

Políticas Sensíveis ao Risco para o Controle da Propagação de Doenças Infecciosas

Henrique Dias Pastor¹,
Valdinei Freire¹, Leliane Nunes de Barros², Karina Valdivia Delgado¹

¹Escola de Artes, Ciências e Humanidades – Universidade de São Paulo (USP)
Rua Arlindo Bettio, 1000 – 03828-000 – São Paulo – SP – Brasil

²Instituto de Matemática e Estatística – Universidade de São Paulo (USP)
Rua do Matão, 1010 – 05508-090 – São Paulo – SP – Brasil

henrique.pastor@usp.br, valdinei.freire@usp.br, leliane@ime.usp.br, kvd@usp.br

Abstract. *The definition of policies to control the spread of infectious diseases, such as COVID19, is an important issue for society and for the government agents, responsible for the controlling measures. The majority of the vast number of recent research on this topic, uses past data to estimate the disease behavior in the future, considering an arbitrary policy. However, the use of automated planning techniques based on Markov Decision Processes (MDPs) has been shown to be more efficient to compute optimized control policies. One of the most common model used to control infectious diseases is the "SIR with vaccination model" in which the dynamics of the susceptible, infected and recovered individuals amount is controlled by vaccination policies. But since this is a problem of saving lives, it is also necessary to take into account the attitude towards risk of a decision maker. Thus, the purpose of this work is to use risk-sensitive MDPs to find optimized policies to control the spread of the considered infectious diseases, taking into account the parameters of the SIR with risk model. The results show that vaccination policies depend on both the basic reproductive rate R_0 , as usual, and the attitude towards the risk of the government agents.*

Resumo. *A definição de políticas de controle do espalhamento de doenças infecciosas, tais como a COVID19, é hoje um tema importante para a sociedade e para os agentes do governo, responsáveis pelas medidas de controle. A grande maioria das pesquisas recentes neste tópico usa dados do passado para estimar a evolução da doença considerando uma política arbitrária. No entanto, o uso de técnicas de planejamento automático baseadas em Processos de Decisão Markovianos (MDPs) tem se mostrado mais eficiente para a computação de políticas de controle otimizadas. Um dos modelos mais utilizados para o controle de doenças infecciosas é o "modelo SIR com vacinação" em que a dinâmica da quantidade de indivíduos suscetíveis, infectados e recuperados é controlada através de políticas de vacinação da população. Mas por se tratar de um problema que envolve vidas, também se faz necessário levar em consideração a atitude frente ao risco dos agentes tomadores de decisão. Assim, a proposta deste trabalho é usar MDPs sensíveis ao risco para encontrar políticas otimizadas de vacinação para o controle da propagação de doenças infecciosas considerando, além dos parâmetros do modelo SIR com vacinação,*

o parâmetro de risco. Os resultados mostram que as políticas de vacinação dependem da taxa de reprodução basal R_0 , como esperado, bem como da atitude frente ao risco dos agentes de tomada de decisões.

1. Introdução

O aparecimento de novas doenças infecciosas (por exemplo, HIV, SARS, COVID), e a persistência de outras, têm causado preocupação nos governos e na população em geral. Além das pesquisas farmacêuticas para a descoberta de vacinas ser um tópico de primordial importância, pesquisas sobre o controle do espalhamento da doença com políticas otimizadas de vacinação na população e medidas não-farmacêuticas (por exemplo, o distanciamento social, fechamento de escolas e espaços públicos), também são consideradas de vital importância [Yaesoubi and Cohen 2016, Yaesoubi and Cohen 2011a, Yaesoubi and Cohen 2011b]. Nessas pesquisas, é necessário considerar recursos governamentais limitados para o diagnóstico médico (por exemplo, o diagnóstico clínico e testes laboratoriais); o tratamento hospitalar (por exemplo, leitos de UTI e aparelhos de respiração); bem como o orçamento para a aplicação de vacinas, caso elas existam. Também é preciso considerar como o fechamento de setores da sociedade afeta a economia de uma cidade (estado ou país). Por fim, e o mais importante, é garantir que tais políticas de controle de espalhamento da doença sejam capazes de salvar um número máximo de vidas humanas.

Um primeiro passo para encontrar políticas de controle é o conhecimento da dinâmica dessas doenças infecciosas. Diversos modelos compartimentais têm sido propostos, entre eles, SI, SIS, SIR, SIRS, SIER, SIERS [Diekmann and Heesterbeek 2000, Keeling and Rohani 2011]. Um dos modelos mais conhecidos, por ser bastante simples, é o SIR [Kermack and McKendrick 1927], que envolve um sistema de equações diferenciais ordinárias descrevendo o espalhamento da doença em termos de fluxo da população entre compartimentos: Suscetíveis (S), Infectados (I) e Recuperados (R) (ou Removidos). A soma dos indivíduos em cada compartimento é o número total N de indivíduos da população em questão. Além disso, o modelo SIR considera uma taxa de recuperação espontânea, uma taxa de infecção e uma taxa de reprodução basal R_0 . Nesse modelo o tempo é contínuo e a dinâmica é determinística.

Existem alguns trabalhos na literatura que resolvem o problema de controle de epidemias modelados com SIR estocástico, considerando ações de vacinação de uma parcela da população a cada período de tempo [Lefèvre 1981, Yaesoubi and Cohen 2011a, Kinathil et al. 2017, Nasir and Rehman 2017]. Podemos modelar este tipo de problema como um Processo de Decisão Markoviano (*Markov Decision Process* – MDPs), cujo critério de otimização, em geral, é encontrar uma política (mapeamento entre estados e ações) com o menor custo acumulado esperado [Puterman 1994]. Quando aplicado ao problema de controle de doenças infecciosas com a existência de vacina, o objetivo é determinar a melhor política de vacinação de indivíduos não infectados (Suscetíveis), antes ou durante uma epidemia, para evitar que a doença se espalhe em uma população, considerando o custo das vacinas e do tratamento dos infectados (isto é, não queremos políticas que vacinem todos os indivíduos de uma população) [Lefèvre 1981, Yaesoubi and Cohen 2011a, Kinathil et al. 2017, Nasir and Rehman 2017]. Neste contexto, o agente tomador de decisão deve escolher o número de indivíduos que deverão

ser vacinados a cada período de tempo, para controlar a propagação da doença infecciosa, minimizando o custo acumulado esperado.

Neste trabalho, resolvemos o problema de controle de SIR estocástico com vacinação modelado como um MDP, porém diferente de outros trabalhos da literatura [Lefèvre 1981, Yaesoubi and Cohen 2011a, Nasir and Rehman 2017], consideramos o risco. Para isso usamos a modelagem dos MDPs sensíveis ao risco (*Risk Sensitive Markov Decision Process*– RSMDP). Em MDPs clássicos o objetivo é encontrar a política que minimiza o custo acumulado esperada. Entretanto, pode ocorrer que o custo acumulado esperado de uma execução da política encontrada seja consideravelmente pior do que o valor médio (esperado) [Mihatsch and Neuneier 2002], ou seja, este critério não leva em consideração a variância e por tal motivo é chamado de neutro a risco. No caso do problema de controle de propagação de doenças é importante levar em conta o risco por incluir decisões que envolvem vidas. Para isso, os RSMDPs utilizam critérios capazes de considerar essa variância.

O artigo está organizado da seguinte forma. Na seção 2 é descrito o modelo SIR de controle da propagação de doenças infecciosas incluindo a descrição da modelagem do custo e transições estocásticas, bem como a descrição do RSMDP utilizado para encontrar as políticas de controle. Na seção 3 são mostradas as políticas encontradas para diferentes valores da taxa de reprodução basal R_0 e diferentes valores do parâmetro de risco k . Finalmente, na seção 4 são apresentadas as conclusões.

2. Modelo para controle da propagação de doenças infecciosas

Existem vários modelos compartimentais determinísticos e estocásticos sobre propagação de doenças infecciosas. Dentre os modelos determinísticos estão [Diekmann and Heesterbeek 2000, Keeling and Rohani 2011]: SI, SIS, SIR, SIRS, SEIR, SEIRS, em que S representa os indivíduos suscetíveis, I infectados, R recuperados e E os indivíduos latentes (que estão infectados mas ainda não transmitem a doença).

A diferença entre os modelos é basicamente a forma com que se representam as epidemias. O modelo SI, e o SIS não levam em consideração o número de indivíduos recuperados, mas o SIS tem uma condição diferente que seria o caso de reinfeção. Os modelos SIR e SIRS adicionam os indivíduos recuperados. No modelo SEIR têm-se uma nova variável que seria o período de incubação (latência) da doença, durante o qual é dito ao indivíduo estar infetado mas não infeccioso.

Neste trabalho é usado o modelo SIR com vacinação [Kermack and McKendrick 1927], em que são conhecidos o número de indivíduos suscetíveis (X_S), infectados (X_I) e recuperados (X_R) e assume-se que o tamanho da população N é constante, ou seja não existem modificações do tamanho devido ao nascimento, morte ou imigração (os indivíduos mortos e os imigrante, permanecem no compartimento R). Além disso, assume-se que os indivíduos adquirem imunidade permanente por infecção ou vacinação.

No modelo SIR são consideradas duas taxas: γ que é a taxa de recuperação espontânea e β que é a taxa de infecção. A Figura 1 mostra a dinâmica do modelo SIR em termos dessas duas taxas, definindo assim o fluxo da população entre esses três compartimentos.

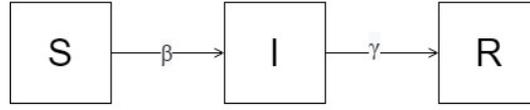


Figura 1. Modelo SIR: fluxo da população.

A evolução do modelo SIR pode ser analisada segundo a taxa de reprodução basal $R_0 = \beta/\gamma$ e é definida como o número esperado de infecções secundárias (indivíduos infectados por outros) decorrentes de um único indivíduo durante todo o seu período infeccioso, em uma população suscetível [Heesterbeek and Dietz 1996, Heffernan et al. 2005, Heesterbeek 2002]. Essa taxa é importante pois permite prever como a infecção se espalhará, além de permitir o estudo da dinâmica da infecção [Heffernan et al. 2005]. Uma epidemia acontece se o número de indivíduos infectados aumenta e isto ocorre quando $R_0 > 1$ [Heffernan et al. 2005, Kinathil et al. 2017].

2.1. Modelagem dos estados, ações e custos

Seja $X_S(t) \in \mathbb{N}$ o número de indivíduos suscetíveis, $X_I(t) \in \mathbb{N}$ infectados e $X_R(t) \in \mathbb{N}$ recuperados no tempo t . Seja ainda uma população de tamanho N , então, em todo tempo t tem-se $X_S(t) + X_I(t) + X_R(t) = N$. O estado do espalhamento da doença no tempo t pode ser representado por $s_t = \langle X_S(t), X_I(t) \rangle$, uma vez que $X_R(t) = N - (X_S(t) + X_I(t))$. O conjunto de estados é denotado por S e consiste de todas as combinações possíveis de X_S e X_I tal que $X_S + X_I \leq N$.

O conjunto de ações de vacinação é denotado por A , sendo que a cada estágio o tomador de decisão está em um estado $s_t = \langle X_S(t), X_I(t) \rangle$ e deve escolher a proporção de indivíduos de $X_S(t)$ que devem ser vacinados para controlar a epidemia. Neste trabalho são consideradas 11 frações de $X_S(t)$ distribuídas uniformemente entre 0 e 1, isto é, $A = \{0, 0.1, 0.2, \dots, 0.9, 1\}$.

O custo de se aplicar uma ação de vacinação $a \in A$ no estado $\langle X_S, X_I \rangle$ é calculado em função do custo de vacinação $cost_{vac}$ e custo de infecção $cost_{inf}$ (que inclui o custo do diagnóstico, custo do tratamento e outros custos indiretos). Assim, consideramos que a função custo é dada por:

$$C(\langle X_S, X_I \rangle, a) = (a \cdot X_S)^{1.5} \cdot cost_{vac} + cost_{inf} \cdot X_I. \quad (1)$$

2.2. Modelagem das transições estocásticas

Neste trabalho, será usado o SIR com vacinação estocástico e discreto. Esse modelo permite que o número de indivíduos que se deslocam entre os compartimentos da população variem de forma estocástica. Para isso, vamos primeiro definir o modelo SIR determinístico.

Considere o estado atual $\langle X_S(t), X_I(t) \rangle$ e as seguintes variáveis aleatórias:

- $D(t)$: quantidade de pessoas suscetíveis que foram vacinadas no passo t ;
- $E(t)$: quantidade de pessoas suscetíveis que foram infectadas no passo t ; e
- $F(t)$: quantidade de pessoas que se tornaram recuperadas no passo t , após estarem infectados.

Dessa forma, o próximo estado é dado por:

$$\begin{aligned} X_S(t+1) &= X_S(t) - D(t) - E(t), \\ X_I(t+1) &= X_I(t) - F(t) + E(t). \end{aligned} \quad (2)$$

Na formulação determinística tem-se que:

$$\begin{aligned} D(t) &= a \cdot X_S(t), \\ E(t) &= \beta \cdot \frac{X_I(t)}{N} \cdot (X_S(t) - D(t)), \\ F(t) &= \gamma \cdot X_I(t). \end{aligned}$$

A formulação probabilística e discreta, usada neste trabalho, considera as mesmas dependências entre as variáveis aleatórias, mas adicionando aleatoriedade. A Figura 2 exibe a Rede Bayesiana Dinâmica que representa a dependência entre estas variáveis aleatórias.

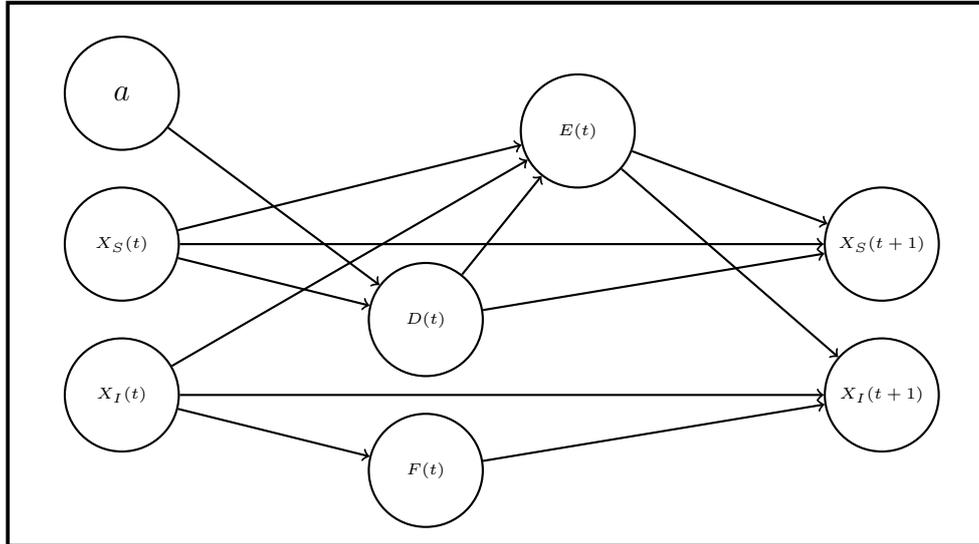


Figura 2. Rede Bayesiana Dinâmica representando as dependências entre as variáveis de estado no tempo t e $t+1$ decompostas nas variáveis intermediárias $D(t)$, $E(t)$ e $F(t)$.

A primeira introdução de incerteza vem do fato de não podermos vacinar frações de indivíduos. Considere $m = \text{floor}(a \cdot X_S(t))$. Então a quantidade de pessoas suscetíveis que foram vacinadas é dada pela seguinte distribuição:

$$\begin{aligned} \Pr(D(t) = m) &= (m+1) - a \cdot X_S(t), \\ \Pr(D(t) = m+1) &= a \cdot X_S(t) - m. \end{aligned} \quad (3)$$

Outra fonte de incerteza é associada à taxa de infecção β no cálculo da quantidade de pessoas suscetíveis e que foram infectadas, $E(t)$, modelada por uma distribuição binomial:

$$\Pr(E(t) = y | D(t) = x) = \text{Binomial} \left(y; n = X_S(t) - x, \beta \cdot \frac{X_I(t)}{N} \right). \quad (4)$$

Finalmente, a terceira fonte de incerteza é devido à taxa de recuperação espontânea γ . Considere $q = \text{floor}(\gamma \cdot X_I(t))$. Então a quantidade de pessoas recuperadas é dada pela seguinte distribuição:

$$\begin{aligned}\Pr(F(t) = q) &= (q + 1) - \gamma \cdot X_I(t), \\ \Pr(F(t) = q + 1) &= \gamma \cdot X_I(t) - q.\end{aligned}\tag{5}$$

Assim, no SIR probabilístico, para especificar a transição probabilística entre dois estados s e s' quando a ação a é executada, isto é, $\Pr(s_{t+1} = s' | s_t = s, a_t = a)$, modificamos as equações para $X_S(t+1)$ e $X_I(t+1)$ (Equação 2) considerando as distribuições de probabilidades para $D(t)$, $E(t)$, e $F(t)$.

2.3. Políticas ótimas sensíveis ao risco

Considerando o conjunto de estados S , o conjunto de ações A , a função de custo C e a função de transição probabilística, pode-se escolher uma política de controle da doença, isto é, a porcentagem de indivíduos que devem ser vacinados para um dado estado $\langle X_S(t), X_I(t) \rangle$.

O critério utilizado em MDPs clássicos é minimizar o custo acumulado esperado para encontrar a política ótima. Porém, este critério não leva em consideração a variância, podendo ocorrer que o custo acumulado de uma execução da política encontrada seja consideravelmente pior do que o valor médio [Mihatsch and Neuneier 2002]. Por tal motivo, esses MDPs são chamados de neutros ao risco.

Por outro lado, para encontrar políticas de controle para o SIR é importante considerar o risco. Assim, acreditamos que um melhor modelo a ser usado no caso de controle de doenças infecciosas é uma extensão de MDPs chamada de processos de decisão markovianos sensíveis ao risco (*Risk Sensitive Markov Decision Process*—RSMDP). O desenvolvimento de algoritmos que consideram o risco na tomada de decisão, é um assunto pouco explorado na literatura dos MDPs [García and Fernández 2015]. Existem diferentes abordagens para quantificar o risco, como: utilidade exponencial esperada [Howard and Matheson 1972, Jaquette 1976, Denardo and Rothblum 1979, Rothblum 1984, Patek 2001], o uso de uma função de transformação linear por partes [Mihatsch and Neuneier 2002, Borges et al. 2018], soma ponderada entre valor esperado e variância [Sobel 1982, Filar et al. 1989] e a estimativa de desempenho em um intervalo de confiança [Filar et al. 1995, Yu et al. 1998, Hou et al. 2014, Hou et al. 2016].

Uma abordagem que permite encontrar uma política estacionária ótima como solução, bem como uma escolha arbitrária do fator de risco sem apresentar problemas de precisão numérica foi introduzida em [Mihatsch and Neuneier 2002]. Essa abordagem, chamada neste trabalho de RSMDP, baseada em uma função linear por partes e usa um fator de risco k . O agente tomador de decisão é averso ao risco se $k \in (0, 1)$, o agente é neutro se $k = 0$ e o agente é propenso ao risco se $k \in (-1, 0)$. Além disso, neste trabalho modelamos o problema como um RSMDP dirigido a meta, isto é, é considerado um conjunto de estados meta G . No problema de controle de doenças infecciosas G pode ser $X_I = 0$ ou $X_I = 0.1 \times N$.

No RSMDP dirigido a meta é usada uma função de transformação $\chi^{(k)}$, que se baseia no valor de entrada $z \in \mathbb{R}$, chamado de diferença temporal, e no fator de risco k ,

sendo essa transformação dada por [Mihatsch and Neuneier 2002]:

$$\chi^{(k)}(z) = \begin{cases} (1 - k)z, & \text{se } z < 0. \\ (1 + k)z, & \text{caso contrário.} \end{cases} \quad (6)$$

A solução de um RSMDP dirigido a meta é uma política estacionária $\pi : S \rightarrow A$. Para avaliar uma política π é usada a função valor $V_\pi^k(s)$, a qual usa a função de transformação $\chi^{(k)}$ sobre a diferença temporal e pode ser obtida resolvendo o seguinte sistema de equações para todo $s \in S$ [Mihatsch and Neuneier 2002]:

$$\sum_{s' \in S} Pr(s'|s, \pi(s)) \chi^{(k)}(C(s, \pi(s)) + \gamma V_\pi^k(s') - V_\pi^k(s)) = 0. \quad (7)$$

Similar aos MDPs neutros ao risco, existem políticas ótimas estacionárias para RSMDPs dirigidos a meta e sua correspondente função valor ótima é única [Mihatsch and Neuneier 2002]. Para cada $k \in (1, -1)$ existe uma única função valor ótima, $V_k^*(s) = \min_{\pi \in \Pi} V_k^\pi(s), \forall s \in S$, que satisfaz a seguinte equação [Mihatsch and Neuneier 2002]:

$$\min_{a \in A} \sum_{s' \in S} Pr(s'|s, a) \mathcal{X}^{(k)}(C(s, a) + V_k^*(s') - V_k^*(s)) = 0, \forall s \in S.$$

Conhecendo-se V_k^* , uma política π^* ótima pode ser obtida por:

$$\pi^*(s) = \arg \min_{a \in A} \sum_{s' \in S} Pr(s'|s, a) \mathcal{X}^{(k)}(C(s, a) + V_k^*(s') - V_k^*(s)).$$

2.4. Algoritmo utilizado para encontrar as políticas

Neste trabalho, para encontrar uma política ótima para RSMDPs dirigidos a meta, foi usado o algoritmo de Iteração de Valor Sensível ao Risco (Risk Sensitive Value Iteration – RSVI) que atualiza a função $Q(s, a)$ para todos os pares estado-ação a cada iteração da seguinte forma [Mihatsch and Neuneier 2002, Borges et al. 2018]:

$$Q^i(s, a) \leftarrow Q^{i-1}(s, a) + \alpha \sum_{s' \in S} Pr(s'|s, a) \chi^{(k)}(C(s, a) + \min_a Q^{i-1}(s', a) - Q^{i-1}(s, a)),$$

em que α é o tamanho do passo. Dado $Q^i(s, a)$, pode ser obtida a função valor na iteração i por:

$$V^i(s) = \min_{a \in A} \{Q^i(s, a)\}$$

e a política gulosa por:

$$\pi(s) = \arg \min_{a \in A} \{Q^i(s, a)\}.$$

O algoritmo RSVI tem garantia de convergência se $0 \leq \alpha \leq (1 + |k|)^{-1}$.

3. Experimentos

Os experimentos foram realizados em Java 8 SE com IntelliJ IDEA 2017.1.2 x64 em um processador AMD 8320e de 3.2GHz, 8GB de RAM 1600mhz DDR3 e 1TB de disco rígido.

Nos experimentos foram utilizados o custo de vacinação $cost_{vac} = 1$ e o custo de infecção $cost_{inf} = 4$, para cada indivíduo. Foram encontradas políticas para cinco valores de R_0 considerando diferentes combinações da taxa de recuperação espontânea $\gamma \in \{0.25; 0.75; 1\}$ e taxa de infecção $\beta \in \{0.2; 0.25; 0.75\}$ e variando o fator de risco k .

As políticas encontradas apresentam o comportamento esperado em termos de atitude ao risco, como é descrito a seguir.

Tomador de decisão averso ao risco. Um agente de controle muito averso ao risco tem uma tendência de vacinar a maior proporção de susceptíveis.

Tomador de decisão propenso ao risco. Um agente de controle muito propenso ao risco tem uma tendência quase que total de vacinar o mínimo possível.

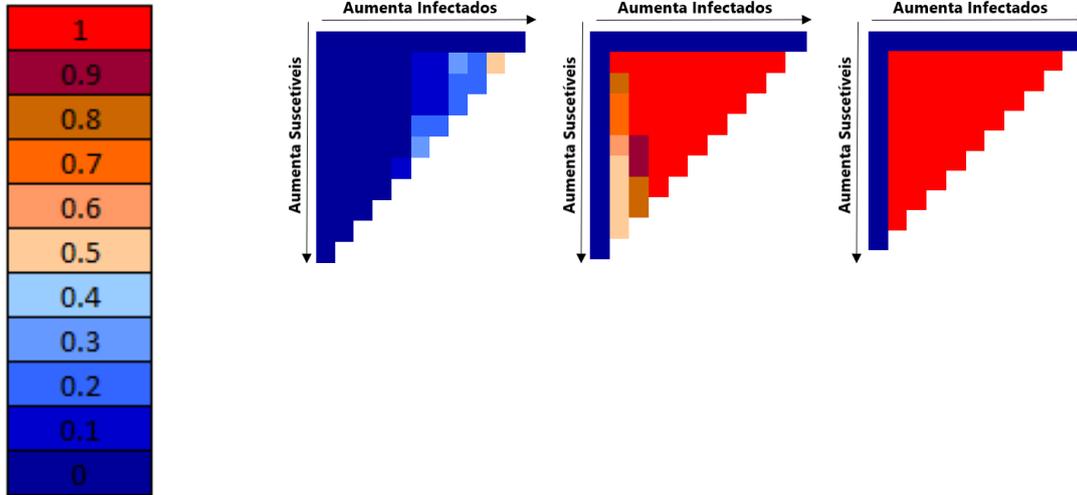
A Tabela 1 mostra três exemplos ilustrativo de políticas geradas para $N=10$, com $\gamma = 0.25$ e $\beta = 0.75$ ($R_0 = 3$) e com $k = -0.99$ (propensa ao risco), $k = 0$ (neutra ao risco) e $k = 0.99$ (aversa ao risco). Cada política é mostrada como uma matriz $\mathbf{B}[i, j] = a_{ij}$ em que as ações $a_{ij} \in \{0, 0.1, 0.2, \dots, 1\}$. Cada um dos valores possíveis das ações é representado com uma cor diferente como mostrado na legenda. As linhas i da matriz são enumeradas de 0 até 10 representando o número de pessoas suscetíveis e as colunas j são enumeradas de 0 até 10 representando o número de pessoas infectadas (o restante dos indivíduos é a quantidade de recuperados, sendo que $X_S + X_I \leq N$). Por exemplo, para $k = -0.99$ e para o estado $\langle X_S = 3, X_I = 7 \rangle$ (quarta linha e oitava coluna da matriz), a ação sugerida pela política é vacinar o 20% dos suscetíveis, isto é, 0.6 pessoas. Assim, de acordo com a Equação 3 a probabilidade de vacinar uma pessoa é 0.6 e de vacinar 0 pessoas é 0.4.

A Figura 3 mostra as políticas para uma população $N = 100$, e com $cost_{inf} = 4$ e $cost_{vac} = 1$ e sete diferentes valores do fator de risco k . O objetivo é mostrar o comportamento da política quando a taxa de reprodução basal R_0 diminui. Nas linhas, são mostradas as políticas para $R_0 = 0.75/0.25 = 3$, $R_0 = 0.8/0.4 = 2$, $R_0 = 0.25/0.25 = 1$, $R_0 = 0.2/0.25 = 0.8$ e $R_0 = 0.25/0.75 = 0.33$. Observando cada coluna individualmente, de cima para baixo, vemos que em geral a proporção escolhida de vacinação também diminui (isto é, aumentam os estados em azul).

Na Figura 3, para um fator $R_0 = 3$ (primeira linha), mesmo para um agente propenso ao risco ($k < 0$), a política sugere vacinar mais do que 50% dos suscetíveis na maioria dos estados pois existe muito risco da doença se espalhar e se tornar uma epidemia. Para $R_0 = 2$ (segunda linha), mesmo para um agente averso ao risco ($k > 0$), essa porcentagem diminui para 30%. As políticas para $R_0 = 1$ e $R_0 = 0.8$ (terceira e quarta linhas) são similares para cada um dos valores de k . A diferença entre elas é que para $R_0 = 0.8$ existe uma quantidade ligeiramente menor de ações que vacinam uma proporção maior ou igual que 20% dos suscetíveis. Para $R_0 = 0.33$ (quinta linha) as

Tabela 1. Exemplos de políticas sensíveis ao risco para $N=10$, $\gamma=0.25$ e $\beta=0.75$

Ações	propensa ao risco	neutra ao risco	aversa ao risco
Prob. de vacinação	$k = -0.99$	$k = 0$	$k = 0.99$



políticas para $k \leq 0$ indicam vacinar 10% ou menos da população na maioria dos estados, enquanto agentes mais aversos ao risco ($k = 0.8$ e $k = 0.9$) indicam vacinar de 50% a 100% dos suscetíveis em alguns estados.

4. Conclusão

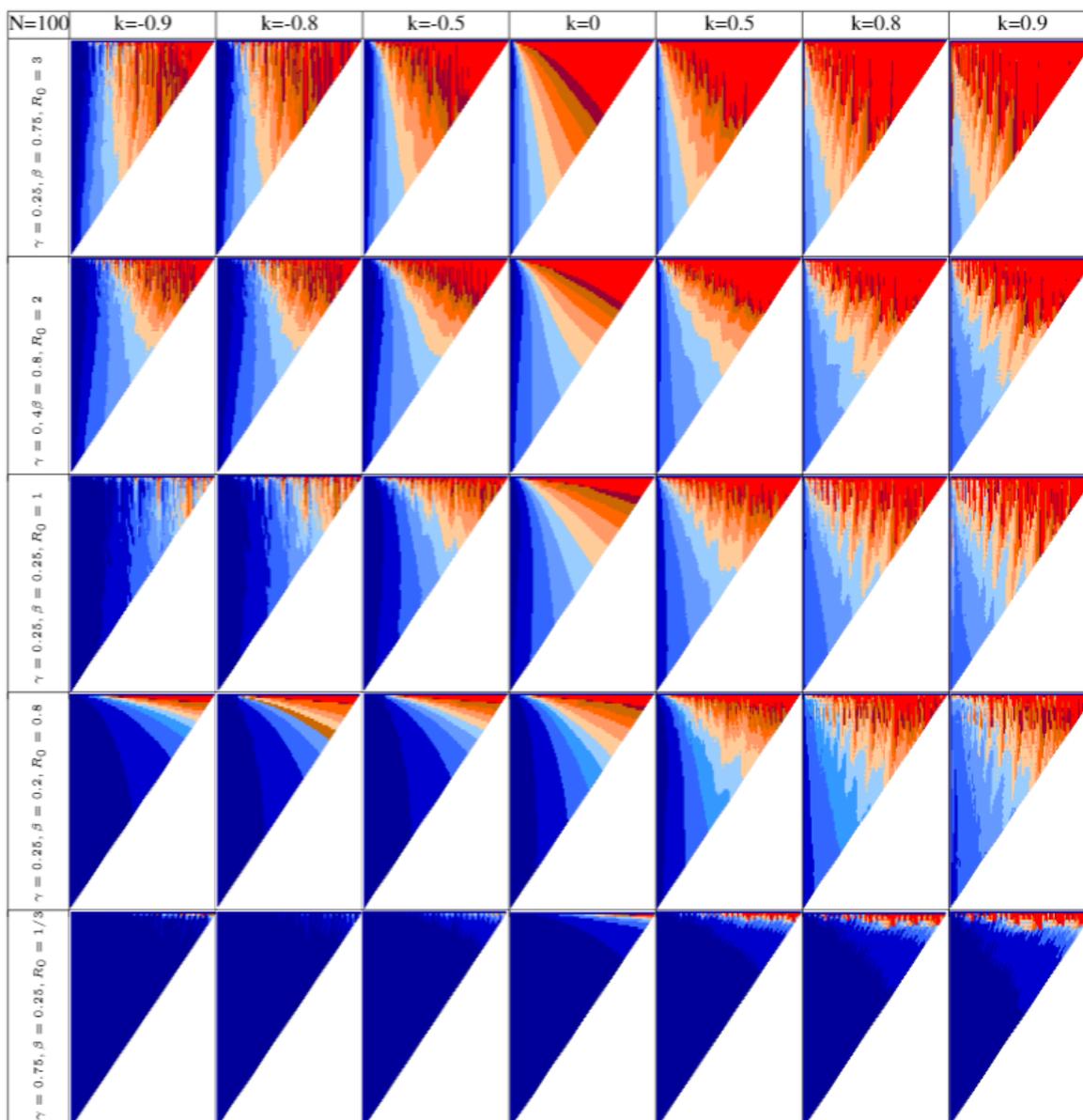
Neste trabalho é proposto o uso de MDPs sensíveis ao risco para encontrar políticas de controle de propagação de doenças infecciosas considerando, além dos parâmetros do modelo SIR estocástico e discreto com vacinação, o parâmetro de risco k . O objetivo principal do trabalho é mostrar a aplicabilidade desse modelo e como considerar a atitude ao risco de agentes responsáveis por definir medidas de controle da doença. Com base na literatura, esse modelo não foi utilizado até então para resolver este problema.

Os resultados mostram que, mesmo para $R_0 > 1$ (taxas que indicam a possibilidade de uma epidemia), nem sempre a política ótima em alguns estados indicam a vacinação de toda a população de suscetíveis. Dependendo da atitude ao risco do agente de controle da epidemia, indicado pelo valor k , a política ótima pode recomendar a vacinação de uma porcentagem menor da população. Por outro lado se $R_0 \leq 1$, nem sempre a política ótima é vacinar menos de 20% da população de suscetíveis, uma vez que esta também vai depender da atitude do agente. Para um agente muito averso ao risco, e com tal taxa de reprodução, a política ótima pode recomendar vacinar uma porcentagem maior da população.

Note ainda que as políticas ótimas sensíveis ao risco, os custos médios acumulados e também o desvio padrão, podem ser analisadas pelo agente de controle diante de informações extra como os orçamentos totais disponíveis para diagnóstico e tratamento da população, bem como o orçamento para aplicação de vacinas para toda a população.

Como trabalhos futuros, pretende-se rodar experimentos com populações mai-

Figura 3. $N=100$, $cost_{inf} = 4$ e $cost_{vac} = 1$



ores. Para isso podemos considerar, no lugar de indivíduos, frações da população de susceptíveis e de infectados, isto é, pode-se considerar os valores $X_S(t)/N$ e $X_I(t)/N$, respectivamente.

Referências

- Borges, I. O., Delgado, K. V., and Freire, V. (2018). Análise do algoritmo de iteração de valor sensível a risco. In *XV Encontro Nacional de Inteligência Artificial e Computacional (ENIAC)*, pages 365–376, São Paulo.
- Denardo, E. V. and Rothblum, U. G. (1979). Optimal stopping, exponential utility, and linear programming. *Mathematical Programming*, 16(1):228–244.

- Diekmann, O. and Heesterbeek, J. (2000). *Mathematical Epidemiology of Infectious Diseases: model building, analysis and interpretation*. John Wiley & Son.
- Filar, J. A., Kallenberg, L. C. M., and Lee, H.-M. (1989). Variance-penalized Markov decision processes. *Mathematics of Operations Research*, 14(1):147–161.
- Filar, J. A., Krass, D., Ross, K. W., and Ross, K. W. (1995). Percentile performance criteria for limiting average Markov decision processes. *IEEE Transactions on Automatic Control*, 40(1):2–10.
- García, J. and Fernández, F. (2015). A comprehensive survey on safe reinforcement learning. *J. Mach. Learn. Res.*, 16(1):1437–1480.
- Heesterbeek, J. and Dietz, K. (1996). The concept of r_0 in epidemic theory. *Statistica Neerlandica*, 50(1):89–110.
- Heesterbeek, J. A. P. (2002). A brief history of r_0 and a recipe for its calculation. *Acta biotheoretica*, 50(3):189–204.
- Heffernan, J. M., Smith, R. J., and Wahl, L. M. (2005). Perspectives on the basic reproductive ratio. *Journal of The Royal Society Interface*, 2:281 – 293.
- Hou, P., Yeoh, W., and Varakantham, P. (2014). Revisiting risk-sensitive MDPs: New algorithms and results. In *Proceedings of the Twenty-Fourth International Conference on Automated Planning and Scheduling, ICAPS 2014, Portsmouth, New Hampshire, USA, June 21-26, 2014*.
- Hou, P., Yeoh, W., and Varakantham, P. (2016). Solving risk-sensitive POMDPs with and without cost observations. In *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, February 12-17, 2016, Phoenix, Arizona, USA.*, pages 3138–3144.
- Howard, R. A. and Matheson, J. E. (1972). Risk-sensitive Markov decision processes. *Management science*, 18(7):356–369.
- Jaquette, S. C. (1976). A utility criterion for Markov decision processes. *Management Science*, 23(1):43–49.
- Keeling, M. J. and Rohani, P. (2011). *Modeling infectious diseases in humans and animals*. Princeton University Press.
- Kermack, W. O. and McKendrick, A. G. (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London A: mathematical, physical and engineering sciences*, 115(772):700–721.
- Kinathil, S., Soh, H., and Sanner, S. (2017). Analytic decision analysis via symbolic dynamic programming for parameterized hybrid MDPs. In *Proceedings of the 27th Conference on Automated Planning and Scheduling (ICAPS-17)*, Pittsburgh, PA.
- Lefèvre, C. (1981). Optimal control of a birth and death epidemic process. *Operations Research*, 29(5):971–982.
- Mihatsch, O. and Neuneier, R. (2002). Risk-sensitive reinforcement learning. *Machine Learning*, 49(2):267–290.

- Nasir, A. and Rehman, H. (2017). Optimal control for stochastic model of epidemic infections. In *2017 14th International Bhurban Conference on Applied Sciences and Technology (IBCAST)*, pages 278–284. IEEE.
- Patek, S. D. (2001). On terminating Markov decision processes with a risk-averse objective function. *Automatica*, 37(9):1379–1386.
- Puterman, M. L. (1994). *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons, Inc., New York, NY, USA, 1st edition.
- Rothblum, U. G. (1984). Multiplicative Markov decision chains. *Mathematics of Operations Research*, 9(1):6–24.
- Sobel, M. J. (1982). The variance of discounted Markov decision processes. *Journal of Applied Probability*, 19(4):794–802.
- Yaesoubi, R. and Cohen, T. (2011a). Dynamic health policies for controlling the spread of emerging infections: influenza as an example. *PloS one*, 6(9):e24043.
- Yaesoubi, R. and Cohen, T. (2011b). Generalized Markov models of infectious disease spread: A novel framework for developing dynamic health policies. *European Journal of Operational Research*, 215(3):679–687.
- Yaesoubi, R. and Cohen, T. (2016). Identifying cost-effective dynamic policies to control epidemics. *Statistics in medicine*, 35(28):5189–5209.
- Yu, S. X., Lin, Y., and Yan, P. (1998). Optimization models for the first arrival target distribution function in discrete time. *Journal of Mathematical Analysis and Applications*, 225(1):193 – 223.