo que os estados de todos os vértices no tempo t sejam $\mathbf{X} \in \mathcal{S}$, então as probabilidades de transição para o vértice i são escritas na forma:

$$\mathbb{P}(X_i(t+h) = 0 | \mathbf{X}(t) = \mathbf{X}) = \delta h + o(h), \quad \text{se } X_i(t) = 1, \quad (1.5)$$

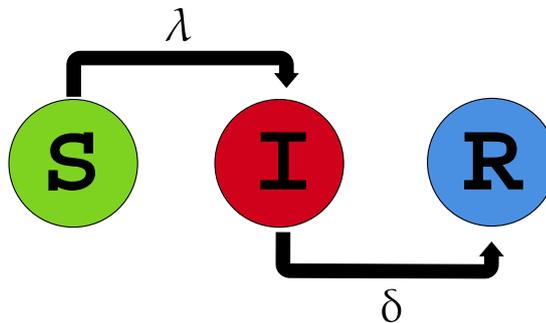
$$\mathbb{P}(X_i(t+h) = 1 | \mathbf{X}(t) = \mathbf{X}) = \lambda N_{\mathbf{X}}(i)h + o(h), \quad \text{se } X_i(t) = 0, \quad (1.6)$$

em que $N_{\mathbf{x}}(i) = |\{j \in V : j \sim i, X_j = 1\}|$ é o número de vizinhos de i que estão infectados. Portanto, cada vértice infectado vai para o estado suscetível a uma taxa δ e infecta qualquer um de seus vizinhos a uma taxa λ .

1.3 MODELO SIR EM GRAFOS

O modelo SIR surgiu em 1927 com Kermack e McKendrick (JUNIOR, 2021; PROCÓPIO-PR, 2020). Como já foi mencionado, esse modelo foi um dos primeiros modelos matemáticos a serem introduzidos no estudo do alastramento de doenças contagiosas, dando origem aos chamados modelos compartimentais. Ao contrário do modelo SIS, o indivíduo adquire imunidade após seu período de infecção. Nesse modelo também são consideradas as mortes pela doença, de modo que indivíduos que morreram contam como recuperados. A Figura 8 apresenta a diagramação do modelo SIR.

Figura 8 – Modelo compartimental SIR.



Fonte: Produzido pelo autor.

Um exemplo típico do modelo SIR é a *catapora* (JUNIOR, 2021), em que o indivíduo a partir do contato com outro infectado contrai a doença, mas, após um certo tempo, esse mesmo indivíduo se recupera, passando a ter imunidade. Desse modo, o mesmo indivíduo não volta a ser infectado ou suscetível.

Seguindo a mesma ideia da Seção 1.2, vamos introduzir as relações espaciais das transmissões considerando que nossa população seja representada por um grafo, em que os indivíduos habitam nos vértices do grafo. Seja $G = (V, A)$ um grafo qualquer contendo N vértices com um número finito de conexões em cada vértice. Também vamos supor que o modelo seja um processo *Markoviano* a tempo contínuo e que seu espaço de estados seja $\mathcal{S} : V \mapsto \{0, 1, 2\}^N$. Os estados de cada vértice são codificados por um vetor *zero-um-dois* $\mathbf{X} = \{X_i : i \in V\}$, em

que, para cada vértice, 0 representa o estado suscetível, 1 representa o infectado e 2 o recuperado. As taxas de cura e infecção são, δ , para todo $i \in V$, e λ , para todo $(i, j) \in A$, respectivamente. Suponha que os estados de todos os vértices no tempo t sejam $\mathbf{X} \in \mathcal{S}$. Portanto, as probabilidades de transição para o vértice i são dadas por:

$$\mathbb{P}(X_i(t+h) = 2 | \mathbf{X}(t) = \mathbf{X}) = \delta h + o(h), \quad \text{se } X_i(t) = 1, \quad (1.7)$$

$$\mathbb{P}(X_i(t+h) = 1 | \mathbf{X}(t) = \mathbf{X}) = \lambda N_{\mathbf{X}}(i)h + o(h), \quad \text{se } X_i(t) = 0, \quad (1.8)$$

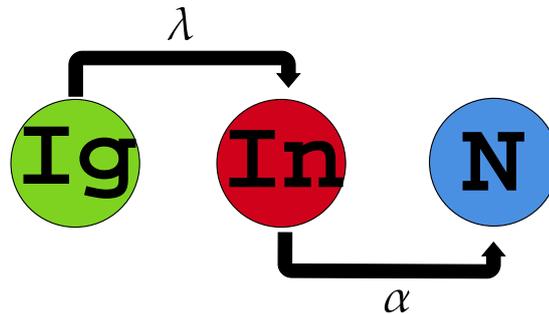
em que $N_{\mathbf{X}}(i) = |\{j \in V : j \sim i, X_j = 1\}|$ é o número de vizinhos de i que estão infectados. Portanto, cada vértice infectado vai para o estado recuperado a uma taxa δ , e infecta qualquer um de seus vizinhos a uma taxa λ .

1.4 MODELO DE MAKI-THOMPSON EM GRAFOS

Além das aplicações em epidemias, os modelos compartimentais podem ser aplicados na modelagem de difusão de informação em uma população. Em 1973, *Maki e Thompson* introduziram um novo modelo para propagação de rumor (MAKI; THOMPSON, 1973). Podemos notar que existem semelhanças entre as dinâmicas do espalhamento de um patógeno em uma população e a disseminação de uma informação na mesma população, como, por exemplo, a necessidade de que haja interação entre os indivíduos do grafo para que o espalhamento do rumor ocorra, do mesmo modo que se faz necessário o contato entre indivíduos para que uma doença se espalhe. Como vimos, em um surto de catapora, a população é dividida em três classes: os suscetíveis, os infectados e os resistentes. Para esse caso, o modelo SIR, apresentado na Seção 1.3 com as devidas considerações, é o mais adequado em descrever a evolução dessa doença.

De modo análogo, Maki e Thompson, em seu modelo, também dividiram em classes: os ignorantes, os informantes e os neutros. Nesse caso, os ignorantes têm o mesmo papel dos suscetíveis do modelo SIR; os informantes são análogos aos infectados; e os neutros são análogos aos resistentes. A dinâmica ocorre por meio da interação entre os indivíduos informantes e os outros indivíduos da população. Nesse caso, quando um indivíduo informante tenta transmitir o rumor para outro informante ou um indivíduo neutro, somente aquele que tentou passar a informação torna-se neutro. Na Figura 9, apresenta a diagramação do modelo MT.

Figura 9 – Modelo compartimental MT.



Fonte: Produzido pelo autor.

Seguindo os mesmos moldes da Seção 1.2, seja $G = (V, A)$ um grafo qualquer com N vértices e um número finito de conexões em cada vértice, e suponhamos que o modelo seja um processo Markoviano a tempo contínuo e que seu espaço de estados seja $\mathcal{S} : V \mapsto \{0, 1, 2\}^N$. Os estados de cada vértice são codificados por um vetor *zero-um-dois* $\mathbf{X} = \{X_i : i \in V\}$, em que 0 representa o estado ignorante, 1 representa o informante e 2 representa o neutro. As taxas de transmissão da informação e com que os indivíduos se tornam neutros são, λ , para todo $(i, j) \in A$, e α , para todo $i \in V$. Suponha que os estados de todos os vértices no tempo t sejam $\mathbf{X} \in \mathcal{S}$, portanto:

$$\mathbb{P}(X_i(t+h) = 2 | \mathbf{X}(t) = \mathbf{X}) = \alpha(N_{\mathbf{X}}(i) + M_{\mathbf{X}}(i))h + o(h), \quad \text{se } X_i(t) = 1, \quad (1.9)$$

$$\mathbb{P}(X_i(t+h) = 1 | \mathbf{X}(t) = \mathbf{X}) = \lambda N_{\mathbf{X}}(i)h + o(h), \quad \text{se } X_i(t) = 0, \quad (1.10)$$

em que $N_{\mathbf{X}}(i) = |\{j \in V : j \sim i, X_j = 1\}|$ é o número de vizinhos de i que são informantes, e $M_{\mathbf{X}}(i) = |\{j \in V : j \sim i, X_j = 2\}|$ é a quantidade de vizinhos de i que são neutros. Portanto, cada vértice no estado informante vai para o estado neutro a uma taxa α , e transmite a informação a qualquer um de seus vizinhos a uma taxa λ .

1.5 CORRELAÇÕES EM MODELOS EPIDÊMICOS

Ao introduzirmos as relações espaciais em modelos de transmissão como os vistos nas seções anteriores, a natureza da dinâmica desses modelos torna seu estudo desafiador. Um caminho que normalmente é usado para reduzir a complexidade é aproximar as trajetórias do processo original por equações obtidas através de campo-médio. Esse caminho é escolhido principalmente quando o modelo é formulado para populações estruturadas. Dizemos que uma população é estruturada quando ela é representada por um grafo diferente do grafo

completo. Nessas condições, assume-se que os estados do processo são não-correlacionados. Se bem é sabido, tal afirmação não é verdadeira, contudo, pode ser considerada uma boa aproximação quando o grafo é muito conectado. Um outro fato importante no que se diz respeito as aproximações obtidas através de campo-médio, é que as se os estados dos vértices adjacentes são não-negativamente correlacionados, tais aproximações podem ser interpretadas como um limite superior das trajetórias do processo. Recentemente, uma demonstração foi desenvolvida por (CATOR; MIEGHEM, 2014) afirmando que a não-negatividade da correlação era satisfeita para os modelos de transmissão SIS e SIR em qualquer grafo. Contudo, foi mostrado por (RODRIGUEZ; ROLDÁN-CORREA; VALENCIA, 2018) que os argumentos usados para a prova da propriedade para o SIR não estão corretos. Devido aos modelos de rumores apresentarem uma dinâmica semelhante ao modelo SIR, entende-se que esta questão ainda é um problema em aberto para ambos os modelos. Diante disso, este trabalho tem como objetivo contribuir com a resolução desse problema.

1.6 ORGANIZAÇÃO DO TRABALHO

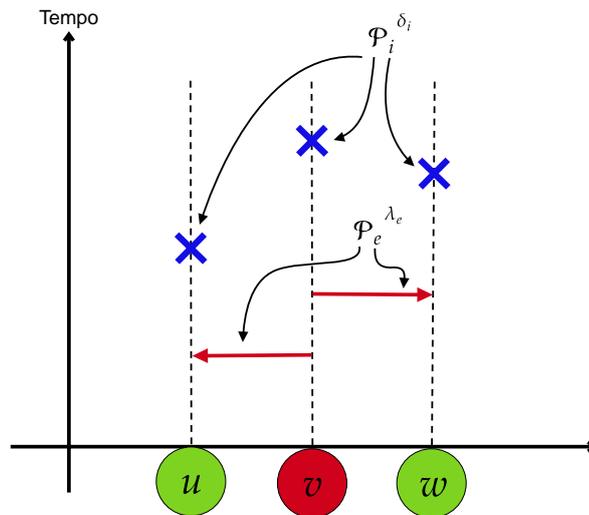
O Capítulo 1 tem caráter introdutório; nele trazemos alguns fatos históricos, bem como os principais conceitos utilizados ao longo deste trabalho. Por fim, construímos a formulação matemática dos modelos SIS, SIR e MT em grafos, em que somos naturalmente levados a uma discussão sobre correlação em modelos epidêmicos e a importância de desenvolver estudos nesse tópico. Por sua vez, o Capítulo 2 tem caráter teórico; nele fazemos uma revisão detalhada dos argumentos matemáticos apresentados por (CATOR; MIEGHEM, 2014) e (RODRIGUEZ; ROLDÁN-CORREA; VALENCIA, 2018). Em seguida, no Capítulo 3, são construídos os algoritmos para simulação dos modelos SIS, SIR e MT, bem como os métodos para estimar as grandezas de interesse. No Capítulo 4, são apresentados e discutidos os resultados obtidos pelos métodos desenvolvidos no Capítulo 3. Finalmente, no Capítulo 5, serão apresentadas as conclusões e perspectivas do trabalho desenvolvido. No Apêndice A, apresentamos as deduções das expressões matemáticas usadas nos algoritmos para simulação dos modelos SIS e SIR. No Apêndice B, apresentamos as deduções das expressões matemáticas usadas no algoritmo para simulação do modelo MT. E, por fim, no Apêndice C, serão fornecidos os códigos em linguagem Python utilizados para obtenção dos resultados deste trabalho.

2 ANÁLISE TEÓRICA

2.1 CONSTRUÇÃO GRÁFICA DE CATOR-VAN MIEGHEM PARA O MODELO SIS

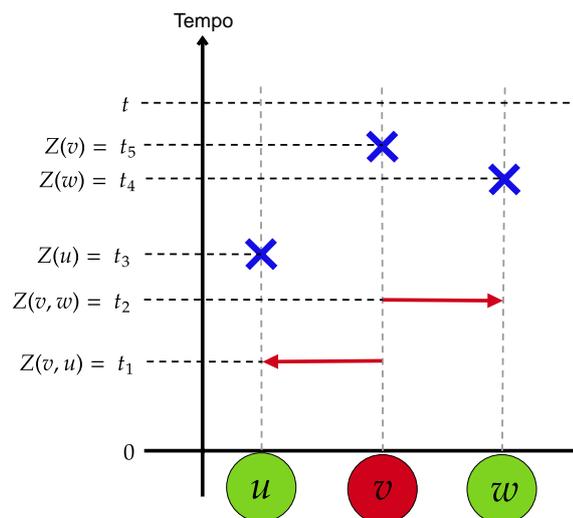
Esta seção será destinada à *construção gráfica de Cator - Van Mieghem* (CATOR; MIEGHEM, 2014) para o modelo SIS. Tal elaboração é baseada na *construção gráfica de Harris* (HARRIS, 1977a) que é bastante utilizada na literatura de *sistemas de partículas interagentes* (GRIMMETT, 2018). De forma geral, vamos considerar um grafo finito $G = (V, A)$ e assumir que cada vértice $i \in V$ pode estar em um de dois estados possíveis no instante de tempo t . Os estados são representados pela variável aleatória $X_i(t)$, sendo $X_i(t) = 0$, quando i estiver no estado suscetível e $X_i(t) = 1$, quando estiver infectado. A ideia que está por trás do método é a de representar o sistema por meio de um diagrama *espaço - tempo*, de modo que se pode identificar em quais momentos houveram tentativas de cura em um certo vértice $i \in V$ que esteja infectado e as respectivas tentativas de transmissão da infecção de i para seu vizinho j . Note que a transmissão ocorre de maneira direcionada do vértice i para j . Denotaremos por $F = \{(i, j) \in V \times V : \{i, j\} \in A\}$ o conjunto formado por todas arestas direcionadas $e = (i, j)$, e seja $W = (V, F)$ a coleção de todos os vértices e arestas direcionadas pertencentes ao grafo. A evolução dos estados de $i \in V$ até o instante t , será apresentada por meio das linhas do tempo, $\{i\} \times [0, t]$, e os instantes de possíveis transmissões pelas linhas de tempo $\{e\} \times [0, t]$. Considere que ao longo dessas linhas do tempo de cada elemento $i \in V$ e $e \in F$ sejam associados um conjunto de pontos seguindo processos de Poisson, $\mathcal{P}_i^{\delta_i}$ e $\mathcal{P}_e^{\lambda_e}$ de parâmetros δ_i e λ_e , respectivamente. Denotaremos tais conjuntos aleatórios como $A_C = \{\mathcal{P}_i^{\delta_i} : i \in V\}$ e $A_I = \{\mathcal{P}_e^{\lambda_e} : e \in F\}$, representando possíveis marcas de cura e setas de infecção, respectivamente. A ideia visual de tal construção é apresentada na Figura 10 para três vértices vizinhos, u , v e w , tais que $X_u(0) = 0$, $X_v(0) = 1$, $X_w(0) = 0$.

Figura 10 – Representação gráfica de Harris.



Fonte: Produzido pelo autor.

Seja $\mathcal{M} = \{\mathcal{B} : \mathcal{B} \subset [0, t] \text{ e } |\mathcal{B}| < \infty\}$ o espaço de todos os subconjuntos finitos de $[0, t]$. Note que $\mathcal{A} = \left[\left(\bigcup_{i \in V} A_I \right) \cup \left(\bigcup_{e \in F} A_C \right) \right]$ representa o conjunto aleatório de todos os eventos de cura e infecção em $[0, t]$. Seja $Z : W \mapsto \mathcal{M}$, tal que: Se $s \in Z(e)$ para $e = (i, j) \in F$, no instante de tempo s o vértice i tenta infectar j . De modo similar, se $s \in Z(i)$, para $i \in V$, então o vértice i é curado no instante de tempo t . A partir da Figura 11 pode-se visualizar o comportamento da função $Z(\cdot)$.

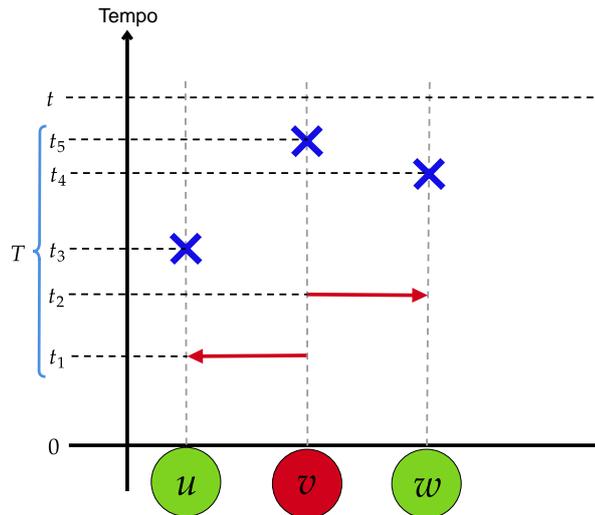
Figura 11 – Representação visual da função Z .

Fonte: Produzido pelo autor.

Agora que $Z(\cdot)$ está bem definida, e será possível traçar a evolução dos estados dos vértices no modelo SIS. Suponha que o vetor $\mathbf{X}(t)$ represente os estados de todos os vérti-

ces no modelo a partir de uma configuração inicial no instante de tempo $t_0 = 0$, em que $\mathbf{X}(0) = (X_1(0), X_2(0), \dots, X_m(0)) \in \{0, 1\}^N$. Vamos determinar os estados de todos os vértices no tempo $s > t_0$ a partir de um instante anterior a s . Considerando a coleção $T = \left[\left(\bigcup_{i \in V} Z(i) \right) \cup \left(\bigcup_{e \in E} Z(e) \right) \right]$ de todos os instantes de tempo que ocorreram eventos de cura e infecção em $[0, t]$.

Figura 12 – Representação visual da coleção de todos os instantes de tempo que ocorreram eventos em $[0, t]$.



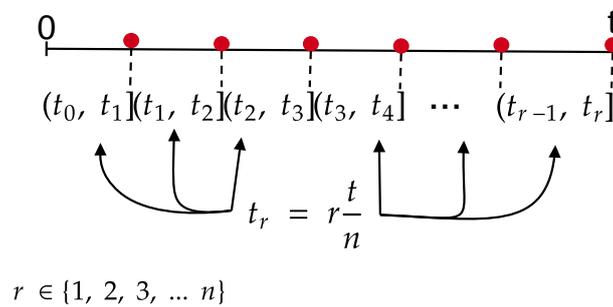
Fonte: Produzido pelo autor.

Note que T é um conjunto finito e pode ser ordenado. Em particular, existe algum $K \in \mathbb{N}$ tal que $T = \{t_1, t_2, \dots, t_K\}$ com $t_1 < t_2 < \dots < t_K$. Definimos $t_0 = 0$ e $t_{K+1} = t$, se o instante s satisfaz $t_k \leq s < t_{k+1}$ sendo $k \in \{0, 1, \dots, K\}$, temos que $X_m(s) = X_m(t_k)$ para todo $m \in V$. Agora suponha que $s \in T$, então, existe um vértice i tal que $s \in Z(i)$, ou existe uma aresta direcionada $e = (i, j)$, tal que $s \in Z(e)$. Se $s \in Z(i)$, então $X_m(s) = X_m(t_{k-1})$, para todo $m \neq i$ e $X_i(s) = 0$. Isto significa que no instante de tempo s , todos os vértices $m \neq i$ conservam seu estado atual enquanto ocorreu uma tentativa de cura no vértice i . De modo análogo, se $s \in Z(e)$, então $X_m(s) = X_m(t_{k-1})$, para todo $m \neq j$ e $X_j(s) = \max(X_i(t_{k-1}), X_j(t_{k-1}))$, que, nesse caso, significa que no instante s o vértice i tenta infectar j . Desse modo, existem duas possibilidades para o estado do vértice j no tempo s . Na primeira delas, se $X_i(t_{k-1}) = 1$ e $X_j(t_{k-1}) = 0$, tem-se que i infecta j , note que neste caso $X_j(s) = \max(1, 0) = 1$. Por outro lado, se $X_i(t_{k-1}) = 1$ e $X_j(t_{k-1}) = 1$, note que neste caso, $X_j(s) = \max(1, 1) = 1$, isto é, o estado de j permanece o mesmo pois ele já encontra-se infectado.

2.2 DISCRETIZAÇÃO DO MODELO SIS

Mais adiante nos depararemos com a necessidade de usar a desigualdade FKG (HOLLANDER; KEANE, 1986; FORTUIN; KASTELEYN; GINIBRE, 1971; FONTES, 1996), e para isto é vantajoso realizar a discretização do tempo para o modelo que estamos trabalhando. Para esta tarefa será definida uma quantidade n de pontos t_1, t_2, \dots, t_n em $[0, t]$ igualmente espaçados, em que o r -ésimo ponto é dado por $t_r = r \frac{t}{n}$ como podemos ver na Figura 13.

Figura 13 – Representação visual da discretização do intervalo $[0, t]$.



Fonte: Produzido pelo autor.

Na Seção 2.1, foi definida a função $Z(\cdot)$, que permitiu descrever a evolução do modelo a tempo contínuo. Para o caso discreto é natural pensar de maneira similar, definindo uma nova função que chamaremos de $Z_n(\cdot)$, que nos auxiliará a descrever os estados do modelo em cada um dos tempos t_r . Sabendo que a função $Z(i)$ atribui eventos de cura para cada vértice $i \in V$ em $[0, t]$, seguindo um processo de *Poisson* de parâmetro δ_i e $Z(e)$, associa eventos de infecção para cada aresta direcionada $e = (i, j) \in F$ em $[0, t]$, segundo um processo de *Poisson* com parâmetro λ_e . Sejam $Z_n(i, t_r)$ e $Z_n(e, t_r)$ definidas da seguinte maneira, respectivamente:

$$Z_n(i, t_r) = \begin{cases} 1, & \text{se } Z(i) \cap (t_{r-1}, t_r] = \emptyset, \\ 0, & \text{se } Z(i) \cap (t_{r-1}, t_r] \neq \emptyset. \end{cases} \quad (2.1)$$

$$Z_n(e, t_r) = \begin{cases} 0, & \text{se } Z(e) \cap (t_{r-1}, t_r] = \emptyset, \\ 1, & \text{se } Z(e) \cap (t_{r-1}, t_r] \neq \emptyset. \end{cases} \quad (2.2)$$

Note que a função $Z_n : W \times \{t_1, \dots, t_r\} \mapsto \{0, 1\}$, sendo $W = (V, F)$, mapeia os eventos de cura e infecção para cada vértice $i \in V$ e aresta direcionada $e = (i, j) \in F$, estabelecendo se há ou não ocorrência de tais eventos no tempo t_r . Note que se Z_n estiver sendo avaliada em um intervalo $(t_{r-1}, t_r]$, que contém duas ou mais marcas de cura ou infecção, todas elas

poderão ser substituídas por apenas uma única marca nesse mesmo intervalo. Essa substituição é justificada pelo fato de tais eventos produzirem sempre o mesmo estado no intervalo em que estão sendo analisados.

Agora que $Z_n(\cdot)$ está bem definida, vamos usá-la para descrever os estados infecciosos dos vértices em cada t_r . Primeiramente, serão descritas as tentativas de cura; para realizar esta tarefa, vamos definir um vetor de estados auxiliar $Y_i^{(n)}(t_r)$ em termos de $X_i^{(n)}(t_{r-1})$ e $Z_n(i, t_r)$, que tem como função descrever o estado do vértice i no tempo t_r a partir do instante t_{r-1} . Seja $Y_i^{(n)}(t_r)$ dado por:

$$Y_i^{(n)}(t_r) := X_i^{(n)}(t_{r-1})Z_n(i, t_r). \quad (2.3)$$

Fazendo uso da Equação 2.3, os estados infecciosos do vértice i podem ser descritos por $X_i^{(n)}(t_r) = Y_i^{(n)}(t_r)$. Note que, em particular para os eventos de cura, temos que $Z(i) \cap (t_{r-1}, t_r] = \emptyset$, implica que $Z_n(i, t_r) = 1$, portanto $Y_i^{(n)}(t_r) = X_i^{(n)}(t_{r-1})(1)$, dessa forma concluímos que $X_i^{(n)}(t_r) = X_i^{(n)}(t_{r-1})$. A interpretação para $X_i^{(n)}(t_r) = X_i^{(n)}(t_{r-1})$ é que se não houver marca de cura no intervalo $(t_{r-1}, t_r]$, o vértice i permanece em seu estado atual. Por outro lado, se $Z(i) \cap (t_{r-1}, t_r] \neq \emptyset$, logo $Z_n(i, t_r) = 0$, implica que $Y_i^{(n)}(t_r) = X_i^{(n)}(t_{r-1})(0) = 0$, portanto $X_i^{(n)}(t_r) = 0$.

Para descrever o estado do vértice j que se encontra na aresta direcionada $e = (i, j)$, definiremos a seguinte equação:

$$X_j^{(n)}(t_r) := Y_j^{(n)}(t_r) + [1 - Y_j^{(n)}(t_r)] \max_{i:(i,j) \in F} \left(Y_i^{(n)}(t_r) Z_n(e, t_r) \right). \quad (2.4)$$

Note que, em particular para os eventos de infecção, temos que $Z(e) \cap (t_{r-1}, t_r] \neq \emptyset$, então $Z_n(e, t_r) = 1$. Se $Y_i^{(n)}(t_r) = 1$, pela Equação 2.4, segue que:

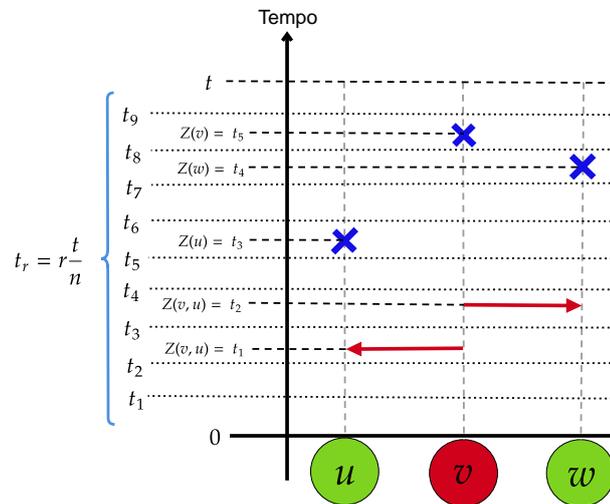
$$\begin{aligned} X_j^{(n)}(t_r) &= Y_j^{(n)}(t_r) + [1 - Y_j^{(n)}(t_r)] \max_{i:(i,j) \in F} \left(Y_i^{(n)}(t_r) Z_n(e, t_r) \right) \\ &= Y_j^{(n)}(t_r) + [1 - Y_j^{(n)}(t_r)] \max_{i:(i,j) \in F} \left((1)(1) \right) \\ &= Y_j^{(n)}(t_r) + [1 - Y_j^{(n)}(t_r)] \max_{i:(i,j) \in F} \left(1 \right) \\ &= 1. \end{aligned} \quad (2.5)$$

Por outro lado, $Z(e) \cap (t_{r-1}, t_r] = \emptyset$ implica que $Z_n(e, t_r) = 0$ e, portanto, pela Equação 2.4, concluímos que:

$$\begin{aligned}
X_j^{(n)}(t_r) &= Y_j^{(n)}(t_r) + [1 - Y_j^{(n)}(t_r)] \max_{i:(i,j) \in F} \left(Y_i^{(n)}(t_r) Z_n(e, t_r) \right) \\
&= Y_j^{(n)}(t_r) + [1 - Y_j^{(n)}(t_r)] \max_{i:(i,j) \in F} \left(Y_i^{(n)}(t_r)(0) \right) \\
&= Y_j^{(n)}(t_r) + [1 - Y_j^{(n)}(t_r)] \max_{i:(i,j) \in F} (0) \\
&= Y_j^{(n)}(t_r).
\end{aligned} \tag{2.6}$$

A interpretação para esse resultado é que se não existem marcas de infecção no intervalo $(t_{r-1}, t_r]$, então o vértice j permanece em seu estado atual. Com a finalidade de fixar o método utilizado para descrever os estados dos vértices i e j fazendo uso das Equações 2.3 e 2.4, respectivamente, vamos aplicá-las em um exemplo concreto. Seja, $\mathbf{X}(0) = \{X_u(0) = 0, X_v(0) = 1, X_w(0) = 0\}$ uma configuração inicial para o modelo em $t = 0$ dos vértices u, v e w , apresentados na Figura 10, em uma realização do processo em $[0, t]$, em que a Figura 14 retrata tal realização, bem como a discretização do intervalo para $n = 10$.

Figura 14 – Representação visual de uma realização do processo e a discretização de $[0, t]$.



Fonte: Produzido pelo autor.

Vamos supor que se faz necessário saber o estado do vértice v em t_3 . Sabemos que v pode vir a ser **curado** ou tentar **infectar** um de seus vizinhos, que são u e w . Pela Figura 14, nota-se que $Z(v) \cap (t_2, t_3] = \emptyset$, então, pela Equação 2.1 $Z_n(v, t_3) = 1$. Portanto, pela Equação 2.3, temos que:

$$\begin{aligned}
Y_v^{(n)}(t_3) &= X_v^{(n)}(t_2)Z_n(v, t_3) \\
&= (1)(1) \\
&= 1.
\end{aligned} \tag{2.7}$$

Concluimos, então, que $X_v^{(n)}(t_3) = Y_v^{(n)}(t_3) = 1$. O vértice v não altera seu estado em t_3 . Por outro lado, temos que, para o vértice u que pertence à aresta direcionada $e = (v, u)$, $Z(e) \cap (t_2, t_3] \neq \emptyset$, logo $Z_n(e, t_3) = 1$, então, pela Equação 2.20, temos que o estado de u é dado por:

$$\begin{aligned}
X_u^{(n)}(t_3) &= Y_u^{(n)}(t_3) + [1 - Y_u^{(n)}(t_3)] \max_{v:(v,u) \in F} \left(Y_v^{(n)}(t_3) Z_n(e, t_3) \right) \\
&= Y_u^{(n)}(t_3) + [1 - Y_u^{(n)}(t_3)] \max_{v:(v,u) \in F} \left((1)(1) \right) \\
&= Y_u^{(n)}(t_3) + [1 - Y_u^{(n)}(t_3)] \max_{v:(v,u) \in F} \left(1 \right) \\
&= Y_u^{(n)}(t_3) + [1 - Y_u^{(n)}(t_3)] \\
&= 1.
\end{aligned} \tag{2.8}$$

Concluimos que o vértice u passa do estado suscetível para infectado no instante de tempo t_3 . Por sua vez, temos que para o vértice w pertencente a aresta direcionada $e = (v, w)$, $Z(e) \cap (t_2, t_3] = \emptyset$, logo $Z_n(e, t_3) = 0$, então, pela Equação 2.20, temos que o estado de w é dado por:

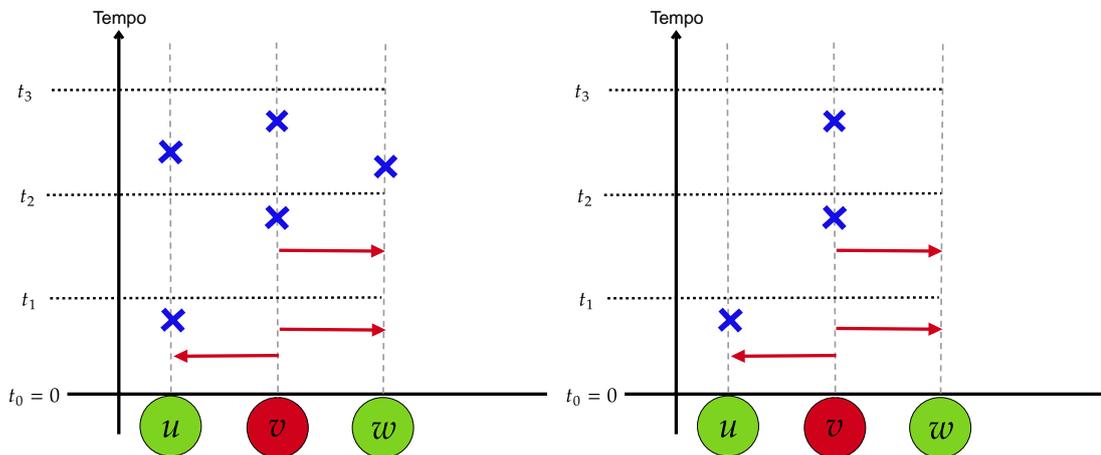
$$\begin{aligned}
X_w^{(n)}(t_3) &= Y_w^{(n)}(t_3) + [1 - Y_w^{(n)}(t_3)] \max_{v:(v,w) \in F} \left(Y_v^{(n)}(t_3) Z_n(e, t_3) \right) \\
&= Y_w^{(n)}(t_3) + [1 - Y_w^{(n)}(t_3)] \max_{v:(v,w) \in F} \left((1)(0) \right) \\
&= Y_w^{(n)}(t_3) + [1 - Y_w^{(n)}(t_3)] \max_{v:(v,w) \in F} \left(0 \right) \\
&= Y_w^{(n)}(t_3).
\end{aligned} \tag{2.9}$$

Portanto, concluimos que o vértice w continua em seu estado suscetível no instante de tempo t_3 . Observe que apenas um dos três vértices mudou de estado, o que já era esperado, pois apenas um evento por vez ocorre em cada tempo t_r . Com esse exemplo, vimos que toda a evolução do processo, em todos os tempos t_1, t_2, \dots, t_r , pode ser descrita fazendo uso da função $Z_n(\cdot)$ em conjunto com as Equações 2.20 e 2.3.

2.2.1 Monotonicidade da função $Z_n(\cdot)$

Nesta seção será dedicada em elucidar, por meio de diagramas, a monotonicidade da função $Z_n(\cdot)$ a partir da construção realizada nas seções anteriores. Considerando os vértices apresentados na Figura 10, suponha que seja fixada uma configuração inicial em $t_0 = 0$ para três vértices vizinhos, u , v e w , tais que $X_u(0) = 0, X_v(0) = 1, X_w(0) = 0$. Considere que o processo seja realizado a partir dessa configuração inicial de acordo com duas funções $Z^{(1)}(\cdot)$ e $Z^{(2)}(\cdot)$ e que, para discretizar o tempo, usemos $n = 3$. Podemos ver tais realizações representadas por meio dos diagramas na Figura 15.

Figura 15 – Realização do processo a partir de $Z^{(1)}(\cdot)$ e $Z^{(2)}(\cdot)$.



(a) Realização do processo **SIS** de acordo com a função $Z^{(1)}(\cdot)$. (b) Realização do processo **SIS** de acordo com a função $Z^{(2)}(\cdot)$.

Fonte: Produzido pelo autor.

Agora vamos analisar o comportamento das funções $Z^{(1)}(\cdot)$ e $Z^{(2)}(\cdot)$ para cada vértice e aresta, para cada intervalo de tempo $(t_{r-1}, t_r]$, tomando como referência os diagramas da Figura 15. Sabendo que $t_r \in \{t_0, t_1, t_2, t_3\}$, para o vértice u temos que $Z^{(1)}(u) \cap (0, t_1] \neq \emptyset$ e $Z^{(2)}(u) \cap (0, t_1] = \emptyset$, logo $Z_n^{(1)}(u, t_1) = 1$ e $Z_n^{(2)}(u, t_1) = 0$, respectivamente, portanto $Z_n^{(1)}(u, t_1) > Z_n^{(2)}(u, t_1)$. Repetindo o processo para t_2 , temos que $Z^{(1)}(u) \cap (t_1, t_2] = \emptyset$ e $Z^{(2)}(u) \cap (t_1, t_2] = \emptyset$, logo $Z_n^{(1)}(u, t_2) = 0$ e $Z_n^{(2)}(u, t_2) = 0$, portanto $Z_n^{(1)}(u, t_2) = Z_n^{(2)}(u, t_2)$. E, por fim, avaliando o estado de u em t_3 , temos que $Z^{(1)}(u) \cap (t_2, t_3] \neq \emptyset$ e $Z^{(2)}(u) \cap (t_2, t_3] = \emptyset$, logo $Z_n^{(1)}(u, t_3) = 1$ e $Z_n^{(2)}(u, t_3) = 0$, portanto, concluímos que $Z_n^{(1)}(u, t_3) > Z_n^{(2)}(u, t_3)$. Repetindo a metodologia para os vértices w e v , chegamos à conclusão que:

$$\forall i \in \{u, v, w\}, r \in \{0, 1, 2, 3\} : Z_n^{(1)}(i, t_r) \leq Z_n^{(2)}(i, t_r). \quad (2.10)$$

Vamos realizar o mesmo procedimento para as arestas (v, u) e (v, w) . Sendo $e = (v, u)$ e sabendo que $t_r \in \{t_0, t_1, t_2, t_3\}$, temos que para t_1 , $Z^{(1)}(e) \cap (0, t_1] \neq \emptyset$ e $Z^{(2)}(e) \cap (0, t_1] \neq \emptyset$, logo $Z_n^{(1)}(e, t_1) = 1$ e $Z_n^{(2)}(e, t_1) = 1$, portanto $Z_n^{(1)}(e, t_1) = Z_n^{(2)}(e, t_1)$. Avaliando para t_2 , temos que $Z^{(1)}(e) \cap (t_1, t_2] = \emptyset$ e $Z^{(2)}(e) \cap (t_1, t_2] = \emptyset$, logo $Z_n^{(1)}(e, t_2) = 1$ e $Z_n^{(2)}(e, t_2) = 1$, portanto $Z_n^{(1)}(e, t_2) = Z_n^{(2)}(e, t_2)$. O mesmo se repete para t_3 , isto é, $Z_n^{(1)}(e, t_3) = Z_n^{(2)}(e, t_3)$. Fazendo a mesma análise para $e = (v, w)$, concluímos que:

$$\forall e \in \{(v, u), (v, w)\}, r \in \{0, 1, 2, 3\} : Z_n^{(1)}(e, t_r) \leq Z_n^{(2)}(e, t_r). \quad (2.11)$$

Podemos concluir que $Z_n(\cdot)$ é monótona crescente se:

$$\forall i \in V, r \in \{1, \dots, n\} : Z_n^{(1)}(i, t_r) \leq Z_n^{(2)}(i, t_r), \quad (2.12)$$

$$\forall e \in F, r \in \{1, \dots, n\} : Z_n^{(1)}(e, t_r) \leq Z_n^{(2)}(e, t_r). \quad (2.13)$$

2.2.2 Monotonicidade dos estados

Como foi visto na Seção 2.2.1, a função $Z_n(\cdot)$ segue uma ordem parcial em que a monotonicidade é dada se:

$$\forall i \in V, r \in \{1, \dots, n\} : Z_n^{(1)}(i, t_r) \leq Z_n^{(2)}(i, t_r), \quad (2.14)$$

$$\forall e \in F, r \in \{1, \dots, n\} : Z_n^{(1)}(e, t_r) \leq Z_n^{(2)}(e, t_r). \quad (2.15)$$

Note que $Z^{(1)}$ e $Z^{(2)}$ levam a diferentes evoluções do processo SIS, isto implica que para cada uma delas vamos ter diferentes valores de estados no mesmo instante de tempo $t = t_r$. Para realizar a prova, vamos estabelecer a seguinte notação $X_i^{(n), Z_n}(t_r)$ para explicitar a dependência de $X_i^{(n)}(t_r)$ por $Z_n(\cdot)$. Seja $Z_n^{(1)}(i, t_r) \leq Z_n^{(2)}(i, t_r)$, então pela construção temos que:

$$X_i^{(n), Z_n^{(1)}}(t_r) \leq X_i^{(n), Z_n^{(2)}}(t_r). \quad (2.16)$$

Essa é uma propriedade importante para o modelo SIS, pois será possível fazer uso da desigualdade FKG. Note que essa propriedade é válida para $t_0 = 0$, em que $X_i^{(n), Z_n^{(1)}}(t_r) =$

$X_i^{(n),Z_n^{(2)}}(t_r)$. Agora provaremos que $Y_i^{(n),Z_n^{(1)}}(t_r) \leq Y_i^{(n),Z_n^{(2)}}(t_r)$. Pela Equação 2.16, temos que $X_i^{(n),Z_n^{(1)}}(t_r) \leq X_i^{(n),Z_n^{(2)}}(t_r)$, sabendo que $Z_n^{(1)}(i, t_r) \leq Z_n^{(2)}(i, t_r)$, então, pela Equação 2.3, temos que:

$$Y_i^{(n),Z_n^{(1)}}(t_r) = X_i^{(n),Z_n^{(1)}}(t_{r-1})Z_n^{(1)}(i, t_r), \quad (2.17)$$

$$Y_i^{(n),Z_n^{(2)}}(t_r) = X_i^{(n),Z_n^{(2)}}(t_{r-1})Z_n^{(2)}(i, t_r). \quad (2.18)$$

Portanto, pela Equação 2.16, concluímos que:

$$Y_i^{(n),Z_n^{(1)}}(t_r) \leq Y_i^{(n),Z_n^{(2)}}(t_r). \quad (2.19)$$

Seguindo adiante, vamos provar para $e = (i, j) \in F$ que $X_j^{(n),Z_n^{(1)}}(t_r) \leq X_j^{(n),Z_n^{(2)}}(t_r)$. Sabendo que $Z_n^{(1)}(e, t_r) \leq Z_n^{(2)}(e, t_r)$, pela Equação 2.4, temos que:

$$X_j^{(n),Z_n^{(1)}}(t_r) = Y_j^{(n),Z_n^{(1)}}(t_r) + [1 - Y_j^{(n),Z_n^{(1)}}(t_r)] \max_{e \in F} \left(Y_i^{(n),Z_n^{(1)}}(t_r) Z_n^{(1)}(e, t_r) \right), \quad (2.20)$$

$$X_j^{(n),Z_n^{(2)}}(t_r) = Y_j^{(n),Z_n^{(2)}}(t_r) + [1 - Y_j^{(n),Z_n^{(2)}}(t_r)] \max_{e \in F} \left(Y_i^{(n),Z_n^{(2)}}(t_r) Z_n^{(2)}(e, t_r) \right). \quad (2.21)$$

Se $Y_j^{(n),Z_n^{(1)}}(t_r) = 1$, a prova é direta, pois, pela Equação 2.19, sabemos que também $Y_j^{(n),Z_n^{(2)}}(t_r) = 1$, então, pela Equação 2.20, concluímos que:

$$X_i^{(n),Z_n^{(1)}}(t_r) = X_i^{(n),Z_n^{(2)}}(t_r). \quad (2.22)$$

Para $Y_j^{(n),Z_n^{(1)}}(t_r) = 0$, temos duas possibilidades, uma para $Y_j^{(n),Z_n^{(2)}}(t_r) = 1$ e outra $Y_j^{(n),Z_n^{(2)}}(t_r) = 0$. Se $Y_j^{(n),Z_n^{(2)}}(t_r) = 1$, então, pela Equação 2.20, temos que $X_j^{(n),Z_n^{(2)}}(t_r) = 1$ e $X_j^{(n),Z_n^{(1)}}(t_r) = \max_{i:(i,j) \in F} \left(Y_j^{(n),Z_n^{(1)}}(t_r) Z_n^{(1)}(e, t_r) \right)$, que, pela Equação 2.19, temos que $\max_{i:(i,j) \in F} \left(Y_j^{(n),Z_n^{(1)}}(t_r) Z_n^{(1)}(e, t_r) \right) \leq 1$. Portanto, concluímos que:

$$X_j^{(n),Z_n^{(1)}}(t_r) \leq X_j^{(n),Z_n^{(2)}}(t_r). \quad (2.23)$$

E, por fim, vamos analisar para o caso em que $Y_j^{(n),Z_n^{(2)}}(t_r) = 0$. De mesmo modo, temos, pela Equação 2.19, que: $X_j^{(n),Z_n^{(1)}}(t_r) = \max_{i:(i,j) \in F} \left(Y_j^{(n),Z_n^{(1)}}(t_r) Z_n^{(1)}(e, t_r) \right)$ e $X_j^{(n),Z_n^{(2)}}(t_r) = \max_{i:(i,j) \in F} \left(Y_j^{(n),Z_n^{(2)}}(t_r) Z_n^{(2)}(e, t_r) \right)$, portanto, concluímos pelas Equações 2.16 e 2.19 que:

$$X_j^{(n),Z_n^{(1)}}(t_r) \leq X_j^{(n),Z_n^{(2)}}(t_r). \quad (2.24)$$

Com esse resultado completamos a prova que $X_j^{(n), Z_n^{(1)}}(t_r) \leq X_j^{(n), Z_n^{(2)}}(t_r)$ e para $X_i^{(n), Z_n^{(1)}}(t_r) \leq X_i^{(n), Z_n^{(2)}}(t_r)$. Com isso, concluímos que os estados do processo SIS são monótonos não-decrescentes.

2.2.3 Correlação entre vizinhos infectados

Os estados infecciosos do processo SIS têm a propriedade de serem positivamente correlacionados, isto é, a ocorrência de um estado infectado em um vértice i , por exemplo, aumenta a probabilidade do seu vizinho j ser infectado. Essa propriedade é consequência da monotonicidade dos estados do processo, como foi mostrado na Seção 2.2.2, os estados do processo são monótonos não-decrescentes. Para provar que os estados do processo SIS são positivamente correlacionados, vamos fazer uso da desigualdade FKG, cujo enunciado se encontra no Lema 2.

A desigualdade FKG surgiu em 1971, quando *Fortuin, Kasteleyn e Ginibre* provaram que para funções monótonas em certos conjuntos de eventos parcialmente ordenados sob uma condição de convexidade adequada na medida de probabilidade, os eventos não-decrescentes estão positivamente correlacionados, enquanto que os decrescentes estão correlacionados negativamente. Essa desigualdade generalizou outras desigualdades obtidas anteriormente por (HARRIS, 1960), em seu trabalho sobre modelos de percolação, e por (GRIFFITHS; HURST; SHERMAN, 1970; GINIBRE, 1970), em seus trabalhos sobre as correlações de spin em modelos de spin ferromagnéticos. Dessa forma, a desigualdade FKG tornou-se uma ferramenta padrão para análise teórica de diversos modelos estocásticos, e também encontrou aplicação em várias outras áreas, como teoria dos grafos e combinatória (HOLLANDER; KEANE, 1986). Mais tarde, em 1974, foram obtidas generalizações da desigualdade FKG, em que foi provada uma desigualdade entre médias de funções monótonas sob duas medidas de probabilidade distintas (HOLLEY, 1974) e, posteriormente, em 1978, também foi obtida uma desigualdade de tipo mais combinatório da qual também a desigualdade provada por (HOLLEY, 1974) pode ser derivada como sendo um caso especial.

Lema 2. *Sejam Z e Y duas variáveis aleatórias crescentes e limitadas em Ω . Então*

$$\mathbb{E}(ZY) \geq \mathbb{E}(Z)\mathbb{E}(Y) \quad (2.25)$$

Sabendo que para todo n , as desigualdades são válidas $X_j^{(n), Z_n^{(1)}}(t_r) \leq X_j^{(n), Z_n^{(2)}}(t_r)$ e $X_i^{(n), Z_n^{(1)}}(t_r) \leq X_i^{(n), Z_n^{(2)}}(t_r)$, e que ao fazermos $n \rightarrow \infty$, recuperamos o processo original a

tempo contínuo. Portanto, temos que as desigualdades mencionadas anteriormente também são válidas para o processo original. Desse modo, concluímos, a partir do Lema 2, que

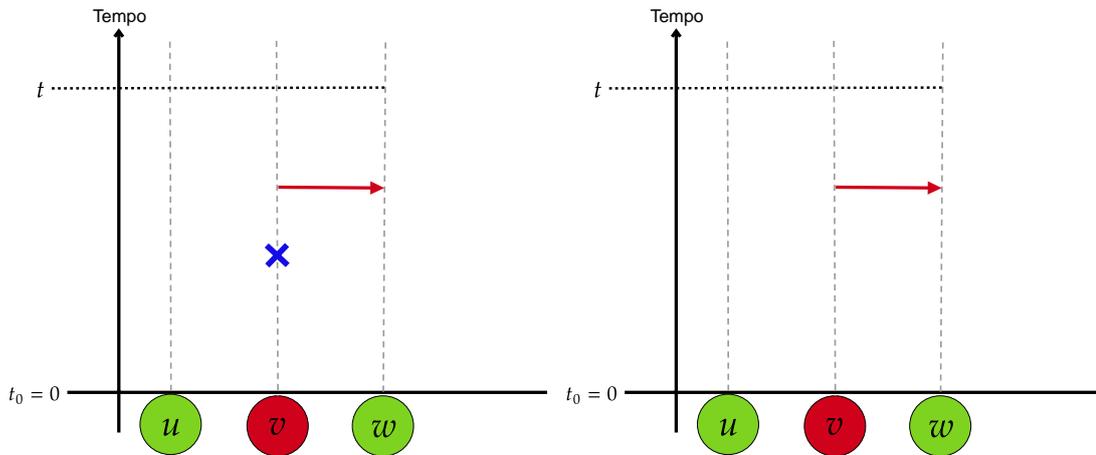
$$\mathbb{E}(X_i(t)X_j(t)) \geq \mathbb{E}(X_i(t))\mathbb{E}(X_j(t)), \quad (2.26)$$

é válido para o processo SIS.

2.3 A NÃO MONOTONICIDADE DO PROCESSO SIR

Como foi visto na Seção 2.1, (CATOR; MIEGHEM, 2014), inspirados pela construção gráfica de *Harris*, nos fornecem uma forma de descrever toda a evolução do processo SIS ao longo de todos os instantes de tempo em $[0, t]$ para todo $t > 0$. Segundo os mesmos autores, a mesma técnica poderia ser usada para provar a desigualdade apresentada no Lema 2 para o processo SIR. De modo que a mesma função $Z(\cdot)$ poderia ser usada, com a única diferença de que se um dado vértice infectado é curado no instante s , seu estado deverá ser alterado para recuperado. De acordo com (RODRIGUEZ; ROLDÁN-CORREA; VALENCIA, 2018), a construção apresentada por (CATOR; MIEGHEM, 2014) só é válida para processos monótonos, o que não é o caso do processo SIR. Em seu exemplo são apresentadas duas realizações do processo SIR, como podem ser vistas na Figura 16, em que são descritas por $Z^{(1)}(\cdot)$ e $Z^{(2)}(\cdot)$, tal que $Z^{(1)} \leq Z^{(2)}$ e $X^{(n), Z_1} = X^{(n), Z_2}(0) = (0, 1, 0)$. Note que $X^{(n), Z_1}(t) = (0, 2, 0)$ e $X^{(n), Z_2}(t) = (0, 1, 1)$, mesmo começando com a mesma configuração inicial e $Z^{(1)} \leq Z^{(2)}$, os respectivos estados do processo no instante t formam duas configurações que não são comparáveis entre si. Desse modo, pode-se concluir que a construção de *Cator-Van Mieghem* não funciona para o modelo SIR. Neste caso, não é possível usar a desigualdade FKG com a mesma abordagem apresentada em (CATOR; MIEGHEM, 2014). Ainda que a construção de *Cator-Van Mieghem* seja interessante, útil para processos que são monótonos e forneça uma prova do Lema 2 para o processo SIS, pontuamos que o mesmo já foi demonstrado para processos monótonos, em (DONNELLY, 1993; HARRIS, 1977b). Deste modo (RODRIGUEZ; ROLDÁN-CORREA; VALENCIA, 2018) fornecem um contraexemplo considerando os valores das variáveis aleatórias do processo SIR original.

Figura 16 – Comparação entre duas realizações do processo SIR de acordo com $Z^{(1)}(\cdot)$ e $Z^{(2)}(\cdot)$ de modo que $Z^{(1)} \leq Z^{(2)}$.



(a) Realização do processo **SIS** de acordo com a função $Z^{(1)}(\cdot)$. (b) Realização do processo **SIS** de acordo com a função $Z^{(2)}(\cdot)$.

Fonte: Produzido pelo autor.

Vamos calcular a covariância entre os estados dos vértices 1 e 2 em $t = 1$, sendo $X_1(0) = 1$ e $X_2(0) = 0$ a configuração inicial dos vértices 1 e 2, respectivamente. O próximo passo será o cálculo das probabilidades do sistema estar em cada uma das suas possíveis configurações em $t = 1$, dado a configuração inicial $X_1(0) = 1$ e $X_2(0) = 0$. Note que o conjunto das possíveis configurações do processo no instante $t = 1$ é:

$$\{(X_1(1) = 1, X_2(1) = 0), (X_1(1) = 1, X_2(1) = 1), (X_1(1) = 1, X_2(1) = 2), \\ (X_1(1) = 2, X_2(1) = 0), (X_1(1) = 2, X_2(1) = 1), (X_1(1) = 2, X_2(1) = 2)\}.$$

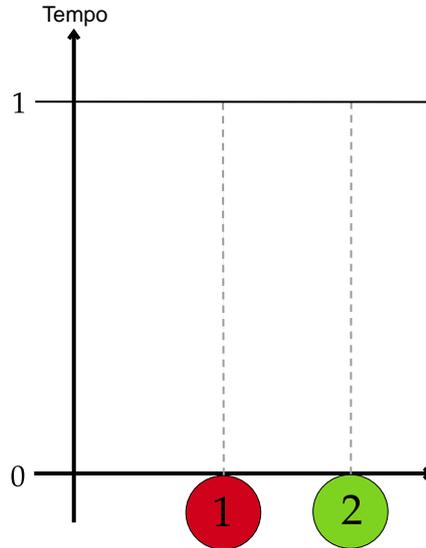
As probabilidades de cada uma dessas configurações podem ser calculadas observando o comportamento das marcas do Processo de Poisson $\mathcal{P}^{\lambda_{(1,2)}}$, \mathcal{P}^{δ_1} e \mathcal{P}^{δ_2} no intervalo de tempo $[0, 1]$, condicionando-as ao evento $\mathcal{X} = \{X_1(0) = 1, X_2(0) = 0\}$. Temos que:

$$\mathbb{P}(X_1(1) = 1, X_2(1) = 0 | \mathcal{X}) = e^{-1}e^{-2}, \quad (2.27)$$

pois, para manter os estados em ambos os vértices no instante de tempo 1, não pode ocorrer nenhuma marca de cura no vértice 1 e nenhuma seta de infecção do vértice 1 para seu vizinho 2, durante a janela de tempo $[0, 1]$. Note que isso é o mesmo que dizer que o evento $\{S_1^\delta > 1, S_{(1,2)}^\lambda > 1\}$ ocorreu, em que S_1^δ é o tempo de ocorrência da primeira marca de cura

no vértice 1 e $S_{(1,2)}^\lambda$, por sua vez, é o tempo de ocorrência da primeira seta de infecção do vértice 1 para 2. Na Figura 17, podemos observar o comportamento do processo nessa ocasião, por meio da representação gráfica de Harris.

Figura 17 – Comportamento das marcas favorável ao evento $\{X_1(1) = 1, X_2(1) = 0\}$.



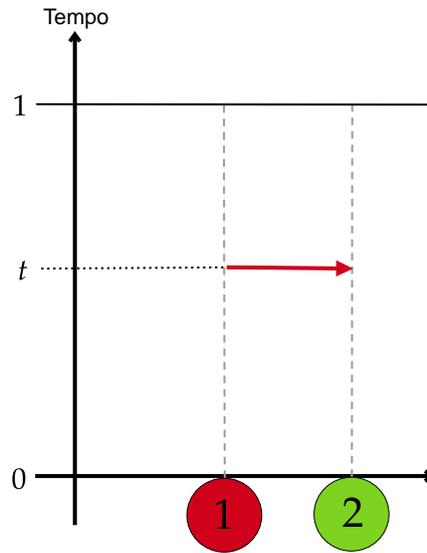
Fonte: Produzido pelo autor.

A probabilidade de $\{X_1(1) = 1, X_2(1) = 1\}$ dado \mathcal{X} , a partir do teorema da probabilidade total para variáveis aleatórias contínuas (ROSS, 2009) pode ser escrita da seguinte forma:

$$\mathbb{P}(X_1(1) = 1, X_2(1) = 1 | \mathcal{X}) = \int_0^1 \mathbb{P}(X_1(1) = 1, X_2(1) = 1 | \mathcal{X}, S_{(1,2)}^\lambda = t) f_{S_{(1,2)}^\lambda}(t) dt, \quad (2.28)$$

em que $f_{S_{(1,2)}^\lambda}(t) = 2e^{-2t} \mathbf{1}_{\{t>0\}}$ é a densidade da variável aleatória $S_{(1,2)}^\lambda$. Note que essa é uma forma conveniente de calcular essa probabilidade, pois sabemos que para o evento $\{X_1(1) = 1, X_2(1) = 1\}$ ocorrer, necessariamente em algum momento durante a janela de tempo $[0, 1]$ ocorreu alguma seta de infecção do vértice 1 para o seu vizinho 2. Também não houve nenhuma marca de cura no vértice 1 durante a janela de tempo $[0, 1]$. Uma possível representação gráfica dessa descrição é mostrada na Figura 18 a partir do diagrama proposto por Harris.

Figura 18 – Comportamento das marcas favorável ao evento $\{X_1(1) = 1, X_2(1) = 1\}$.



Fonte: Produzido pelo autor.

Note que o evento $\{X_1(1) = 1, X_2(1) = 1\}$ é equivalente ao evento $\{S_1^\delta > 1, S_1^\delta > 1 - t\}$; ver Figura 18. Portanto, temos que:

$$\begin{aligned}
 \mathbb{P}(X_1(1) = 1, X_2(1) = 1 | \mathcal{X}, S_{(1,2)}^\lambda = t) &= \mathbb{P}(S_1^\delta > 1, S_1^\delta > 1 - t | \mathcal{X}, S_{(1,2)}^\lambda = t) \quad (2.29) \\
 &= \underbrace{\mathbb{P}(S_1^\delta > 1) \mathbb{P}(S_1^\delta > 1 - t)}_{\text{independência}} \\
 &= e^{-1} e^{-(1-t)} \mathbb{1}_{\{0 < t < 1\}},
 \end{aligned}$$

substituindo esse resultado na Equação 2.28, concluímos que:

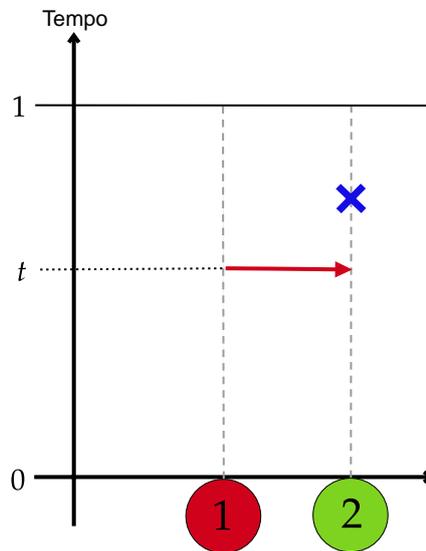
$$\begin{aligned}
 \mathbb{P}(X_1(1) = 1, X_2(1) = 1 | \mathcal{X}) &= \int_0^1 e^{-1} e^{-(1-t)} 2e^{-2t} dt \quad (2.30) \\
 &= 2e^{-2} \int_0^1 e^{-t} dt \\
 &= 2e^{-2} (1 - e^{-1}).
 \end{aligned}$$

De modo análogo, podemos calcular a probabilidade da configuração $\{X_1(1) = 1, X_2(1) = 2\}$, dado a condição inicial \mathcal{X} condicionado ao evento $\{S_{(1,2)}^\lambda = t\}$. Logo:

$$\mathbb{P}(X_1(1) = 2, X_2(1) = 0 | \mathcal{X}) = \int_0^1 \mathbb{P}(X_1(1) = 2, X_2(1) = 0 | \mathcal{X}, S_1^\delta = t) f_{S_1^\delta}(t) dt, \quad (2.31)$$

note que, mais uma vez, é conveniente fazer esse condicionamento, pois para que o vértice 2 esteja recuperado em $t = 1$, necessariamente ocorreu uma seta de infecção do vértice 1 para 2, na janela de tempo $[0, 1]$ e, em algum momento, ocorreu alguma marca de cura em 2, na janela $[t, 1]$. Podemos ver na Figura 19 um diagrama de uma das possíveis formas de como esses eventos podem ocorrer de modo a gerar a configuração de interesse.

Figura 19 – Comportamento das marcas favorável ao evento $\{X_1(1) = 1, X_2(1) = 2\}$.



Fonte: Produzido pelo autor.

Note que o evento $\{X_1(1) = 1, X_2(1) = 2\}$ é equivalente ao evento $\{S_1^\delta > 1, S_2^\delta < 1 - t\}$; ver Figura 19. Portanto, temos que:

$$\begin{aligned} \mathbb{P}(X_1(1) = 1, X_2(1) = 2 | \mathcal{X}, S_{(1,2)}^\lambda = t) &= \mathbb{P}(S_1^\delta > 1, S_2^\delta < 1 - t | \mathcal{X}, S_{(1,2)}^\lambda = t) \quad (2.32) \\ &= \underbrace{\mathbb{P}(S_1^\delta > 1) \mathbb{P}(S_2^\delta < 1 - t)}_{\text{independência}} \\ &= e^{-1} (1 - e^{-(1-t)}) \mathbb{1}_{\{0 < t < 1\}}, \end{aligned}$$

substituindo esse resultado na equação 2.31, concluímos que:

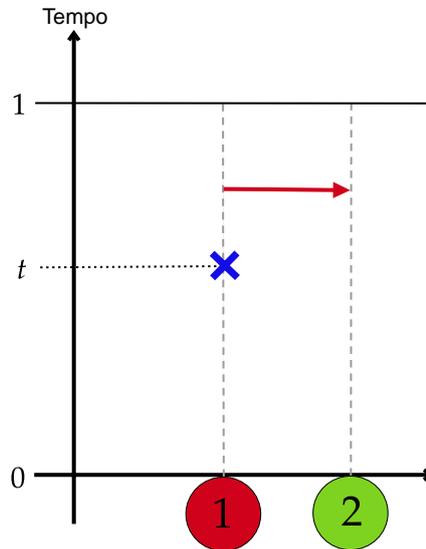
$$\begin{aligned} \mathbb{P}(X_1(1) = 1, X_2(1) = 2 | \mathcal{X}) &= \int_0^1 e^{-1} (1 - e^{-(1-t)}) 2e^{-2t} dt & (2.33) \\ &= 2e^{-1} \int_0^1 (e^{-2t} - e^{-1-t}) dt \\ &= e^{-3} (e - 1)^2. \end{aligned}$$

Seguindo as mesmas ideias, podemos calcular a probabilidade da configuração $\{X_1(1) = 2, X_2(1) = 0\}$, dado a condição inicial \mathcal{X} condicionada ao evento $\{S_1^\delta = t\}$. Logo:

$$\mathbb{P}(X_1(1) = 2, X_2(1) = 0 | \mathcal{X}) = \int_0^1 \mathbb{P}(X_1(1) = 2, X_2(1) = 0 | \mathcal{X}, S_1^\delta = t) f_{S_1^\delta}(t) dt, \quad (2.34)$$

em que $f_{S_1^\delta}(t) = e^{-t} \mathbf{1}_{\{t > 0\}}$ é a densidade da variável aleatória S_1^δ . Vemos que, mais uma vez, essa é uma forma conveniente de calcular essa probabilidade, pois sabemos que para que o evento $\{X_1(1) = 2, X_2(1) = 0\}$ ocorra, necessariamente em algum momento ocorreu alguma marca de cura em 1 na janela $[0, 1]$, antes de ocorrer alguma seta de infecção do vértice 1 para 2, como mostrado na Figura 20.

Figura 20 – Comportamento das marcas favorável ao evento $\{X_1(1) = 2, X_2(1) = 0\}$.



Fonte: Produzido pelo autor.

Note que o evento $\{X_1(1) = 2, X_2(1) = 0\}$ é equivalente ao evento $\{S_1^\delta < S_{(1,2)}^\lambda\}$; ver Figura 20. Portanto, temos que:

$$\mathbb{P}(X_1(1) = 2, X_2(1) = 0 | \mathcal{X}) = \int_0^1 \mathbb{P}(S_1^\delta < S_{(1,2)}^\lambda | \mathcal{X}, S_1^\delta = t) f_{S_1^\delta}(t) dt, \quad (2.35)$$

por outro lado,

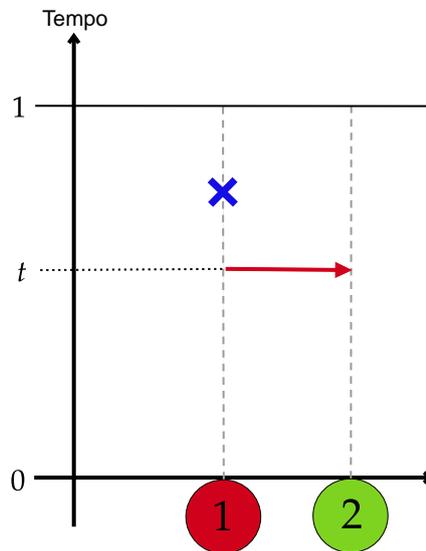
$$\begin{aligned} \mathbb{P}(S_1^\delta < S_{(1,2)}^\lambda | \mathcal{X}, S_1^\delta = t) &= \mathbb{P}(t < S_{(1,2)}^\lambda | \mathcal{X}) \\ &= \mathbb{P}(t < S_{(1,2)}^\lambda) \\ &= e^{-2t} \mathbf{1}_{\{0 < t < 1\}}, \end{aligned} \quad (2.36)$$

substituindo a expressão final da Equação 2.37 na Equação 2.35, concluímos que:

$$\begin{aligned} \mathbb{P}(X_1(1) = 2, X_2(1) = 0 | \mathcal{X}) &= \int_0^1 e^{-2t} e^{-t} dt \\ &= \frac{1 - e^{-3}}{3}. \end{aligned} \quad (2.37)$$

Prosseguindo, vamos calcular a probabilidade da configuração $\{X_1(1) = 2, X_2(1) = 1\}$, dado a condição inicial \mathcal{X} seguindo as mesmas ideias anteriores. Mais uma vez, vamos condicionar ao evento $\{S_1^\delta = t\}$, pois para que o vértice 2 chegue no tempo 1 no estado infectado, necessariamente tem que ocorrer uma marca de infecção do vértice 1 para o 2, como podemos ver na Figura 21 uma possível configuração das marcas que geram o evento de interesse.

Figura 21 – Comportamento das marcas favorável ao evento $\{X_1(1) = 2, X_2(1) = 1\}$.



Fonte: Produzido pelo autor.

Aplicando o que foi dito, temos que:

$$\mathbb{P}(X_1(1) = 2, X_2(1) = 1 | \mathcal{X}) = \int_0^1 \mathbb{P}(X_1(1) = 2, X_2(1) = 0 | \mathcal{X}, S_{(1,2)}^\lambda = t) f_{S_{(1,2)}^\lambda}(t) dt. \quad (2.38)$$

note que o evento $\{X_1(1) = 2, X_2(1) = 1\}$ é equivalente ao evento $\{S_1^\delta < 1 - t, S_2^\delta > 1\}$, como pode ser visto na Figura 21. Desse modo temos que:

$$\begin{aligned} \mathbb{P}(X_1(1) = 2, X_2(1) = 1 | \mathcal{X}, S_{(1,2)}^\lambda = t) &= \mathbb{P}(S_1^\delta > 1, S_2^\delta < 1 - t | \mathcal{X}, S_{(1,2)}^\lambda = t) \quad (2.39) \\ &= \underbrace{\mathbb{P}(S_1^\delta > 1) \mathbb{P}(S_2^\delta < 1 - t)}_{\text{independência}} \\ &= e^{-1} (1 - e^{-(1-t)}) \mathbb{1}_{\{0 < t < 1\}}, \end{aligned}$$

substituindo esse resultado na Equação 2.38, concluímos que:

$$\begin{aligned} \mathbb{P}(X_1(1) = 2, X_2(1) = 1 | \mathcal{X}) &= \int_0^1 e^{-1} (1 - e^{-(1-t)}) 2e^{-2t} dt \quad (2.40) \\ &= 2e^{-1} \int_0^1 (e^{-2t} - e^{-1-t}) dt \\ &= e^{-3} (e - 1)^2. \end{aligned}$$

Finalmente, calcularemos a probabilidade de $\{X_1(1) = 2, X_2(1) = 2\}$ dado \mathcal{X} . Vamos utilizar o fato de que:

$$\sum_{\substack{1 \leq x_1 \leq 2 \\ 0 \leq x_2 \leq 2}} \mathbb{P}(X_1(1) = x_1, X_2(1) = x_2 | \mathcal{X}) = 1, \quad (2.41)$$

portanto, a probabilidade $\mathbb{P}(X_1(1) = 2, X_2(1) = 2 | \mathcal{X})$ pode ser calculada por complemento, isto é:

$$\begin{aligned} \mathbb{P}(X_1(1) = 2, X_2(1) = 2 | \mathcal{X}) &= 1 - \sum_{\substack{1 \leq x_1 \leq 2 \\ 0 \leq x_2 < 2}} \mathbb{P}(X_1(1) = x_1, X_2(1) = x_2 | \mathcal{X}) \quad (2.42) \\ &= 1 - \left(e^{-1} e^{-2} + 2e^{-2} (1 - e^{-1}) \right. \\ &\quad \left. + e^{-3} (e - 1)^2 + \frac{1 - e^{-3}}{3} + e^{-3} (e - 1)^2 \right) \\ &= \frac{2e^{-3} (e - 1)^3}{3}. \end{aligned}$$

Para o cálculo das esperanças de $X_1(1)$ e $X_2(1)$, precisaremos das suas respectivas distribuições marginais. Sabendo que:

$$\mathbb{P}(X_1(1) = x_1 | \mathcal{X}) = \sum_{0 \leq x_2 \leq 2} \mathbb{P}(X_1(1) = x_1, X_2(1) = x_2 | \mathcal{X}), \quad (2.43)$$

logo,

$$\begin{aligned} \mathbb{P}(X_1(1) = 1 | \mathcal{X}) &= \sum_{0 \leq x_2 \leq 2} \mathbb{P}(X_1(1) = 1, X_2(1) = x_2 | \mathcal{X}) \\ &= e^{-3} + 2e^{-2}(1 - e^{-1}) + e^{-3}(e - 1)^2 \\ &= e^{-1}, \end{aligned} \quad (2.44)$$

$$\begin{aligned} \mathbb{P}(X_1(1) = 2 | \mathcal{X}) &= \sum_{0 \leq x_2 \leq 2} \mathbb{P}(X_1(1) = 2, X_2(1) = x_2 | \mathcal{X}) \\ &= \frac{1 - e^{-3}}{3} + e^{-3}(e - 1)^2 + \frac{2e^{-3}(e - 1)^3}{3} \\ &= \frac{3e^3 - 3e^2}{3e^3}, \end{aligned} \quad (2.45)$$

$$\begin{aligned} \mathbb{P}(X_2(1) = 0 | \mathcal{X}) &= \sum_{1 \leq x_1 \leq 2} \mathbb{P}(X_1(1) = x_1, X_2(1) = 0 | \mathcal{X}) \\ &= e^{-3} + \frac{1 - e^{-3}}{3} \\ &= \frac{1 + 2e^{-3}}{3}, \end{aligned} \quad (2.46)$$

$$\begin{aligned} \mathbb{P}(X_2(1) = 1 | \mathcal{X}) &= \sum_{1 \leq x_1 \leq 2} \mathbb{P}(X_1(1) = x_1, X_2(1) = 1 | \mathcal{X}) \\ &= 2e^{-2}(1 - e^{-1}) + e^{-3}(e - 1)^2 \\ &= e^{-1} - e^{-3}, \end{aligned} \quad (2.47)$$

$$\begin{aligned} \mathbb{P}(X_2(1) = 2 | \mathcal{X}) &= \sum_{1 \leq x_1 \leq 2} \mathbb{P}(X_1(1) = x_1, X_2(1) = 2 | \mathcal{X}) \\ &= e^{-3}(e - 1)^2 + \frac{2e^{-3}(e - 1)^3}{3} \\ &= \frac{2e^3 - 3e^2 + 1}{3e^3}. \end{aligned} \quad (2.48)$$

Agora podemos calcular as esperanças, fazendo:

$$\begin{aligned}
\mathbb{E}(X_1(1)|\mathcal{X}) &= \sum_{1 \leq x_1 \leq 2} x_1 \mathbb{P}(X_1(1) = x_1 | \mathcal{X}) & (2.49) \\
&= (1)\mathbb{P}(X_1(1) = 1 | \mathcal{X}) + (2)\mathbb{P}(X_1(1) = 2 | \mathcal{X}) \\
&= e^{-1} + 2 \left(\frac{3e^3 - 3e^2}{3e^3} \right) \\
&= \frac{6e^3 - 3e^2}{3e^3},
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}(X_2(1)|\mathcal{X}) &= \sum_{0 \leq x_2 \leq 2} x_2 \mathbb{P}(X_2(1) = x_2 | \mathcal{X}) & (2.50) \\
&= (0)\mathbb{P}(X_2(1) = 0 | \mathcal{X}) + (1)\mathbb{P}(X_2(1) = 1 | \mathcal{X}) + (2)\mathbb{P}(X_2(1) = 2 | \mathcal{X}) \\
&= e^{-1} - e^{-3} + 2 \left(\frac{2e^3 - 3e^2 + 1}{3e^3} \right) \\
&= \frac{4e^3 - 3e^2 - 1}{3e^3},
\end{aligned}$$

por sua vez,

$$\begin{aligned}
\mathbb{E}(X_1(1), X_2(1)|\mathcal{X}) &= \sum_{\substack{1 \leq x_1 \leq 2 \\ 0 \leq x_2 \leq 2}} x_1 x_2 \mathbb{P}(X_1(1) = x_1, X_2(1) = x_2 | \mathcal{X}) \\
&= 2e^{-2}(1 - e^{-1}) + e^{-3}(e - 1)^2 + e^{-3}(e - 1)^2 + 4 \frac{2e^{-3}(e - 1)^3}{3} \\
&= \frac{8e^3 - 18e^2 + 18e - 8}{3e^3}.
\end{aligned}$$

Agora temos todas as expressões necessárias para o cálculo da covariância, logo:

$$\begin{aligned}
COV(X_1(1), X_2(1)|\mathcal{X}) &= \mathbb{E}(X_1(1), X_2(1)|\mathcal{X}) - \left(\mathbb{E}(X_1(1)|\mathcal{X}) \mathbb{E}(X_2(1)|\mathcal{X}) \right) & (2.51) \\
&= \frac{8e^3 - 18e^2 + 18e - 8}{3e^3} - \left[\left(\frac{6e^3 - 3e^2}{3e^3} \right) \left(\frac{4e^3 - 3e^2 - 1}{3e^3} \right) \right] \\
&= \frac{8e^3 - 18e^2 + 18e - 8}{3e^3} - \frac{24e^6 - 30e^5 + 9e^4 - 6e^3 + 3e^2}{9e^6} \\
&= -\frac{1}{3e^4} - \frac{2}{e^3} + \frac{5}{e^2} - \frac{8}{3e} < 0,
\end{aligned}$$

portanto, concluímos que:

$$\mathbb{E}(X_1(1), X_2(1)) < \mathbb{E}(X_1(1))\mathbb{E}(X_2(1)). \quad (2.52)$$

Desse modo, vemos que o contra-exemplo apresentado por (RODRIGUEZ; ROLDÁN-CORREA; VALENCIA, 2018), mostra que os argumentos apresentados por (CATOR; MIEGHEM, 2014) não podem ser estendidos para o processo SIR.

3 SIMULAÇÃO DE PROCESSOS DE TRANSMISSÃO

3.1 ALGORITMOS

3.1.1 Algoritmo para simulação do processo SIS

Para a realização de simulações de processos estocásticos se faz necessário converter as taxas em probabilidades, de modo que seja preservado na média o número de eventos por unidade de tempo (COTA, 2017). Inicialmente, assumimos que para o modelo SIS em um grafo $G = (V, A)$, o vértice infectado $i \in V$ torna-se suscetível a uma taxa δ_i e transmite a doença para seu vizinho $j \in V$ a uma taxa $\lambda_{(i,j)}$. Em nosso trabalho, vamos focar no modelo em que todas as taxas de infecção e cura são iguais para todos os vértices e arestas do grafo. Portanto, considerando que o processo é uma dinâmica *Markoviana*, o modelo SIS a tempo contínuo pode ser simulado de maneira simples a partir dos seguintes passos.

Passo 1 — Inicialização: Defina um grafo $G = (V, A)$ com N vértices. Fixe as taxas de cura e infecção de modo que δ_i e $\lambda_{(i,j)}$ sejam constantes para todo $i \in V$ e $(i, j) \in A$, respectivamente. Defina uma configuração inicial para o sistema, fazendo com que para todo $i \in V$, assuma o estado infectado com uma probabilidade p . Após a inicialização, os vértices infectados e suscetíveis serão armazenados nas listas $\mathcal{E}^{(inf)} = \{\mathcal{E}_i^{(inf)}\}$ e $\mathcal{E}^{(sus)} = \{\mathcal{E}_i^{(sus)}\}$, respectivamente. Defina as listas X_I, X_S, T_{tempo} em que serão armazenados os números de infectados, suscetíveis e o tempo decorrido, respectivamente. Defina a posição inicial dos dados a serem armazenados na lista como $n = 0$.

Passo 2: Selecione um dos vértices $i \in \mathcal{E}^{(inf)}$ com probabilidade:

$$\epsilon_i = \frac{\delta + n_i^S \lambda}{N_I \delta + N_S \lambda} \quad \forall i \in \mathcal{E}^{(inf)}, \quad (3.1)$$

em que n_i^S é a quantidade de vizinhos do vértice i no estado suscetível, N_I é a quantidade total de vértices infectados e N_S é a soma dos vizinhos suscetíveis de todos os vértices infectados dada por:

$$N_S = \sum_{i \in \mathcal{E}^{(inf)}} n_i^S. \quad (3.2)$$

As deduções das expressões usadas **Passo 2** encontram-se no Apêndice A.

Passo 3: O vértice escolhido no **Passo 2** torna-se **suscetível** com probabilidade:

$$\mu_i = \frac{\delta}{\delta + n_i^S \lambda}, \quad (3.3)$$

em que n_i^S é a quantidade de vizinhos suscetíveis do vértice i escolhido. Caso contrário, com igual probabilidade, escolha um vértice $j \in \mathcal{E}^{(sus)}$ vizinho a i e o torne **infectado**. Por fim, atualize as listas $\mathcal{E}^{(sus)}$ e $\mathcal{E}^{(inf)}$, respectivamente. Ver Apêndice A.

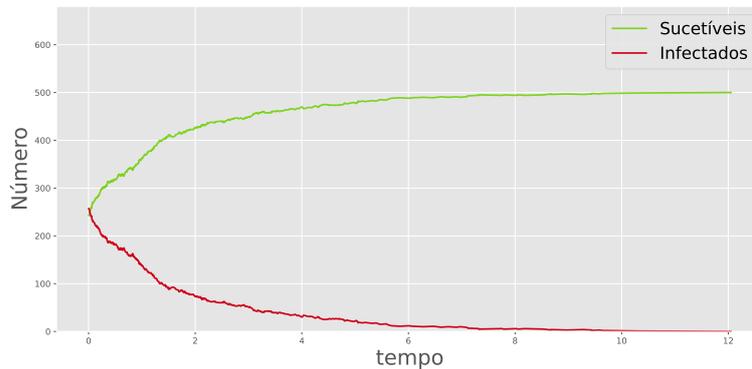
Passo 4: Faça $t = t + \Delta t$, em que $\Delta t \sim \exp(N_I \delta + N_S \lambda)$.

Passo 5: Salve as quantidades de infectados, suscetíveis e o tempo decorrido na posição n das listas X_I , X_S , T_{tempo} e faça $n = n + 1$, respectivamente.

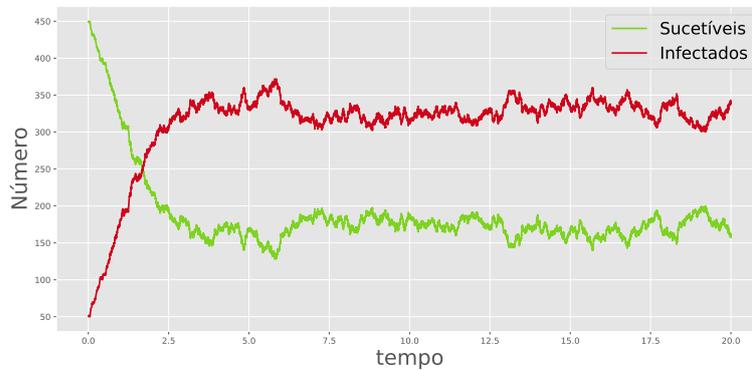
Passo 6: Enquanto houver vértices infectados, **repita** os passos 2, 3, 4 e 5.

Na Figura 22 apresentamos algumas simulações da dinâmica do modelo SIS obtidas a partir da aplicação do algoritmo apresentado. Para isso, foi escolhido um grafo completo com $N = 500$ vértices e dois pares distintos de λ e δ .

Figura 22 – Simulações para o modelo SIS no grafo completo de 500 vértices.



(a) $\lambda = 0,001$, $\delta = 1$ e $p = 0,5$.



(b) $\lambda = 0,003$, $\delta = 1$ e $p = 0,1$.

Fonte: Produzido pelo autor.

Note que, nesse caso, obtemos duas dinâmicas com comportamentos distintos. No primeiro gráfico, Figura 22a, o número de infectados vai a zero, pois o mesmo se encontra no regime subcrítico, em razão de $\lambda = 0,001 < \lambda_{crítico}$. Por outro lado, a dinâmica do processo apresentado na Figura 22b se encontra no regime supercrítico, pois $\lambda = 0,006 > \lambda_{crítico}$. Note, também, que o processo encontra-se variando em torno de uma média, então dizemos que o mesmo está em seu estado estacionário.

3.1.2 Algoritmo para simulação do processo SIR

Para o modelo SIR, a simulação tem os mesmos passos e parâmetros apresentados na Seção 3.1.1, com a diferença que no **Passo 3** o vértice torna-se recuperado e que em $t = 0$ apenas um vértice encontra-se no estado infectado. Uma vez que algum vértice assume o estado recuperado, o mesmo vértice não poderá assumir os estados suscetível e infectado ao longo de todo o resto do processo. Portanto, temos os seguintes passos para simular o modelo SIR:

Passo 1 — Inicialização: Defina um grafo $G = (V, A)$ com N vértices. Fixe as taxas de cura e infecção de modo que δ_i e $\lambda_{(i,j)}$ sejam constantes para todo $i \in V$ e $(i, j) \in A$, respectivamente. Defina uma configuração inicial para o sistema, sorteando um vértice da rede que começará no estado infectado. Após a inicialização, os vértices infectados e suscetíveis serão armazenados nas listas $\mathcal{E}^{(inf)} = \{\mathcal{E}_i^{(inf)}\}$ e $\mathcal{E}^{(sus)} = \{\mathcal{E}_i^{(sus)}\}$, respectivamente. Defina as listas X_S , X_I , X_R e T_{tempo} , em que serão armazenadas as quantidades de suscetíveis, infectados, recuperados e o tempo decorrido. Defina a posição inicial dos dados a serem armazenados na lista como $n = 0$.

Passo 2: Selecione um dos vértices $i \in \mathcal{E}^{(inf)}$ com probabilidade:

$$\epsilon_i = \frac{\delta + n_i^S \lambda}{N_I \delta + N_S \lambda} \quad \forall i \in \mathcal{E}^{(inf)}, \quad (3.4)$$

em que n_i^S é a quantidade de vizinhos do vértice i no estado suscetível, N_I é a quantidade total de vértices infectados e N_S é a soma dos vizinhos suscetíveis de todos os vértices infectados dada por:

$$N_S = \sum_{i \in \mathcal{E}^{(inf)}} n_i^S. \quad (3.5)$$

As deduções das expressões usadas **Passo 2** encontram-se no Apêndice A.

Passo 3: O vértice escolhido no **Passo 2** torna-se **recuperado** com probabilidade:

$$\mu_i = \frac{\delta}{\delta + n_i^S \lambda}, \quad (3.6)$$

em que n_i^S é a quantidade de vizinhos do vértice i no estado suscetível. Caso contrário, com igual probabilidade, escolha um vértice $j \in \mathcal{E}^{(sus)}$ vizinho a i e o torne **infectado**. Por fim, atualize as listas $\mathcal{E}^{(sus)}$, $\mathcal{E}^{(inf)}$ e $\mathcal{E}^{(rec)}$, respectivamente. Ver Apêndice A.

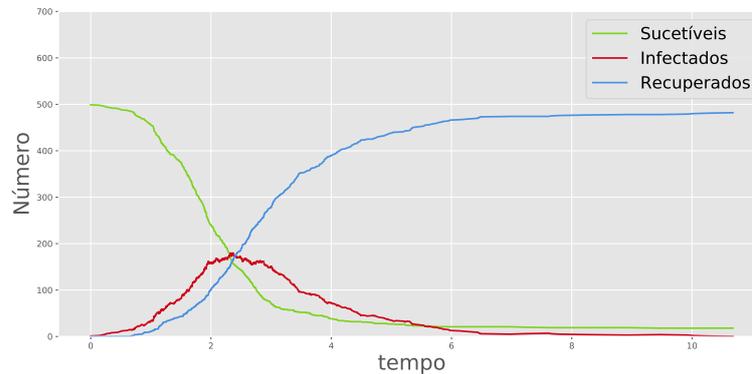
Passo 4: Faça $t = t + \Delta t$ em que $\Delta t \sim \exp(N_I \delta + N_S \lambda)$.

Passo 5: Salve as quantidades de infectados, suscetíveis, recuperados e o tempo decorrido na posição n das listas X_S , X_I , X_R , T_{tempo} e faça $n = n + 1$, respectivamente.

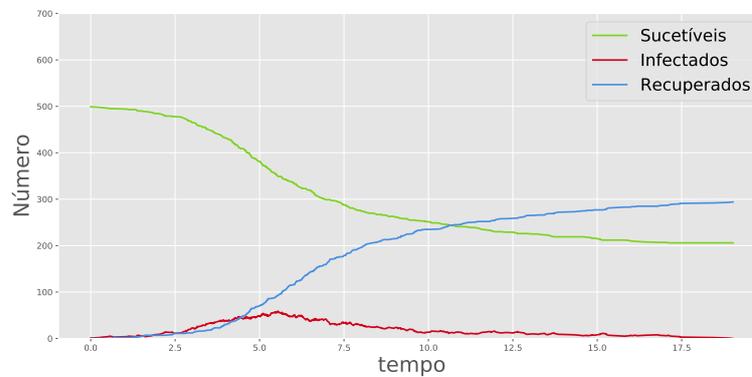
Passo 6: Enquanto houver vértices infectados, **repita** os passos 2, 3, 4 e 5.

Na Figura 23 apresentamos algumas simulações da dinâmica do modelo SIR obtidas a partir da aplicação do algoritmo apresentado. Para isso, foi escolhido um grafo completo com $N = 500$ vértices e dois pares distintos de λ e δ .

Figura 23 – Simulações para o modelo SIR no grafo completo de 500 vértices.



(a) $\lambda = 0,007$ e $\delta = 1$



(b) $\lambda = 0,0025$ e $\delta = 1$

Fonte: Produzido pelo autor.

Note que a dinâmica do processo apresentado na Figura 23a exibe um pico no número de infectados em um tempo menor que o processo da Figura 23b, devido ao fato de $\lambda_a > \lambda_b$. Também o tempo de extinção do processo da Figura 23a ocorre de maneira mais rápida, essa diferença também pode ser explicada observando o valor de λ usado em cada um dos processos.

3.1.3 Algoritmo para simulação do processo de Maki-Thompson

O método para simulação do modelo MT também segue os mesmos moldes do SIR apresentados na Seção 3.1.2, porém com pequenas mudanças nas expressões 3.4, 3.6 e nas taxas, que, nesse caso, α é a taxa com que um ignorante torna-se neutro e λ é a taxa de propagação.

Passo 1 — Inicialização: Defina um grafo $G = (V, A)$ com N vértices. Fixe as taxas de neutralização e propagação, de modo que α_i e $\lambda_{(i,j)}$ sejam constantes para todo $i \in V$ e $(i, j) \in A$, respectivamente. Defina uma configuração inicial para o sistema, sorteando um vértice da rede que começará no estado informante. Após a inicialização, os vértices informantes e ignorantes serão armazenados nas listas $\mathcal{E}^{(inf)} = \{\mathcal{E}_i^{(inf)}\}$ e $\mathcal{E}^{(ig)} = \{\mathcal{E}_i^{(ig)}\}$, respectivamente. Defina as listas X_{In} , X_{Ig} , X_N e T_{tempo} , em que serão armazenadas as quantidades de ignorantes, informantes, neutros e o tempo decorrido. Defina a posição inicial dos dados a serem armazenados na lista como $n = 0$.

Passo 2: Selecione um dos vértices $i \in \mathcal{E}^{(inf)}$ com probabilidade:

$$\epsilon_i = \frac{(k_i - n_i^{Ig})\alpha + n_i^{Ig}\lambda}{N_{IN}\alpha + N_{Ig}\lambda} \quad \forall i \in \mathcal{E}^{(inf)}, \quad (3.7)$$

em que n_i^{Ig} é a quantidade de vizinhos do vértice i no estado ignorante, k_i o grau do vértice i , N_{Ig} é a soma dos vizinhos ignorantes de todos os vértices informantes. E, por fim, N_{IN} é a soma dos não ignorantes vizinhos de todos os vértices informantes, dado por:

$$N_{Ig} = \sum_{i \in \mathcal{E}^{(inf)}} n_i^{Ig}, \quad (3.8)$$

$$N_{IN} = \sum_{i \in \mathcal{E}^{(inf)}} (k_i - n_i^{Ig}). \quad (3.9)$$

As deduções das expressões usadas no **Passo 2** encontram-se no Apêndice B.

Passo 3: O vértice escolhido no **Passo 2** torna-se **neutro** com probabilidade:

$$\mu_i = \frac{(k_i - n_i^{Ig})\alpha}{(k_i - n_i^{Ig})\alpha + n_i^{Ig}\lambda}, \quad (3.10)$$

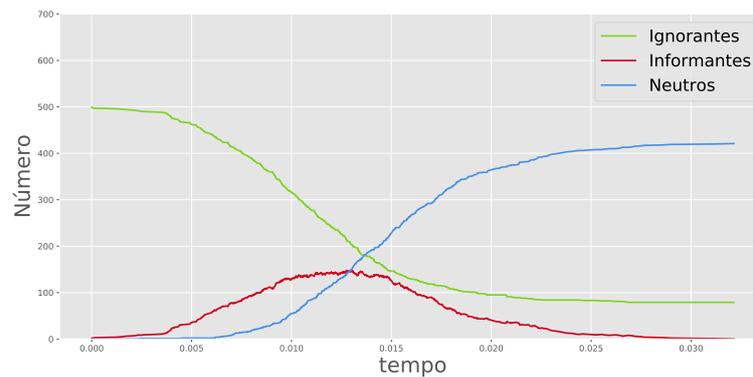
em que n_i^{Ig} é a quantidade de vizinhos do vértice i no estado ignorante e k_i é o grau do vértice i (ver Apêndice B). Caso contrário, com igual probabilidade, escolha um vértice $j \in \mathcal{E}^{(ig)}$ vizinho a i e o torne **informante**. Por fim, atualize as listas $\mathcal{E}^{(ig)}$ e $\mathcal{E}^{(inf)}$, respectivamente.

Passo 4: Faça $t = t + \Delta t$, em que $\Delta t \sim \exp(N_{IN}\alpha + N_{Ig}\lambda)$.

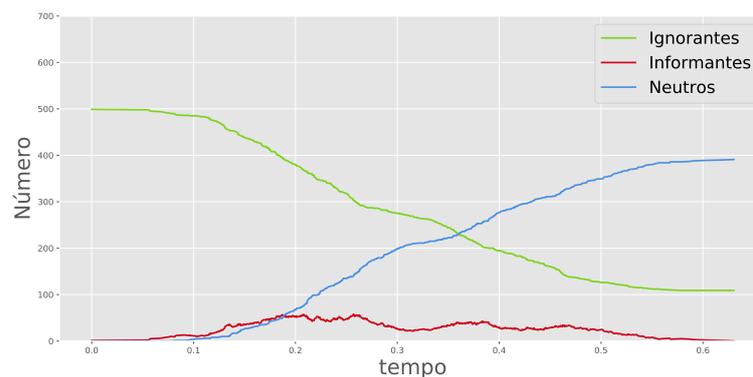
Passo 5: Salve as quantidades de informantes, ignorantes, neutros e o tempo decorrido na posição n das listas X_{In} , X_{Ig} , X_N , T_{tempo} e faça $n = n + 1$, respectivamente. **Passo 6:** Enquanto houver vértices informantes, **repita** os passos **2**, **3**, **4** e **5**.

Na Figura 24 apresentamos algumas simulações da dinâmica do modelo MT obtidas a partir da aplicação do algoritmo apresentado. Para isso, foi escolhido um grafo completo com $N = 500$ vértices (Figura 24a) e um grafo k -regular com $k = 50$ e $N = 500$ (Figura 24b).

Figura 24 – Simulações para o modelo MT no grafo completo e k -regular.



(a) $\alpha = 1$ e $\lambda = 1$



(b) $\alpha = 1$ e $\lambda = 1$

Fonte: Produzido pelo autor.

Note que, como já foi mencionado, existe uma forte semelhança entre as dinâmicas do modelo SIR e MT. Contudo, o processo MT acaba mais rapidamente em comparação com o SIR. Isso se dá pela mudança na forma como acontecem as mudanças do estado informante

para neutro. No SIR, essa mudança ocorre de maneira espontânea com uma taxa δ , não dependendo da topologia do grafo. Por outro lado, no MT ocorre a partir da interação entre informante - informante e informante - neutro. Desse modo, existe uma dependência intrínseca com a topologia do grafo.

3.1.3.1 Validação do Algoritmo para simulação do Modelo Maki-Thompson

Um resultado interessante para o modelo Maki-Thompson é que a densidade final de ignorantes no grafo completo para valores de $N \rightarrow \infty$ tem distribuição normal com média $\mu = 0,20318787$ e variância $\sigma^2 = 0,27273575$ (WATSON, 1987). Esse resultado foi posteriormente generalizado por (LEBENSZTAYN; MACHADO; RODRÍGUEZ, 2011). Essa conclusão teórica é uma ferramenta útil e pode ser usada para verificar se o algoritmo apresentado na Seção 3.1.3 de fato reproduz a dinâmica do modelo MT.

Para verificação da normalidade foi aplicado o teste de *Shapiro-Wilk* (SHAPIRO; WILK, 1965; YAP; SIM, 2011). O teste de *Shapiro-Wilk* testa a hipótese nula de que a amostra $\{x_1, x_2, \dots, x_n\}$ é extraída de uma população com distribuição normal. A estatística de teste é dada por:

$$W = \frac{\left(\sum_{i=1}^n a_i x_{(i)}\right)^2}{\sum_{i=1}^n (x_i - \bar{x})^2}, \quad (3.11)$$

em que $x_{(i)}$ é o i -ésimo menor valor da amostra, e \bar{x} é a média amostral dos valores extraídos. Os coeficientes a_i são dados por:

$$(a_1, \dots, a_n) = \frac{m^T V^{-1}}{C}, \quad (3.12)$$

em que C é a norma do vetor dada por:

$$C = \|V^{-1}m\| = \sqrt{(m^T V^{-1} V^{-1} m)}, \quad (3.13)$$

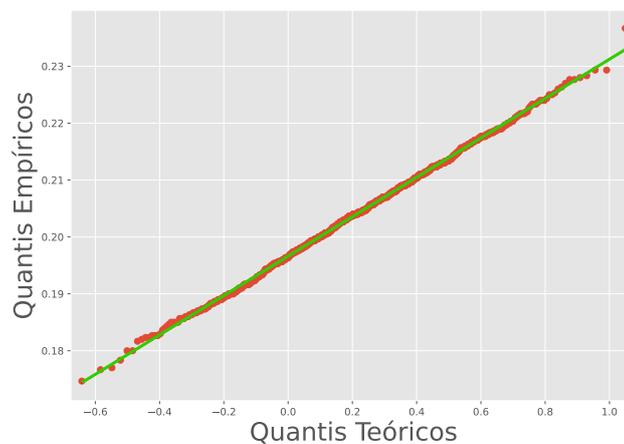
temos que $m = (m_1, \dots, m_n)^T$ é o vetor composto dos valores esperados das estatísticas de ordem de variáveis aleatórias independentes e identicamente distribuídas com distribuição normal padrão. Por último, V é a matriz de covariância dessas estatísticas de ordem normalmente distribuídas.

A hipótese nula desse teste é que a população é normalmente distribuída. Portanto, se o p -valor for menor que o nível alfa selecionado, a hipótese nula é rejeitada e há evidências de que os dados testados não seguem uma distribuição normal. Por outro lado, se o p -valor for maior

que o nível alfa selecionado, a hipótese nula não pode ser rejeitada. Desse modo, usamos o algoritmo em $M = 1000$ simulações do modelo MT com $\alpha = \lambda = 1,0$ para um grafo completo de tamanho $N = 3000$. As densidades finais dos suscetíveis de todas as M simulações foram computadas obtendo as respectivas estimativas $\hat{\mu} = 0,2036$ e $\hat{\sigma}^2 = 0,2735$ para média e a variância. Por fim, aplicou-se o teste de *Shapiro-Wilk*, em que foi obtido um *p-valor* de 0,8999. Dessa forma como o *p-valor* não é menor que 0,05, deixamos de rejeitar a hipótese nula. Também foi aplicado um teste gráfico por meio do *gráfico Q-Q* apresentado na Figura 25. A utilização de *gráficos Q-Q* para comparação de duas amostras pode ser interpretado como um método não-paramétrico para realizar essa tarefa (WILK; GNANADESIKAN, 1968). Essa abordagem consiste em plotar pontos (x, y) no gráfico, em que cada coordenada x corresponde a um dos quantis de uma das distribuições, e a cada coordenada y corresponde a um dos quantis da outra distribuição. Se os dados das duas distribuições que estão sendo comparadas são semelhantes, então os pontos do *gráfico Q-Q* terão um comportamento aproximadamente de $y = x$, isto é, formarão uma reta do tipo $y = x$.

Note que a linearidade dos pontos do gráfico da Figura 25 sugere que os dados são normalmente distribuídos, como esperado. Portanto, a partir desses resultados e do teste de *Shapiro-Wilk*, não temos evidências suficientes para dizer que a densidade final de ignorantes no modelo MT gerado pelo algoritmo desenvolvido na Seção 3.1.2 não segue uma distribuição normal com $\mu = 0,20$ e $\sigma^2 = 0,27$. Concluímos que o algoritmo apresentado na Seção 3.1.3 é um método confiável para a simulação do modelo MT.

Figura 25 – Gráfico Q-Q das densidades finais de indivíduos ignorantes do modelo Maki-Thompson.



Fonte: Produzido pelo autor.

3.2 CÁLCULO DAS CORRELAÇÕES E COVARIÂNCIAS

3.2.1 Métodos Discretos

Para o cálculo das correlações e covariâncias entre dois vértices vizinhos, faremos algumas modificações nos algoritmos apresentados na Seção 3.1. O método que foi desenvolvido nos retorna como resultado o valor médio das grandezas de interesse em um intervalo de tempo fixo. Para isso, iremos fixar dois vértices vizinhos i e j e, ao longo de um intervalo de tempo $[t_i, t_f]$ predefinido, os estados infecciosos de i e j em cada transição do sistema serão salvos nas listas:

$$\mathcal{I}^{(i)} = \{X_0^i, X_1^i, X_2^i, \dots, X_k^i\}, \quad (3.14)$$

$$\mathcal{I}^{(j)} = \{X_0^j, X_1^j, X_2^j, \dots, X_k^j\}, \quad (3.15)$$

em que X_k^i e X_k^j são variáveis aleatórias indicadoras que descrevem o estado infectado dos vértices i e j no intervalo de tempo entre a sua $(n-1)$ -ésima e a n -ésima transição. Como estamos interessados na correlação e covariância entre vizinhos que estão infectados, então:

$$X^i = \begin{cases} 1, & \text{se } i \text{ estiver infectado,} \\ 0, & \text{caso contrário.} \end{cases} \quad (3.16)$$

Também será definida uma terceira lista, dada por:

$$\mathcal{T} = \{\Delta t_0, \Delta t_1, \Delta t_2, \dots, \Delta t_k\}, \quad (3.17)$$

contendo os intervalos de tempo Δt_k entre $(n-1)$ -ésima e a n -ésima transição. Como as transições ocorrem em intervalos de tempo de comprimentos distintos, as expressões para as estimativas dos primeiros e segundos momentos são calculadas fazendo uma média ponderada entre a variável X_k^i e o intervalo de tempo Δt_k . Portanto temos, que as expressões necessárias para estimar as covariâncias e correlações desejadas são dadas por:

$$\hat{\mu}_{X^i} = \frac{\sum_{k=0}^N x_k^i \Delta t_k}{\sum_{n=0}^N \Delta t_k}, \quad (3.18)$$

$$\hat{\mu}_{X^i}'' = \frac{\sum_{k=0}^N (x_k^i)^2 \Delta t_k}{\sum_{k=0}^N \Delta t_k}, \quad (3.19)$$

$$\hat{\mu}_{X^i X^j} = \frac{\sum_{k=0}^N x_k^j x_k^i \Delta t_k}{\sum_{k=0}^N \Delta t_k}, \quad (3.20)$$

$$\hat{\sigma}_{X^i}^2 = \hat{\mu}_{X^i}'' - [\hat{\mu}_{X^i}]^2, \quad (3.21)$$

em que a covariância e a correlação, $\hat{\sigma}_{X^i X^j}$ e $\hat{\rho}_{X^i X^j}$, são dadas por:

$$\hat{\sigma}_{X^i X^j} = \hat{\mu}_{X^i X^j} - (\hat{\mu}_{X^i} \hat{\mu}_{X^j}), \quad (3.22)$$

$$\hat{\rho}_{X^i X^j} = \frac{\hat{\sigma}_{X^i X^j}}{\hat{\sigma}_{X^i} \hat{\sigma}_{X^j}}. \quad (3.23)$$

3.2.2 Cálculo da Correlação e Covariância: Modelo SIS

Passo 1 — Inicialização: Fixe $[t_i, t_f]$ e i, j vizinhos. Defina um grafo $G = (V, A)$ com N vértices. Fixe as taxas de cura e infecção de modo que δ_i e $\lambda_{(i,j)}$ sejam constantes para todo $i \in V$ e $(i, j) \in A$, respectivamente. Defina uma configuração inicial para o sistema, fazendo com que todo $i \in V$ assuma o estado infectado com uma probabilidade p . Após a inicialização, os vértices infectados e suscetíveis serão armazenados nas listas $\mathcal{E}^{(inf)} = \{\mathcal{E}_i^{(inf)}\}$ e $\mathcal{E}^{(sus)} = \{\mathcal{E}_i^{(sus)}\}$, respectivamente. Defina as listas $\mathcal{I}^{(i)}$, $\mathcal{I}^{(j)}$ e \mathcal{T} para armazenar os valores de X_k^i , X_k^j e Δt_k , respectivamente.

Passo 2: Selecione um dos vértices $i \in \mathcal{E}^{(inf)}$ com probabilidade:

$$\epsilon_i = \frac{\delta + n_i^S \lambda}{N_I \delta + N_S \lambda} \quad \forall i \in \mathcal{E}^{(inf)}, \quad (3.24)$$

em que n_i^S é a quantidade de vizinhos do vértice i no estado suscetível, N_I é a quantidade total de vértices infectados e N_S é a soma de todos vizinhos suscetíveis de todos os vértices infectados dada por:

$$N_S = \sum_{i \in \mathcal{E}^{(inf)}} n_i^S. \quad (3.25)$$

Passo 3: O vértice escolhido no **Passo 2** torna-se **suscetível** com probabilidade:

$$\mu_i = \frac{\delta}{\delta + n_i^S \lambda}, \quad (3.26)$$

em que n_i^S é a quantidade de vizinhos suscetíveis do vértice i escolhido. Caso contrário, com igual probabilidade, escolha um vértice $j \in \mathcal{E}^{(sus)}$ vizinho a i e o torne **infectado**. Por fim, atualize as listas $\mathcal{E}^{(sus)}$ e $\mathcal{E}^{(inf)}$, respectivamente.

Passo 4: Faça $t = t + \Delta t_k$, em que $\Delta t_k \sim \exp(N_I \delta + N_S \lambda)$.

Passo 5: Se $t \geq t_i$, então faça:

$$X_k^i = \begin{cases} 1, & \text{se } i \in \mathcal{E}^{(inf)}, \\ 0, & \text{caso contrário.} \end{cases} \quad (3.27)$$

$$X_k^j = \begin{cases} 1, & \text{se } j \in \mathcal{E}^{(inf)}, \\ 0, & \text{caso contrário.} \end{cases} \quad (3.28)$$

então, atualize as listas $\mathcal{I}^{(i)}$, $\mathcal{I}^{(j)}$ e \mathcal{T} com os valores de X_k^i , X_k^j e Δt_k , respectivamente.

Passo 6: Enquanto $t < t_f$, repita os passos **2**, **3**, **4** e **5**;

Passo 7: Quando $t \geq t_f$, use os valores das listas $\mathcal{I}^{(i)}$, $\mathcal{I}^{(j)}$ e \mathcal{T} nas Equações 3.23 e 3.22 para o cálculo da covariância e correlação entre i e j , respectivamente.

3.2.3 Cálculo da Correlação e Covariância: Modelo SIR

Passo 1 — Inicialização: Fixe $[t_i, t_f]$ e i, j vizinhos. Defina um grafo $G = (V, A)$ com N vértices. Fixe as taxas de cura e infecção de modo que δ_i e $\lambda_{(i,j)}$ sejam constantes para todo $i \in V$ e $(i, j) \in A$, respectivamente. Defina uma configuração inicial para o sistema sorteando um vértice da rede que começará no estado infectado. Após a inicialização, os vértices infectados e suscetíveis serão armazenados nas listas $\mathcal{E}^{(inf)} = \{i\}$ e $\mathcal{E}^{(sus)} = \{i\}$, respectivamente. Defina as listas $\mathcal{I}^{(i)}$, $\mathcal{I}^{(j)}$ e \mathcal{T} para armazenar os valores de X_k^i , X_k^j e Δt_k , respectivamente.

Passo 2: Selecione um dos vértices $i \in \mathcal{E}^{(inf)}$ com probabilidade:

$$\epsilon_i = \frac{\delta + n_i^S \lambda}{N_I \delta + N_S \lambda} \quad \forall i \in \mathcal{E}^{(inf)}, \quad (3.29)$$

em que n_i^S é a quantidade de vizinhos do vértice i no estado suscetível, N_I é a quantidade total de vértices infectados e N_S é a soma de todos os vizinhos suscetíveis de todos os vértices infectados dada por:

$$N_S = \sum_{i \in \mathcal{E}^{(inf)}} n_i^S. \quad (3.30)$$

Passo 3: O vértice escolhido no **Passo 2** torna-se **Recuperado** com probabilidade:

$$\mu_i = \frac{\delta}{\delta + n_i^S \lambda}, \quad (3.31)$$

em que n_i^S é a quantidade de vizinhos suscetíveis do vértice i escolhido. Caso contrário, com igual probabilidade, escolha um vértice $j \in \mathcal{E}^{(sus)}$ vizinho a i e o torne **infectado**. Por fim, atualize as listas $\mathcal{E}^{(sus)}$ e $\mathcal{E}^{(inf)}$, respectivamente.

Passo 4: Faça $t = t + \Delta t_k$, em que $\Delta t_k \sim \exp(N_I \delta + N_S \lambda)$.

Passo 5: Se $t \geq t_i$, então faça:

$$X_k^i = \begin{cases} 1, & \text{se } i \in \mathcal{E}^{(inf)}, \\ 0, & \text{caso contrário.} \end{cases} \quad (3.32)$$

$$X_k^j = \begin{cases} 1, & \text{se } j \in \mathcal{E}^{(inf)}, \\ 0, & \text{caso contrário.} \end{cases} \quad (3.33)$$

Então, atualize as listas $\mathcal{I}^{(i)}$, $\mathcal{I}^{(j)}$ e \mathcal{T} com os valores de X_k^i , X_k^j e Δt_k , respectivamente.

Passo 6: Enquanto $t < t_f$, repita os passos **2**, **3**, **4** e **5**;

Passo 7: Quando $t \geq t_f$, use os valores das listas $\mathcal{I}^{(i)}$, $\mathcal{I}^{(j)}$ e \mathcal{T} nas Equações 3.22 e 3.23 para o cálculo da covariância e correlação entre i e j , respectivamente.

3.2.4 Cálculo da Correlação e Covariância: Modelo MT

Passo 1 — Inicialização: Fixe $[t_i, t_f]$ e i, j vizinhos. Defina um grafo $G = (V, A)$ com N vértices. Fixe as taxas de neutralização e propagação de modo que α_i e $\lambda_{(i,j)}$ sejam constantes para todo $i \in V$ e $(i, j) \in A$, respectivamente. Defina uma configuração inicial para o sistema, sorteando um vértice da rede que começará no estado informante. Após a inicialização, os vértices informantes e ignorantes serão armazenados nas listas $\mathcal{E}^{(inf)} = \{\mathcal{E}_i^{(inf)}\}$ e $\mathcal{E}^{(ig)} = \{\mathcal{E}_i^{(ig)}\}$, respectivamente. Defina as listas $\mathcal{I}^{(i)}$, $\mathcal{I}^{(j)}$ e \mathcal{T} para armazenar os valores de X_k^i , X_k^j e Δt_k , respectivamente.

Passo 2: Selecione um dos vértices $i \in \mathcal{E}^{(inf)}$ com probabilidade:

$$\epsilon_i = \frac{(k_i - n_i^{Ig})\alpha + n_i^{Ig}\lambda}{N_{IN}\alpha + N_{Ig}\lambda} \quad \forall i \in \mathcal{E}^{(inf)}, \quad (3.34)$$

em que n_i^{Ig} é a quantidade de vizinhos do vértice i no estado ignorante, k_i o grau do vértice i , N_{Ig} é a soma dos vizinhos ignorantes de todos os vértices informantes. E, por fim, N_{IN} é

a soma dos não ignorantes vizinhos de todos os vértices informantes, dado por:

$$N_{Ig} = \sum_{i \in \mathcal{E}^{(inf)}} n_i^{Ig}, \quad (3.35)$$

$$N_{IN} = \sum_{i \in \mathcal{E}^{(inf)}} (k_i - n_i^{Ig}). \quad (3.36)$$

Passo 3: O vértice escolhido no **Passo 2** torna-se **neutro** com probabilidade:

$$\mu_i = \frac{(k_i - n_i^{Ig})\alpha}{(k_i - n_i^{Ig})\alpha + n_i^{Ig}\lambda}, \quad (3.37)$$

em que n_i^{Ig} é a quantidade de vizinhos do vértice i no estado ignorante e k_i é o grau do vértice i (ver Apêndice B). Caso contrário, com igual probabilidade, escolha um vértice $j \in \mathcal{E}^{(ig)}$ vizinho a i e o torne **informante**. Por fim, atualize as listas $\mathcal{E}^{(ig)}$ e $\mathcal{E}^{(inf)}$, respectivamente.

Passo 4: Faça $t = t + \Delta t$, em que $\Delta t \sim \exp(N_{IN}\alpha + N_{Ig}\lambda)$.

Passo 5: Se $t \geq t_i$, então faça:

$$X_n^i = \begin{cases} 1, & \text{se } i \in \mathcal{E}^{(inf)}, \\ 0, & \text{caso contrário.} \end{cases} \quad (3.38)$$

$$X_n^j = \begin{cases} 1, & \text{se } j \in \mathcal{E}^{(inf)}, \\ 0, & \text{caso contrário.} \end{cases} \quad (3.39)$$

então, atualize as listas $\mathcal{I}^{(i)}$, $\mathcal{I}^{(j)}$ e \mathcal{T} com os valores de X_k^i , X_k^j e Δt_k respectivamente.

Passo 6: Enquanto $t < t_f$, repita os passos **2**, **3**, **4** e **5**;

Passo 7: Quando $t \geq t_f$, use os valores das listas $\mathcal{I}^{(i)}$, $\mathcal{I}^{(j)}$ e \mathcal{T} nas Equações 3.22 e 3.23 para o cálculo da covariância e correlação entre i e j , respectivamente.

3.2.5 Método Instantâneo

Tendo em vista que nos grafos k -regulares as covariâncias e correlações entre vértices adjacentes infectados são as mesmas, pois tais vértices possuem quantidade iguais de vizinhos, um outro método para aferição dessas grandezas foi desenvolvido. Para isso, serão feitas pequenas modificações nos algoritmos apresentados nas Seções 3.1.1, 3.1.2 e 3.1.3. Dado que a variável aleatória que descreve os estados dos vértices segue uma distribuição de *Bernoulli*, temos que a covariância entre dois vértices infectados pode ser expressa como:

$$COV(X_i(t), X_j(t)) = \mathbb{P}(X_i(t) = 1, X_j(t) = 1) - \left[\mathbb{P}(X_i(t) = 1)\mathbb{P}(X_j(t) = 1) \right]. \quad (3.40)$$

Note que $\mathbb{P}(X_i(t) = 1, X_j(t) = 1)$ pode ser interpretada como a probabilidade de se encontrar dois vértices infectados que estão interligados pela mesma aresta. Por outro lado $\mathbb{P}(X_i(t) = 1)$ e $\mathbb{P}(X_j(t) = 1)$ é a probabilidade de se encontrar um vértice infectado. Desse modo, note que $\mathbb{P}(X_i(t) = 1) = \mathbb{P}(X_j(t) = 1)$, portanto, a Equação 3.40 pode ser escrita na forma:

$$COV(X_i(t), X_j(t)) = \mathbb{P}(X_A(t) = 1) - \left[\mathbb{P}(X_V(t) = 1) \right]^2, \quad (3.41)$$

em que $\{X_A(t) = 1\}$ é o evento: *encontrar uma aresta infectada* e $\{X_V(t) = 1\}$ é o evento: *encontrar um vértice infectado*. Nesse ponto, deduziremos expressões para $\mathbb{P}(X_A(t) = 1)$ e $\mathbb{P}(X_V(t) = 1)$, respectivamente. Sabendo que a quantidade total de arestas para um grafo k -regular é dada pela Equação 1.4, portanto, temos que:

$$\mathbb{P}(X_A(t) = 1) = \frac{2N_A(t)}{Nk}, \quad (3.42)$$

$$\mathbb{P}(X_V(t) = 1) = \frac{N_i(t)}{N}, \quad (3.43)$$

em que $N_A(t)$ é o número de arestas infectadas no tempo t e $N_i(t)$ é a quantidade total de vértices infectados no tempo t , k é o grau do grafo e N é a quantidade total de vértices. Portanto, substituindo as Expressões 3.42 e 3.43 na Equação 3.41, obtemos:

$$COV(X_i(t), X_j(t)) = \frac{2N_A(t)}{Nk} - \left[\frac{N_i(t)}{N} \right]^2. \quad (3.44)$$

3.2.6 Método Instantâneo — Cálculo da Covariância: Modelo SIS

Passo 1 — Inicialização: Fixe t_f e i, j vizinhos. Defina um grafo $G = (V, A)$ k -regular com N vértices. Fixe as taxas de cura e infecção de modo que δ_i e $\lambda_{(i,j)}$ sejam constantes para todo $i \in V$ e $(i, j) \in A$, respectivamente. Defina uma configuração inicial para o sistema, fazendo com que todo $i \in V$ assumo o estado infectado com uma probabilidade p . Após a inicialização, os vértices infectados e suscetíveis serão armazenados nas listas $\mathcal{E}^{(inf)} = \{\mathcal{E}_i^{(inf)}\}$ e $\mathcal{E}^{(sus)} = \{\mathcal{E}_i^{(sus)}\}$, respectivamente. Defina as listas COV_t , T_{tempo} , que armazenarão os valores das covariâncias e o tempo decorrido.

Passo 2: Selecione um dos vértices $i \in \mathcal{E}^{(inf)}$ com probabilidade:

$$\epsilon_i = \frac{\delta + n_i^S \lambda}{N_I \delta + N_S \lambda} \quad \forall i \in \mathcal{E}^{(inf)}, \quad (3.45)$$

em que n_i^S é a quantidade de vizinhos do vértice i no estado suscetível, N_I é a quantidade total de vértices infectados e N_S é a soma de todos os vizinhos suscetíveis de todos os vértices infectados dada por:

$$N_S = \sum_{i \in \mathcal{E}^{(inf)}} n_i^S. \quad (3.46)$$

Passo 3: O vértice escolhido no **Passo 2** torna-se **suscetível** com probabilidade:

$$\mu_i = \frac{\delta}{\delta + n_i^S \lambda}, \quad (3.47)$$

em que n_i^S é a quantidade de vizinhos suscetíveis do vértice i escolhido. Caso contrário, com igual probabilidade, escolha um vértice $j \in \mathcal{E}^{(sus)}$ vizinho a i e o torne **infectado**. Por fim, atualize as listas $\mathcal{E}^{(sus)}$ e $\mathcal{E}^{(inf)}$, respectivamente.

Passo 4: Faça $t = t + \Delta t$, em que $\Delta t \sim \exp(N_I \delta + N_S \lambda)$.

Passo 5: Contabilize as grandezas $N_A(t)$ e $N_i(t)$ e, pela Equação 3.44, calcule a covariância. Salve o resultado na posição n da lista COV_t e t na posição n da lista T_{tempo} . Finalmente, faça $n = n + 1$.

Passo 6: Então os passos **2**, **3**, **4** e **5** são repetidos enquanto houver vértices infectados na rede ou $t < t_f$.

3.2.7 Método Instantâneo — Cálculo da Covariância: Modelo SIR

Passo 1 — Inicialização: Fixe t_f e i, j vizinhos. Defina um grafo $G = (V, A)$ k -regular com N vértices. Fixe as taxas de cura e infecção de modo que δ_i e $\lambda_{(i,j)}$ sejam constantes para todo $i \in V$ e $(i, j) \in A$, respectivamente. Defina uma configuração inicial para o sistema sorteando um vértice da rede que começará no estado infectado. Após a inicialização, os vértices infectados e suscetíveis serão armazenados nas listas $\mathcal{E}^{(inf)} = \{\mathcal{E}_i^{(inf)}\}$ e $\mathcal{E}^{(sus)} = \{\mathcal{E}_i^{(sus)}\}$, respectivamente. Defina as listas COV_t , T_{tempo} , que armazenarão os valores das covariâncias e o tempo decorrido.

Passo 2: Selecione um dos vértices $i \in \mathcal{E}^{(inf)}$ com probabilidade:

$$\epsilon_i = \frac{\delta + n_i^S \lambda}{N_I \delta + N_S \lambda} \quad \forall i \in \mathcal{E}^{(inf)}, \quad (3.48)$$

em que n_i^S é a quantidade de vizinhos do vértice i no estado suscetível, N_I é a quantidade total de vértices infectados e N_S é a soma de todos os vizinhos suscetíveis de todos os vértices infectados dada por:

$$N_S = \sum_{i \in \mathcal{E}^{(inf)}} n_i^S. \quad (3.49)$$

Passo 3: O vértice escolhido no **Passo 2** torna-se **recuperado** com probabilidade:

$$\mu_i = \frac{\delta}{\delta + n_i^S \lambda}, \quad (3.50)$$

em que n_S é a quantidade de vizinhos suscetíveis do vértice i escolhido. Caso contrário, com igual probabilidade, escolha um vértice $j \in \mathcal{E}^{(sus)}$ vizinho a i e o torne **infectado**. Por fim, atualize as listas $\mathcal{E}^{(sus)}$ e $\mathcal{E}^{(inf)}$, respectivamente.

Passo 4: Faça $t = t + \Delta t$, em que $\Delta t \sim \exp(N_I \delta + N_S \lambda)$.

Passo 5: Contabilize as grandezas $N_A(t)$ e $N_i(t)$ e, pela Equação 3.44, calcule a covariância. Salve o resultado na posição n da listas COV_t e t na posição n da lista T_{tempo} . Finalmente, faça $n = n + 1$.

Passo 6: Então os passos **2**, **3**, **4** e **5** são repetidos enquanto houver vértices infectados na rede ou $t < t_f$.

3.2.8 Método Instantâneo — Cálculo da Covariância: Modelo MT

Passo 1 — Inicialização: Fixe t_f e i, j vizinhos. Defina um grafo $G = (V, A)$ k -regular com N vértices. Fixe as taxas de cura e infecção de modo que α_i e $\lambda_{(i,j)}$ sejam constantes para todo $i \in V$ e $(i, j) \in A$, respectivamente. Defina uma configuração inicial para o sistema sorteando um vértice da rede que começará no estado informante. Após a inicialização, os vértices informantes e ignorantes serão armazenados nas listas $\mathcal{E}^{(inf)} = \{\mathcal{E}_i^{(inf)}\}$ e $\mathcal{E}^{(ig)} = \{\mathcal{E}_i^{(ig)}\}$, respectivamente. Defina as listas COV_t , T_{tempo} , que armazenarão os valores das covariâncias e o tempo decorrido.

Passo 2: Selecione um dos vértices $i \in \mathcal{E}^{(inf)}$ com probabilidade:

$$\epsilon_i = \frac{(k_i - n_i^{Ig})\alpha + n_i^{Ig}\lambda}{N_{IN}\alpha + N_{Ig}\lambda} \quad \forall i \in \mathcal{E}^{(inf)}, \quad (3.51)$$

em que n_i^{Ig} é a quantidade de vizinhos do vértice i no estado ignorante, k_i o grau do vértice i , N_{Ig} é a soma dos vizinhos ignorantes de todos os vértices informantes. E, por fim, N_{IN} é

a soma dos não ignorantes vizinhos de todos os vértices informantes, dado por:

$$N_{Ig} = \sum_{i \in \mathcal{E}^{(inf)}} n_i^{Ig}, \quad (3.52)$$

$$N_{IN} = \sum_{i \in \mathcal{E}^{(inf)}} (k_i - n_i^{Ig}). \quad (3.53)$$

Passo 3: O vértice escolhido no **Passo 2** torna-se **neutro** com probabilidade:

$$\mu_i = \frac{(k_i - n_i^{Ig})\alpha}{(k_i - n_i^{Ig})\alpha + n_i^{Ig}\lambda}, \quad (3.54)$$

em que n_i^{Ig} é a quantidade de vizinhos do vértice i no estado ignorante e k_i é o grau do vértice i (ver Apêndice B). Caso contrário, com igual probabilidade, escolha um vértice $j \in \mathcal{E}^{(ig)}$ vizinho a i e o torne **informante**. Por fim, atualize as listas $\mathcal{E}^{(ig)}$ e $\mathcal{E}^{(inf)}$, respectivamente.

Passo 4: Faça $t = t + \Delta t$, em que $\Delta t \sim \exp(N_{IN}\alpha + N_{Ig}\lambda)$.

Passo 5: Contabilize as grandezas $N_A(t)$ e $N_i(t)$ e, pela Equação 3.44 calcule a covariância. Salve o resultado na posição n da listas COV_t e t na posição n da lista T_{tempo} . Finalmente, faça $n = n + 1$.

Passo 6: Então os passos **2**, **3**, **4** e **5** são repetidos enquanto houver vértices infectados na rede ou $t < t_f$.

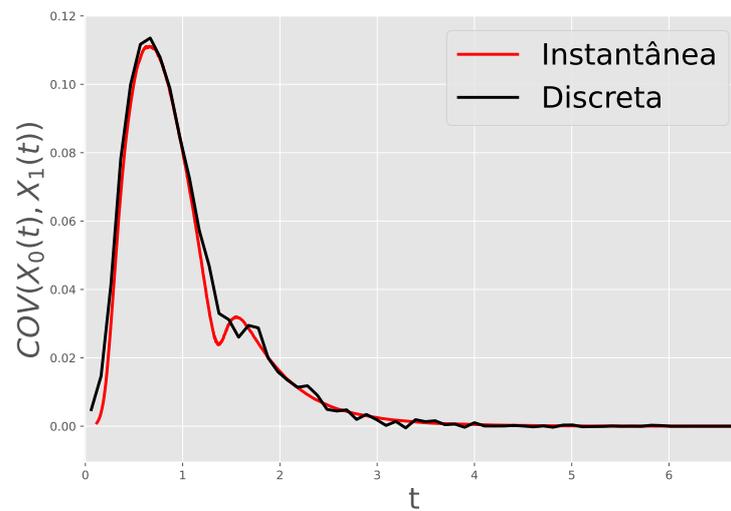
3.2.9 Comparação Entre os Métodos Discretos e Instantâneos

De modo geral, desenvolvemos dois métodos para o cálculo da covariância. O método instantâneo nos retorna a covariância em cada instante de tempo; já o método discreto nos retorna o valor médio da covariância em um intervalo de tempo $[t_i, t_f]$ fixo. Por outro lado, o método discreto pode ser adaptado para realizar os cálculos de forma equivalente ao método instantâneo, basta fazer com que o intervalo $[t_i, t_f]$ seja dividido em pequenos subintervalos, possibilitando sua aplicação em cada uma das partes menores. Note que a forma instantânea só é válida para grafos k -regulares, enquanto o método discreto é mais geral e pode ser aplicado em qualquer família de grafos, pois não foi concebido a partir considerações sobre a topologia do grafo.

Por meio do método instantâneo, estimamos os valores das covariâncias em função do tempo para um grafo k -regular com $N = 300$ vértices e $k = 22$. Para isso, foram feitas simulações de Monte-Carlo com 1000 amostras no mesmo intervalo de tempo $[0.01, 6.778]$. O resultado é apresentado na curva vermelha da Figura 26. O procedimento foi repetido para

o método discreto, sendo o intervalo $[0.01, 6.778]$ subdividido em pequenos subintervalos de largura $\Delta t = 0.101$. Fixamos os vizinhos $i = 0$ e $j = 0$ e, por fim, foram realizadas simulações de Monte-Carlo com 2×10^4 amostras para cada subintervalo. Os resultados estão dispostos na Figura 26 (curva preta). Note que o comportamento da covariância foi preservado indicando que os métodos são equivalentes.

Figura 26 – Comparação entre os métodos discretos e contínuos na aferição das covariâncias no modelo SIR.

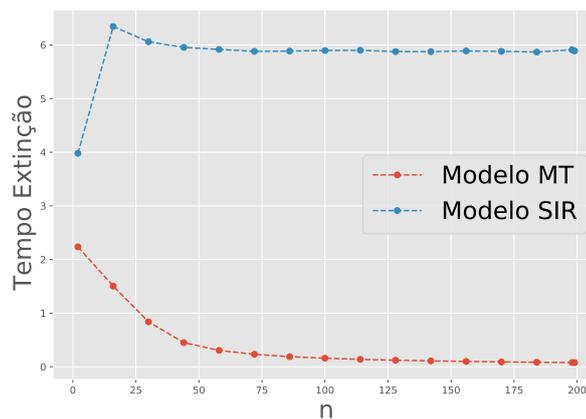


Fonte: Produzido pelo autor.

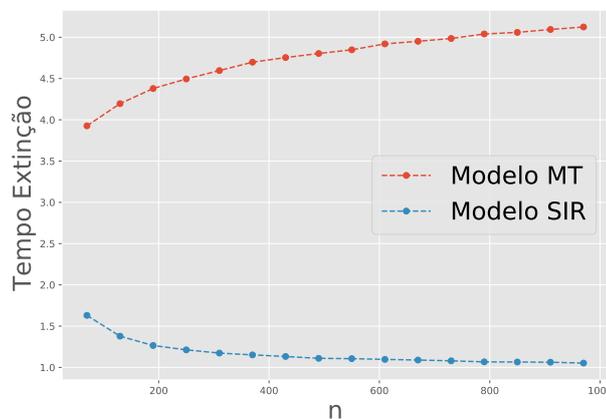
4 RESULTADOS E DISCUSSÕES

A partir dos algoritmos desenvolvidos no Capítulo 3, foram realizadas simulações de Monte Carlo para estudar o comportamento da correlação e covariância em função dos parâmetros de cada um dos grafos escolhidos. Inicialmente, foram estimados os tempos médios de extinção dos processos SIR e MT para cada grafo por meio de simulações de Monte Carlo com 2×10^4 amostras. Sendo $N = 200$ o número de vértices do grafo k -regular com k assumindo valores de $k = 2, 16, 30, \dots, N - 1$, respectivamente. Para o grafo estrela $n \times n$ foram usados os valores, $n = 70, 130, 190, \dots, 970$, respectivamente. Os resultados são apresentados nas Figuras 27a e 27b.

Figura 27 – Resultados das estimativas dos tempos de extinção do processo SIR e MT.



(a) Tempos de extinção dos processos SIR e MT definidos no grafo k -regular.



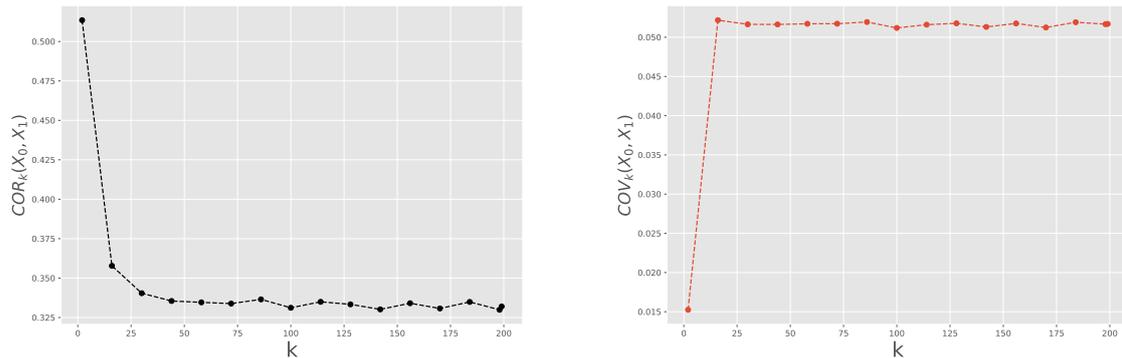
(b) Tempos de extinção dos processos SIR e MT definidos no grafo estrela $n \times n$.

Fonte: Produzido pelo autor.

4.1 CORRELAÇÕES E COVARIÂNCIAS

4.1.1 Modelo SIR

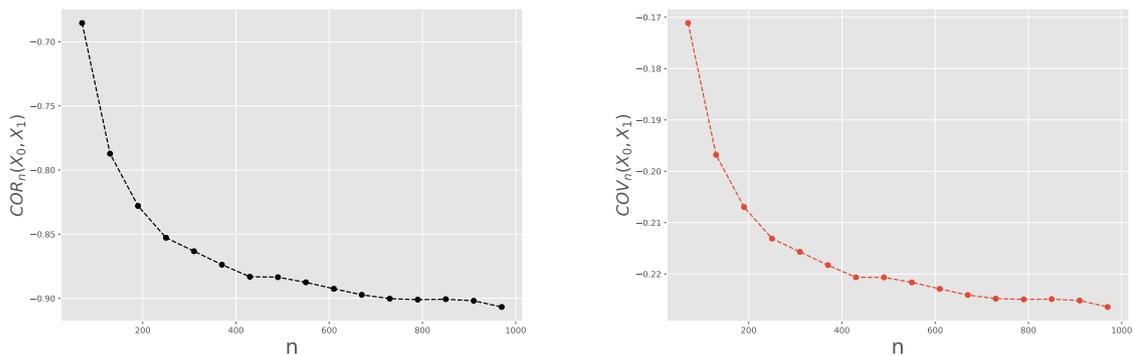
Figura 28 – Estimativas das covariâncias e correlações médias entre os vértices 0 e 1 no grafo k -regular para o modelo SIR com $\alpha = 1.0$ e $\lambda = 20/k$.



(a) Correlações médias em função de k para o grafo k -regular. (b) Covariâncias médias em função de k para o grafo k -regular.

Fonte: Produzido pelo autor.

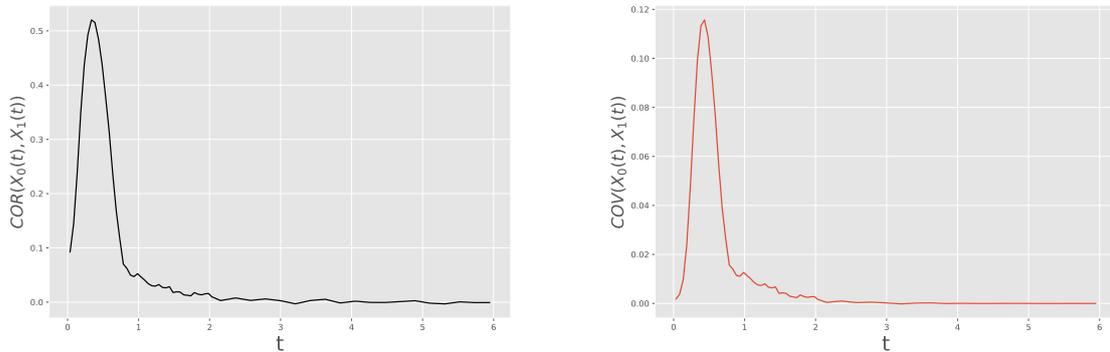
Figura 29 – Estimativas das covariâncias e correlações médias entre os vértices 0 e 1 no grafo estrela $n \times n$ para o modelo SIR com $\delta = 1$ e $\lambda = 20.0/n$.



(a) Correlações médias em função de n para o grafo estrela $n \times n$. (b) Covariâncias médias em função de n para o grafo estrela $n \times n$.

Fonte: Produzido pelo autor.

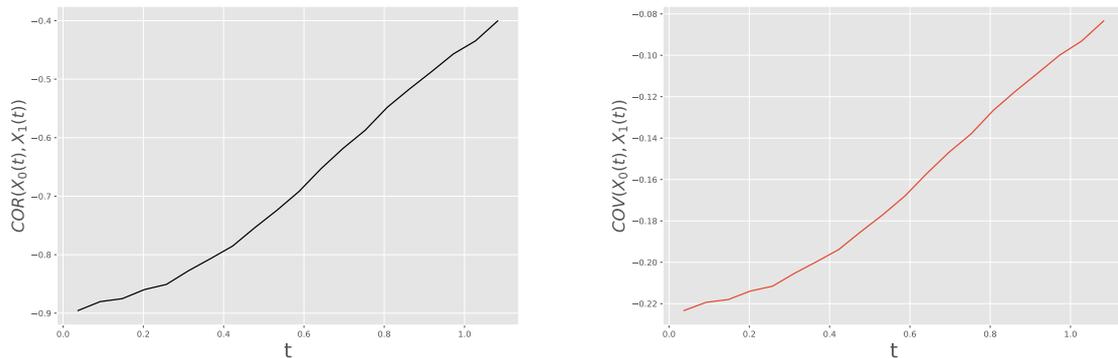
Figura 30 – Estimativas das covariâncias e correlações em função do tempo entre os vértices 0 e 1 no grafo k -regular com $k = 30$, $\delta = 1$, $\lambda = 20/30$ para o modelo SIR.



(a) Correlações em função do tempo para o grafo k -regular com $k = 30$. (b) Covariâncias em função do tempo para o grafo k -regular com $k = 30$.

Fonte: Produzido pelo autor.

Figura 31 – Estimativas das covariâncias e correlações em função do tempo entre os vértices 0 e 1 no grafo k -regular com $k = 30$, $\delta = 1$, $\lambda = 20/610$ para o modelo SIR.



(a) Correlações em função do tempo para o grafo estrela 610×610 . (b) Covariância em função do tempo para o grafo estrela 610×610 .

Fonte: Produzido pelo autor.

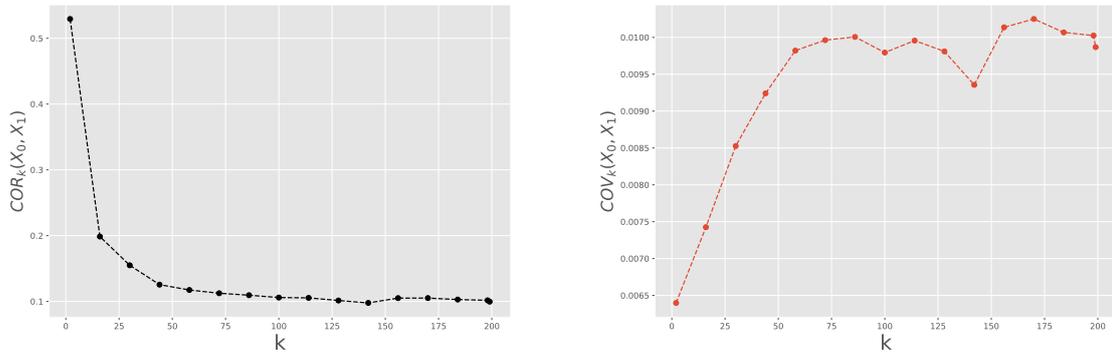
De posse dos tempos de extinção do processo, foram realizadas simulações de Monte Carlo para estimar a covariância e correlação média entre o vértice $i = 0$ e seu vizinho $j = 1$, no intervalo de tempo $[0.01, t_{ex}]$, para cada valor de n e k , respectivamente. Os resultados para o modelo SIR no grafo k -regular são exibidos na Figura 28. Notemos que, apesar do método apresentado por (CATOR; MIEGHEM, 2014) para mostrar que a correlação e covariância entre os estados de vértices adjacentes não ser aplicável ao modelo SIR, os autores sugeriram que tais grandezas apresentariam o comportamento visto na Figura 28. Note que, a partir de $k \geq 30$, as correlações não apresentam grandes mudanças, tendo um comportamento oscilatório em torno do valor $3,32 \times 10^{-1}$, o mesmo acontece para as covariâncias, que oscilam em torno de

$5,17 \times 10^{-2}$. Observe que os valores das covariâncias e correlações oscilam em torno do valor estimado para o grafo completo. Nesse caso, podemos interpretar que o processo definido no grafo k -regular a partir de $k \geq 30$ apresenta o mesmo comportamento do processo definido no grafo completo. Para o grafo estrela $n \times n$, os resultados são exibidos na Figura 29. Note que as correlações e covariâncias apresentam um comportamento monótono decrescente negativo, indicando que o Lema 2 não é satisfeito. Portanto, as trajetórias do processo obtidas a partir da aproximação de Campo Médio não são boas estimativas para as trajetórias do processo original definido nesse tipo de grafo.

Para o cálculo das covariâncias e correlações em função do tempo, foram fixados $k = 30$ para o grafo k -regular e $n = 610$ para o grafo estrela $n \times n$. Com isso, os intervalos de tempo $[0,01, t_{ext}]$ relativos ao k e n escolhidos foram subdivididos em subintervalos menores, em que, por meio de simulação Monte Carlo, as correlações e covariâncias foram calculadas em cada uma dessas partes. Para o grafo 30-regular, os resultados são apresentados na Figura 30. Como já esperado, os valores assumidos pelas covariâncias em todo o intervalo são positivos, tendendo a zero para $t \geq 2$. Nesse caso, temos que a desigualdade apresentada no Lema 2 é respeitada em todo t . O processo foi repetido para o grafo estrela 610×610 , que apresentou um comportamento suave e crescente negativo, como pode ser visto na Figura 31. Sendo assim, nota-se que a condição imposta pelo Lema 2, nesse caso, não é satisfeita para todo t .

4.1.2 Modelo MT

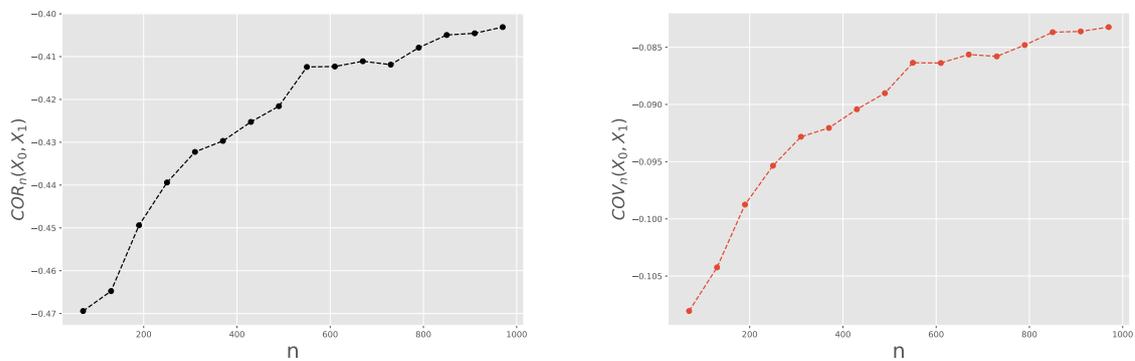
Figura 32 – Estimativas das covariâncias e correlações médias entre os vértices 0 e 1 no grafo k -regular para o modelo MT com $\alpha = \lambda = 1.0$.



(a) Correlações médias em função de k para o grafo k -regular. (b) Covariâncias médias em função de k para o grafo k -regular.

Fonte: Produzido pelo autor.

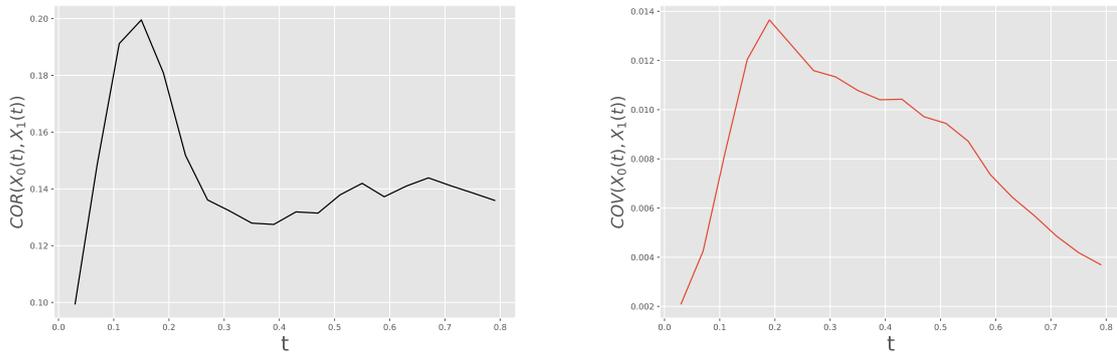
Figura 33 – Estimativas das covariâncias e correlações médias entre os vértices 0 e 1 no grafo estrela $n \times n$ para o modelo MT com $\alpha = \lambda = 1$.



(a) Correlações médias em função de n para o grafo estrela $n \times n$. (b) Covariâncias médias em função de n para o grafo estrela $n \times n$.

Fonte: Produzido pelo autor.

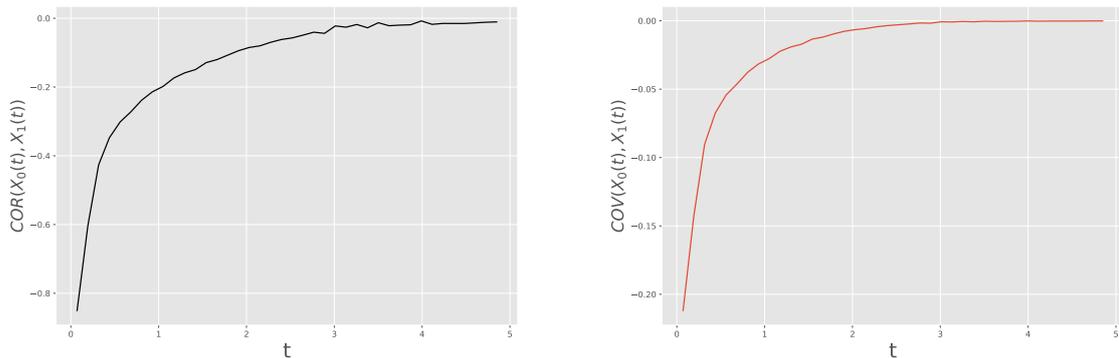
Figura 34 – Estimativas das covariâncias e correlações em função do tempo entre os vértices 0 e 1 no grafo k -regular com $k = 30$, $\delta = \lambda = 1,0$ para o modelo MT.



(a) Correlações em função do tempo para o grafo k -regular com $k = 30$. (b) Covariâncias em função do tempo para o grafo k -regular com $k = 30$.

Fonte: Produzido pelo autor.

Figura 35 – Estimativas das covariâncias e correlações médias entre os vértices 0 e 1 no grafo estrela 610×610 com $\alpha = \lambda = 1,0$ para o modelo MT.



(a) Correlação entre os vértices 0 e 1 para o modelo Maki-Tompson no grafo estrela conectada $n \times n$. (b) Covariância entre os vértices 0 e 1 para o modelo Maki-Tompson no grafo estrela conectada $n \times n$.

Fonte: Produzido pelo autor.

Os resultados para o modelo MT definido no grafo k -regular são apresentados na Figura 32, tais resultados foram obtidos aplicando a mesma metodologia vista na Seção 4.1.1. Note que as covariâncias e correlações permaneceram positivas para todos os valores de k , satisfazendo a equação apresentada no Lema 2. Veja também que a correlação exibe poucas mudanças para $k > 50$, sendo que para $k \geq 100$ a mesma oscila em torno do valor de 9×10^{-2} (grafo completo). Esse resultado indica que não há grandes variações no comportamento do processo para todo $k \geq 100$. Para o grafo estrela $n \times n$, por sua vez, o comportamento das correlações e covariâncias apresentadas se diferem do grafo completo. Neste caso, temos que a curva é monótona crescente negativa, indicando que o Lema 2 para esse caso não é satisfeito, ver

Figura 33.

Para esse modelo, também foram calculadas as correlações e covariâncias em função do tempo. Sendo o processo definido no grafo 30-regular, cujos resultados podem ser vistos na Figura 34, o comportamento da correlação se assemelha com o modelo SIR, apresentando um pico e estabilizando em um valor constante. Contudo, no caso do modelo MT, a correlação estabiliza em torno de 0,14, diferente do SIR, que converge para zero. A covariância, por sua vez, apresenta um comportamento completamente distinto em comparação ao modelo SIR, mas continua sendo positiva em todo o intervalo. Por fim, podemos concluir que, neste caso, a desigualdade apresentada no Lema 2 é respeitada em todo t . Finalmente, repetimos o procedimento para o modelo definido no grafo estrela $n \times n$, (Figura 35). Os valores das covariâncias e correlações inicialmente são negativos, convergindo para zero de forma suave a partir de $t \geq 3$.

5 CONCLUSÕES

Esta dissertação teve como foco estudar o comportamento da covariância e correlação entre vértices adjacentes nos processos SIS, SIR e MT definidos em grafos finitos por meio de simulações computacionais e resultados analíticos. Tendo como objetivo geral contribuir com as técnicas atuais de modelagem dos fenômenos de transmissão. Para isso, foram revisados os argumentos teóricos que normalmente são empregados para aproximar as trajetórias dos processos originais por meio de Campo Médio. Tais argumentos envolvem a suposição da existência de independência entre os estados de diferentes vértices do grafo. Apesar dessa suposição não ser verdadeira, ela pode ser considerada uma boa aproximação se os estados dos diferentes vértices forem positivamente correlacionados, sendo um limite superior das trajetórias do processo estocástico envolvido. Tendo em vista tais ideias, no Capítulo 4 foram estudados de maneira detalhada os argumentos apresentados por (CATOR; MIEGHEM, 2014), em que é afirmado, para os modelos SIS e SIR, que as respectivas variáveis aleatórias apresentam correlações positivas. Nesse mesmo capítulo também apresentamos de maneira detalhada o contra exemplo dado por (RODRIGUEZ; ROLDÁN-CORREA; VALENCIA, 2018), em que é sugerido que os argumentos usados na prova para o modelo SIR apresentados por (CATOR; MIEGHEM, 2014) não estão corretos, de modo que nos leva a concluir que as mesmas afirmações também não são verdadeiras para o modelo MT, devido as semelhanças compartilhadas com o SIR.

Nossa contribuição com este trabalho pode ser dividida em dois pontos. Primeiro, o desenvolvimento de métodos e algoritmos computacionais para estimar a correlação e covariância em função do tempo entre dois vértices adjacentes definidos em qualquer grafo finito nos modelos SIS, SIR e MT. Com isso, estudamos o comportamento médio de tais grandezas em um k -regular e no grafo estrela $n \times n$. Por fim, apresentamos diversos resultados numéricos por meio de gráficos, que corroboram com as alegações teóricas trazidas por (RODRIGUEZ; ROLDÁN-CORREA; VALENCIA, 2018), sugerindo que a construção apresentada por (CATOR; MIEGHEM, 2014) não está correta para o modelo SIR. Esses resultados também foram verificados para o modelo MT, confirmando a hipótese de que a justificativa apresentada por (CATOR; MIEGHEM, 2014) também não é válida para esse modelo, devido ao comportamento similar ao SIR. Desse modo, conseguimos reunir um conjunto de métodos e novos resultados numéricos que nos leva a concluir que este problema ainda está em aberto e necessita de uma prova teórica.

Como trabalhos futuros, pretende-se repetir as simulações numéricas que apresentamos neste trabalho utilizando um quantitativo maior de amostras. Também pretendemos estender os estudos do comportamento das covariâncias e correlações para outras famílias de grafos, bem como estudar este problema de maneira teórica. Portanto, a execução desse trabalho nos levou a expandir nossos conhecimentos na área de processos de transmissão e trouxe novas pistas para uma demonstração teórica do problema.

REFERÊNCIAS

- CATOR, E.; MIEGHEM, P. V. Nodal infection in markovian susceptible-infected-susceptible and susceptible-infected-removed epidemics on networks are non-negatively correlated. *Physical Review E*, APS, v. 89, n. 5, p. 052802, 2014.
- COTA, W. Métodos de simulação de processos epidêmicos markovianos em redes complexas. *Researchgate*, 02 2017.
- DONNELLY, P. The correlation structure of epidemic models. *Mathematical biosciences*, Elsevier, v. 117, n. 1-2, p. 49–75, 1993.
- DURRETT, R.; DURRETT, R. *Essentials of stochastic processes*. [S.l.]: Springer, 1999. v. 1.
- EMERITO, A. d. P. Modelos matemáticos determinísticos em doenças. 1989.
- FARIAS, A. V. Um estudo da modelagem epidemiológica sir usando conceitos de derivadas de ordem inteira e fracionária. *Monografia, UFRG*, 2017.
- FONTES, L. R. G. *Notas em percolação*. [S.l.]: Instituto de Matemática Pura e Aplicada, 1996.
- FORTUIN, C. M.; KASTELEYN, P. W.; GINIBRE, J. Correlation inequalities on some partially ordered sets. *Communications in Mathematical Physics*, Springer, v. 22, n. 2, p. 89–103, 1971.
- GINIBRE, J. General formulation of griffiths' inequalities. *Communications in mathematical physics*, Springer, v. 16, n. 4, p. 310–328, 1970.
- GRIFFITHS, R. B.; HURST, C. A.; SHERMAN, S. Concavity of magnetization of an ising ferromagnet in a positive external field. *Journal of Mathematical Physics*, American Institute of Physics, v. 11, n. 3, p. 790–795, 1970.
- GRIMMETT, G. *Probability on Graphs: Random Processes on Graphs and Lattices*. 2. ed. [S.l.]: Cambridge University Press, 2018. (Institute of Mathematical Statistics Textbooks).
- HARRIS, T. E. A lower bound for the critical probability in a certain percolation process. In: CAMBRIDGE UNIVERSITY PRESS. *Mathematical Proceedings of the Cambridge Philosophical Society*. [S.l.], 1960. v. 56, n. 1, p. 13–20.
- HARRIS, T. E. A correlation inequality for markov processes in partially ordered state spaces. *The Annals of Probability*, JSTOR, p. 451–454, 1977.
- HARRIS, T. E. A correlation inequality for markov processes in partially ordered state spaces. *The Annals of Probability*, JSTOR, p. 451–454, 1977.
- HOLLANDER, W. T. F. D.; KEANE, M. Inequalities of fkg type. *Physica A: Statistical Mechanics and its Applications*, Elsevier, v. 138, n. 1-2, p. 167–182, 1986.
- HOLLEY, R. Remarks on the fkg inequalities. *Communications in Mathematical Physics*, Springer, v. 36, n. 3, p. 227–231, 1974.
- JUNIOR, F. P. de P. *Modelagem Matemática e Epidemiologia Com Aplicações*. 1. ed. [S.l.: s.n.], 2021. (Ensino de Ciência). ISBN 978-65-250-0689-5.

JURKIEWICZ, S. Grafos - uma introdução. *São Paulo: OBMEP*, 2009.

LEBENSZTAYN, E.; MACHADO, F. P.; RODRÍGUEZ, P. M. Limit theorems for a general stochastic rumour model. *SIAM Journal on Applied Mathematics*, SIAM, v. 71, n. 4, p. 1476–1486, 2011.

MAKI, D. P.; THOMPSON, M. *Mathematical models and applications* prentice-hall. *Englewood Cliffs (NJ)*, 1973.

NETTO, P. O. B.; JURKIEWICZ, S. *Grafos: introdução e prática*. [S.l.]: Editora Blucher, 2017.

PROCÓPIO-PR, C. C. A matemática nos modelos do tipo sir. *Jornal Eletrônico de Ensino e Pesquisa de Matemática*, v. 4, 2020. Disponível em: <<http://www.dma.uem.br/kit/jeepepa>>.

RODRIGUEZ, P. M.; ROLDÁN-CORREA, A.; VALENCIA, L. A. Comment on “nodal infection in markovian susceptible-infected-susceptible and susceptible-infected-removed epidemics on networks are non-negatively correlated”. *Physical Review E*, APS, v. 98, n. 2, p. 026301, 2018.

ROSS, S. *Probabilidade: um curso moderno com aplicações*. [S.l.]: Bookman Editora, 2009.

SHAPIRO, S.; WILK, M. An analysis of variance test for normality. *Biometrika*, v. 52, n. 3, p. 591–611, 1965.

WATSON, R. On the size of a rumour. *Stochastic processes and their applications*, Elsevier, v. 27, p. 141–149, 1987.

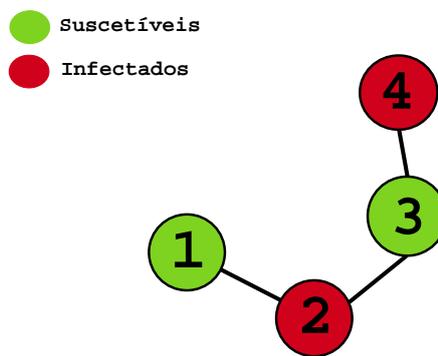
WILK, M. B.; GNANADESIKAN, R. Probability plotting methods for the analysis for the analysis of data. *Biometrika*, Oxford University Press, v. 55, n. 1, p. 1–17, 1968.

YAP, B. W.; SIM, C. H. Comparisons of various types of normality tests. *Journal of Statistical Computation and Simulation*, Taylor & Francis, v. 81, n. 12, p. 2141–2155, 2011.

APÊNDICE A – DEDUÇÕES DAS EXPRESSÕES 3.1 E 3.3

A expressão 3.1 pode ser justificada partindo da construção gráfica apresentada na Seção 2. Considere a realização do modelo SIS em um certo grafo $G = (V, A)$, com vértices e arestas $V = \{1, 2, 3, 4\}$ e $A = \{(1, 2), (2, 3), (3, 4)\}$, respectivamente (ver Figura 36). Sendo $X_1(t_0) = 0, X_2(t_0) = 1, X_3(t_0) = 0, X_4(t_0) = 1$, a configuração do processo no instante de tempo t_0 , definimos que 0 significa que o vértice está no estado suscetível e 1 infectado.

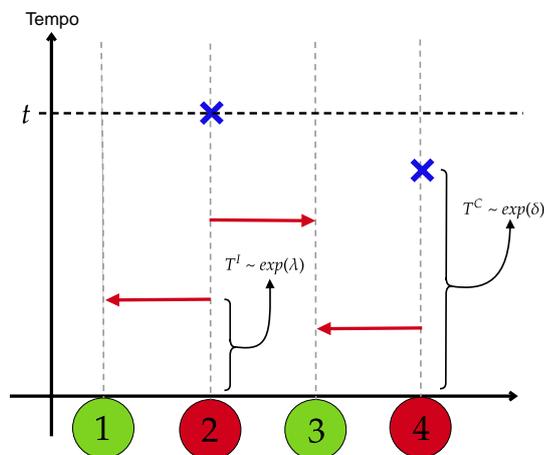
Figura 36 – Estados do vértice do grafo G em t_0



Fonte: Produzido pelo autor.

Note que as listas dos vértices suscetíveis e infectados são dadas por $\mathcal{E}^{(sus)} = \{1, 3\}$ e $\mathcal{E}^{(inf)} = \{2, 4\}$, respectivamente. O diagrama da realização do modelo até o tempo t é apresentado na Figura 37.

Figura 37 – Representação visual de uma realização do processo no tempo t .



Fonte: Produzido pelo autor.

Agora calcularemos a probabilidade de uma marca de cura ou seta de infecção ocorrer antes em $i = 2$ ao invés de $i = 4$. Podemos pensar que o tempo necessário para ocorrência de algum desses eventos deve ser menor em $i = 2$ ao invés de $i = 4$. Isso pode ser traduzido em termos de probabilidade da seguinte forma:

$$\epsilon_2 = \mathbb{P}\left(\min\{T_2^C, T_{21}^I, T_{23}^I\} < \min\{T_4^C, T_{43}^I\}\right). \quad (\text{A.1})$$

Sabendo que $T_2^C \sim \text{Exp}(\delta)$ é o tempo que o vértice $i = 2$ leva pra ser curado e $T_{21}^I, T_{23}^I \sim \text{Exp}(\lambda)$, o tempo que $i = 2$ tenta transmitir a infecção para os vértices $j = 1$ e $j = 3$, respectivamente, temos que encontrar a distribuição de $\min\{T_2^C, T_{21}^I, T_{23}^I\}$. Para isso, faremos uso do seguinte argumento: se o mínimo é maior que t , então todas as variáveis também necessitam ser maiores que t . Assim,

$$\mathbb{P}\left(\min\{T_2^C, T_{21}^I, T_{23}^I\} > t\right) = \mathbb{P}\left(T_2^C > t, T_{21}^I > t, T_{23}^I > t\right), \quad (\text{A.2})$$

por independência das variáveis, a expressão A.3 pode ser escrita como:

$$\begin{aligned} \mathbb{P}\left(T_2^C > t, T_{21}^I > t, T_{23}^I > t\right) &= \mathbb{P}\left(T_2^C > t\right)\mathbb{P}\left(T_{21}^I > t\right)\mathbb{P}\left(T_{23}^I > t\right) \\ &= e^{-\delta t} e^{-\lambda t} e^{-\lambda t} \\ &= e^{-(\delta+2\lambda)t}, \end{aligned} \quad (\text{A.3})$$

logo,

$$\begin{aligned} \mathbb{P}\left(\min\{T_2^C, T_{21}^I, T_{23}^I\} \leq t\right) &= 1 - \mathbb{P}\left(\min\{T_2^C, T_{21}^I, T_{23}^I\} > t\right) \\ &= 1 - e^{-(\delta+2\lambda)t}. \end{aligned} \quad (\text{A.4})$$

Portanto, concluímos que $\min\{T_2^C, T_{21}^I, T_{23}^I\} \sim \text{Exp}(\delta + 2\lambda)$. De modo análogo, sabemos que $T_4^C \sim \text{Exp}(\delta)$ e $T_{43}^I \sim \text{Exp}(\lambda)$ são independentes. Dessa forma, temos que:

$$\mathbb{P}\left(\min\{T_4^C, T_{43}^I\} \leq t\right) = 1 - e^{-(\delta+\lambda)t}. \quad (\text{A.5})$$

Portanto, concluímos que $\min \{T_4^C, T_{43}^I\} \sim Exp(\delta + \lambda)$. Fazendo as seguintes substituições $T^* = \min \{T_2^C, T_{21}^I, T_{23}^I\}$ e $T^{**} = \min \{T_4^C, T_{43}^I\}$, na Equação A.1, obtemos:

$$\begin{aligned} \epsilon_2 &= \mathbb{P}\left(\min \{T_2^C, T_{21}^I, T_{23}^I\} < \min \{T_4^C, T_{43}^I\}\right) \\ &= \mathbb{P}(T^* < T^{**}). \end{aligned} \quad (\text{A.6})$$

Sabendo que T^* e T^{**} , respectivamente, são independentes a partir das propriedades de probabilidade condicional, $\mathbb{P}(T^* < T^{**})$ pode ser calculada como segue adiante:

$$\begin{aligned} \mathbb{P}(T^* < T^{**}) &= \int_0^\infty \mathbb{P}(T^* < T^{**} | T^* = t) (\delta + 2\lambda) e^{-(\delta+2\lambda)t} dt \\ &= \int_0^\infty \mathbb{P}(T^* > t) (\delta + 2\lambda) e^{-(\delta+2\lambda)t} dt \\ &= \int_0^\infty e^{-(\delta+\lambda)t} (\delta + 2\lambda) e^{-(\delta+2\lambda)t} dt \\ &= \int_0^\infty (\delta + 2\lambda) e^{-(\delta+\lambda+\delta+2\lambda)t} dt \\ &= \int_0^\infty (\delta + 2\lambda) e^{-(2\delta+3\lambda)t} dt \\ &= \frac{\delta + 2\lambda}{2\delta + 3\lambda}. \end{aligned} \quad (\text{A.7})$$

Portanto, concluímos que $\epsilon_2 = \frac{\delta+2\lambda}{2\delta+3\lambda}$, em que $n_S = 2$, $N_I = 2$ e $N_S = 3$, como queríamos demonstrar. De modo análogo, a probabilidade para $i = 4$ ser escolhido é:

$$\begin{aligned} \epsilon_4 &= \mathbb{P}\left(\min \{T_4^C, T_{43}^I\} < \min \{T_2^C, T_{21}^I, T_{23}^I\}\right) \\ &= \frac{\delta + \lambda}{2\delta + 3\lambda}. \end{aligned} \quad (\text{A.8})$$

De modo geral, ϵ_i pode ser escrita da seguinte forma:

$$\epsilon_i = \frac{\delta + n_i^S \lambda}{N_I \delta + N_S \lambda}, \quad (\text{A.9})$$

em que N_I é a quantidade total de vértices infectados e N_S é a soma da quantidade total de vizinhos suscetíveis de todos os vértices infectados, que pode ser escrita como:

$$N_S = \sum_{i \in \mathcal{E}^{(inf)}} n_i^S. \quad (\text{A.10})$$

Note que ϵ_i , por ser uma probabilidade, tem que satisfazer a seguinte propriedade:

$$\sum_{i \in \mathcal{E}^{(inf)}} \epsilon_i = 1. \quad (\text{A.11})$$

Verificando para as expressões do nosso exemplo, temos que $\mathcal{E}^{(inf)} = \{2, 4\}$, logo:

$$\begin{aligned} \sum_{i \in \mathcal{E}^{(inf)}} \epsilon_i &= 1 & (\text{A.12}) \\ \sum_{i \in \{2,4\}} \epsilon_i &= 1 \\ \epsilon_2 + \epsilon_4 &= 1 \\ \frac{\delta + 2\lambda}{2\delta + 3\lambda} + \frac{\delta + \lambda}{2\delta + 3\lambda} &= 1 \\ \frac{\delta + 2\lambda + \delta + \lambda}{2\delta + 3\lambda} &= 1 \\ \frac{2\delta + 3\lambda}{2\delta + 3\lambda} &= 1 \\ 1 &= 1. \blacksquare \end{aligned}$$

Como foi visto na seção 3.1.1, a probabilidade de um vértice infectado tornar-se suscetível ou recuperado, no caso do modelo SIR, é dada por:

$$\mu_i = \frac{\delta}{\delta + n_i^S \lambda}. \quad (\text{A.13})$$

Essa probabilidade também pode ser justificada usando as mesmas ideias aplicadas para a dedução da expressão 3.1. Por exemplo, suponha que o vértice $i = 2$ apresentado no diagrama da Figura 37 seja sorteado. A probabilidade de que ele seja curado antes de infectar um de seus vizinhos pode ser dada calculando a probabilidade do tempo de ocorrência de uma marca de cura associada a $i = 2$ seja menor que todas as marcas de infecção de $i = 2$ para seus vizinhos $j = 1$ e $j = 3$. Isto é, o tempo que o vértice $i = 2$ leva pra ser curado é menor que o tempo até que ele transmita a infecção para um de seus vizinhos. Em termos matemáticos, essa probabilidade é dada expressa por:

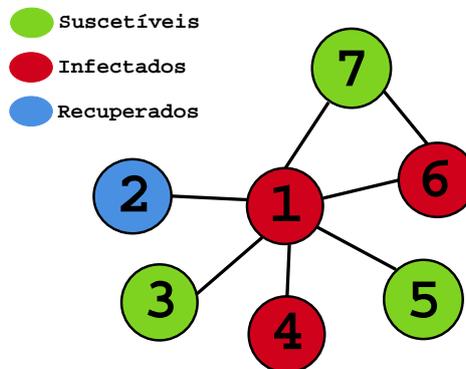
$$\begin{aligned} \mu_2 &= \mathbb{P}(T_2^C < T_{21}^I, T_2^C < T_{23}^I) & (\text{A.14}) \\ &= \mathbb{P}(T_2^C = \min(T_{21}^I, T_{23}^I)) \\ &= \frac{\delta}{\delta + \lambda + \lambda} \\ &= \frac{\delta}{\delta + 2\lambda}. \end{aligned}$$

Como queríamos mostrar, pois o vértice $i = 2$ tem apenas $n_2^S = 2$ vizinhos suscetíveis. Para o modelo SIR, temos as mesmas expressões, com a diferença de que, quando o vértice é curado, não pode assumir nenhum outro estado ao longo do tempo.

APÊNDICE B – DEDUÇÕES DAS EXPRESSÕES 3.7 E 3.10

Vamos seguir as mesmas ideias da Seção A para demonstrar as expressões 3.7 e 3.10. Considere a realização do modelo MT no grafo $G = (V, A)$ apresentado na Figura 38, em que temos, $V = \{1, 2, 3, 4, 5, 6, 7\}$ e $A = \{(1, 2), (1, 3), (1, 4), (1, 5), (1, 6), (1, 7), (6, 7)\}$. Considerando que o processo em um certo instante t_0 apresenta a configuração $X_1(t_0) = 1, X_2(t_0) = 2, X_3(t_0) = 0, X_4(t_0) = 1, X_5(t_0) = 0, X_6(t_0) = 1, X_7(t_0) = 0$, em que 0 significa que o vértice está no estado suscetível (*ignorante*), 1 infectado (*informante*) e 2 recuperado (*neutro*).

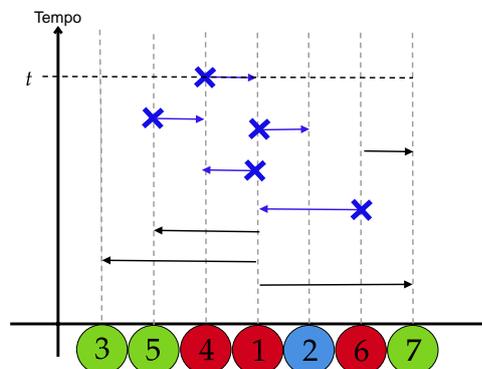
Figura 38 – Representação visual de uma realização do processo no tempo t .



Fonte: Produzido pelo autor.

O diagrama da Figura 39 apresenta a realização do modelo MT até o instante t . Nesse caso, as setas de infecção seguem um processo de Poisson com parâmetro λ (setas pretas), e as marcas de cura seguem um processo de Poisson com parâmetro α (marcas azuis).

Figura 39 – Representação visual de uma realização do processo no tempo t .



Fonte: Produzido pelo autor.

Neste caso, as marcas de cura representam as interações entre os pares de vértices *informante - informante* e *informante - neutro*. Note que $\mathcal{E}^{(ig)} = \{3, 5, 7\}$ e $\mathcal{E}^{(inf)} = \{1, 4, 6\}$, respectivamente. Então, a probabilidade do vértice no estado informante $i = 1$ ser escolhido antes de $j = 4$ e $j = 6$ é dada por:

$$\begin{aligned} \epsilon_1 &= \mathbb{P}\left(\min\{T_{12}^C, T_{14}^C, T_{16}^C, T_{13}^I, T_{15}^I, T_{17}^I\} < \min\{T_{41}^C, T_{61}^C, T_{67}^I\}\right) \\ &= \frac{3\alpha + 3\lambda}{3\alpha + 3\lambda + 2\alpha + \lambda} \\ &= \frac{3\alpha + 3\lambda}{5\alpha + 4\lambda}. \end{aligned} \quad (\text{B.1})$$

Note que 3α da expressão $\frac{3\alpha+3\lambda}{5\alpha+4\lambda}$ é dado por $(k_1 - n_1^S)\alpha = (6-3)\alpha = 3\alpha$, em que k_1 é o grau do vértice $i = 1$ e n_1^S é o número de vizinhos de $i = 1$ que estão no estado ignorante, 3λ é dado por $n_1^{Ig}\lambda$. Também, note que $5\alpha = N_{IN}\alpha$, em que $N_{IN} = [(k_1 - n_1^S) + (k_4 - n_4^S) + (k_6 - n_6^S)]$ que, por sua vez, é $N_{IN} = [(6 - 3) + (1 - 0) + (2 - 1)] = (3 + 1 + 1) = 5$. E, por fim, $4\lambda = N_{Ig}\lambda$, em que $N_{Ig} = (n_1^{Ig} + n_4^{Ig} + n_6^{Ig}) = (3 + 0 + 1) = 4$. Portanto, podemos concluir que a probabilidade de um dado vértice i informante ser escolhido é dada pelas seguintes expressões:

$$\epsilon_i = \frac{(k_i - n_i^{Ig})\alpha + n_i^{Ig}\lambda}{N_{IN}\alpha + N_{Ig}\lambda}, \quad (\text{B.2})$$

em que

$$\begin{aligned} N_{Ig} &= \sum_{i \in \mathcal{E}^{(inf)}} n_i^{Ig}, \\ N_{IN} &= \sum_{i \in \mathcal{E}^{(inf)}} (k_i - n_i^{Ig}), \end{aligned} \quad (\text{B.3})$$

em que N_{Ig} é a soma dos vizinhos ignorantes de todos os vértices informantes. Por fim, N_{IN} é a soma de todos os vizinhos não ignorantes de todos os vértices informantes. Para calcular a probabilidade de um dado vértice informante tornar-se neutro, vamos partir da suposição que o vértice $i = 1$ é escolhido. Do mesmo modo que vimos na Seção A, temos que essa probabilidade é dada por:

$$\begin{aligned}\mu_1 &= \mathbb{P}\left(\min\{T_{12}^C, T_{14}^C, T_{17}^C\} < \min\{T_{13}^I, T_{15}^I, T_{17}^I\}\right) \\ &= \frac{3\alpha}{3\alpha + 3\lambda}.\end{aligned}\tag{B.4}$$

Que, de modo geral, pode ser escrita como:

$$\mu_i = \frac{(k_i - n_i^{Ig})\alpha}{(k_i - n_i^{Ig})\alpha + n_i^{Ig}\lambda},\tag{B.5}$$

em que k_i é o grau do vértice i e n_i^{Ig} é a quantidade de ignorantes vizinhos ao vértice i .

APÊNDICE C – CORRIDA DAS EXPONENCIAIS

As expressões deduzidas nos apêndices A e B podem ser facilmente reproduzidas como casos particulares do teorema da corrida das exponenciais (DURRETT; DURRETT, 1999), que enunciamos a seguir.

Teorema 3 (Corrida das Exponenciais). *Seja a variável aleatória $V = \min\{T_1, \dots, T_n\}$, com:*

$$\mathbb{P}(V > t) = \exp\left(-\sum_{i=1}^n \lambda_i t\right), \quad (\text{C.1})$$

seja I a variável aleatória que descreve o índice do menor T_i , então:

$$\mathbb{P}(I = i) = \frac{\lambda_i}{\sum_{k \neq i}^n \lambda_k}. \quad (\text{C.2})$$

APÊNDICE D – CÓDIGO FONTE

Acesse o seguinte link do repositório com os códigos fonte: <https://github.com/luciano-mendes-jr/simulacoes_dissertacao>.