Modelos de Epidemia: SI e SIR : Lista 2 - Probabilidades

Nestor Caticha

March 2020

1 Introdução

Apresento abaixo dois modelos de epidemias. A dedução é feita usando a regra do produto e soma lógica da probabilidade para manipular asserções associadas às variáveis do problema. Alguns pontos serão esclarecidos e outros vocês devem completar. Estes serão indicados por ($\otimes$). O objetivo desta lista é adquirir experência em modelagem de sistemas complexos usando técnicas de probabilidade, tentando deixar claras as hipóteses subjascentes em cada decisão sobre como construir o modelo. A análise destas hipóteses pode ser usada para criar outros modelos. Estes modelos são muito simplificados, mas a partir deles podemos facilmente colocar mais ingredientes de forma a que sejam úteis -se não para administrar a crise atual- pelo menos para que entendam o quê os especialistas estão fazendo. Esta modelagem também deixa claro que tipo de informação, caso a tivessemos, seria útil para obter melhores modelos.

2 Modelo SI

O estado de um agente (uma pessoa?) é descrito por uma variável aleatória X, que toma valores em $\{S, I\}$, que significam susceptíveis (ou sadios) e infectados (ou doentes) respectivamente. Precisamos um índice temporal t que evolui de forma discreta - talvez dia a dia, ou por hora. Também poderemos considerar o limite zero do intervalo. Por simplicidade, algumas vezes escreveremos a asserção $X_t = S$ de forma simplificada como S_t.

As variáveis de interesse são as probabilidades

$$P(S_t|A) \ e \ P(I_t|A), \tag{1}$$

onde A é a informação disponível - tal como condições iniciais e modelo de transmissão. Obviamente $1 = P(S_t|A) + P(I_t|A)$

O modelo de transmissão é definido por $P(I_{t^+}|S_tA)$ que por agora supomos constante e independente do tempo. O significado de t^+ é um tempo intermediário de transição entre t e $t+1$ [1]. É em t^+ que ocorrem as transições. Neste modelo vamos supor que não há chance de sarar e portanto temos a seguinte relação entre asserções

$\otimes$ "$X_{t+1} = I$" = "$X_t = I$" OU ("$X_t = S$" E "infectou em $t^+ : t \to t+1$").

A asserção "está infectado hoje" é equivalente a "estava infectado ontem " OU "estava sadio ontem E ficou doente hoje".

Ou seja, há duas formas de estar doente hoje: já estava ou acabou de fazer a transição.

Então a probabilidade $P(I_{t+1}|A)$ pode ser escrita como

$$\begin{aligned} P(X_{t+1} = I|A) &= P((X_t = I) \vee ((X_t = S) \wedge (X_t = I))|A), \\ P(X_{t+1} = I|A) &= P((X_t = I|A) + P((X_t = S) \wedge (X_t = I_{t^+})|A) \\ P(X_{t+1} = I|A) &= P((X_t = I|A) + P((X_t = S)|A)P((X_{t^+} = I)|(X_t = S) \wedge A), \\ P(I_{t+1}|A) &= P(I_t|A) + P(S_t|A)P(I_{t^+}|S_tA), \end{aligned} \tag{2}$$

onde a última expressão está escrita com uma notação mais simples. Acima simplesmente usamos a regra da soma e do produto lógicos. Também temos, da normalização

$$P(S_{t+1}|A) = 1 - P(I_{t+1}|A) = P(S_t|A) - P(S_t|A)P(I_{t^+}|S_tA) \tag{3}$$

Se multiplicarmos estas probabilidades por N, o número de agentes, teremos uma estimativa das populações infectadas e sadias.

A probabilidade de transição de sadio para doente será parametrizada por ($\otimes$):

$$P(I_{t^+}|S_tA) = \alpha \Delta t P(I_t|A), \tag{4}$$

proporcional a Δt em primeira ordem. α, a taxa de transição por unidade de tempo do estado sadio para doente é o único parâmetro do modelo. Note, e aqui está um ponto muito importante, que também aparece o fator $P(I_t|A)$, Se não houvesse doentes o agente não ficaria infectado, portanto em primeira ordem é proporcional à probabilidade de outros agentes estarem infectados. Como supomos uma sociedade homogênea, os outros agentes estão infestados com probabilidade igual entre si.

Se denotarmos as populações por letras minúsculas as populações evoluem de acordo com (($\otimes$)):

$$s_{t+1} = s_t\left(1 - \frac{\alpha \Delta t}{N} i_t\right), \tag{5}$$

$$i_{t+1} = i_t\left(1 + \frac{\alpha \Delta t}{N} s_t\right) \tag{6}$$

Podemos escrever as diferenças dividas por Δt e tomar o limite para tempo contínuo para obter o sistema de equações diferenciais

$$\dot{s} = -\frac{\alpha}{N} si, \tag{7}$$

[1] Em processos estocásticos a tempo contínuo a escolha de $P(I_{t^+}|S_tA)$ ou $P(I_{t+1}|S_tA)$ pode levar a resultados diferentes. Temos o cuidado de escrever t^+ para denotar o instante de transição. As escolhas diferentes estão associadas aos nomes de Itô e Stratonovich.

$$\dot{i} = \frac{\alpha}{N} s i, \tag{8}$$

$$\tag{9}$$

sujeito a condições iniciais que satisfaçam $s(0) + i(0) = N$.

($\otimes$) É fácil mostrar que

$$i(t) = N \frac{i(0)}{i(0) + e^{-\alpha t}(N - i(0))} \tag{10}$$

substituindo $s = N - i$ em 8 e usando frações parciais para fazer a integral.

($\otimes$) Faça um programa para entender a evolução do sistema a tempo discreto e a tempo contínuo.

($\otimes$) No sistema 6 faça a substituição $x_t = i_t \frac{\alpha \Delta t}{N(1+\alpha\Delta t)}$ e $s_t = N - i_t$ para obter o mapa logísitco:

$$x_{t+1} = r x_t (1 - x_t), \tag{11}$$

onde $r = 1 + \alpha\Delta t$. A diferença entre este mapa e as equações diferencias obtidas no limite $\Delta t \to 0$ é que para o mapa podemos ter comportamento muito complexo incluindo um regime de sensibilidade às condições de contorno que recebe o nome de caos. Mostre que se $\alpha\Delta t < 1$ a população se mantém positiva. Para não ter problemas com valores negativos de populações, que não ocorrem na versão 8 a tempo contínuo, podemos estudar o mapar discreto

$$i_{t+1} = N\lambda \frac{i_t}{N + (\lambda - 1) i_t} \tag{12}$$

Ao fazer a escolha 4, entre outras coisas definimos t^+ para poder tomar o limite contínuo. Equação acima pode ser obtida fazendo outra prescrição que a que leva à equação 10.

3 Modelo SIR

A variável X que descreve o estado de um agente agora toma 3 possíveis valores S, I, R, que significam susceptíveis, infectados e recuperados. Temos agora a esperança de recuperação e neste modelo o recuperado não volta a ficar infectado, desenvolvendo imunidade permanente.

Podemos ter neste modelo transições de
susceptível $\to$ infectado
infectado $\to$ recuperado
e nenhuma outra. Portanto introduzimos os parâmetros α e γ por $\otimes$:

$$P(I_{t^+}|S_t A) = \alpha \Delta t P(I_t|A), \tag{13}$$

$$P(R_{t^+}|I_t A) = \gamma \Delta t, \tag{14}$$

que modelam as taxas de infecção e cura. A probabilidade de cura não depende de quantos já estão recuperados na população.

Precisamos entender quais as asserções importantes para entender o problema e atribuir-lhes probabilidades, que neste caso serão três.

Primeiro $\otimes$:

"Agente esta recuperado em $t+1$" é equivalente a "Agente esta recuperado em t" OU ("Agente esta infestado em t" E "Agente se curou em $t^+ : t \to t+1$")

$$\begin{aligned}
P(X_{t+1} = R|A) &= P((X_t = R) \vee ((X_t = I) \wedge (X_{t^+} = R))|A) \\
P(X_{t+1} = R|A) &= P(X_t = R|A) + P(((X_t = I) \wedge (X_{t^+} = R))|A) \\
P(X_{t+1} = R|A) &= P(X_t = R|A) + P(X_t = I|A)P(X_{t^+} = R|(X_t = I) \wedge A) \\
P(R_{t+1}|A) &= P(R_t|A) + P(I_t|A)P(R_{t^+}|I_t A) \qquad (15)
\end{aligned}$$

onde a equação 15 está escrita na notação simplificada.

Segundo, a asserção

"Agente está infectado em $t+1$" é equivalente a "("agente esta infectado em t") E ("agente não recupera em t^+ ")" OU "("agente esta suscetivel em t ") E ("agente ficou doente em t_{t^+} ")" .

Traduzindo para uma linguagem mais simples:

$$P(I_{t+1}|A) = P(I_t|A)(1 - P(R_{t^+}|I_t A) + P(S_t|A)P(I_{t^+}|S_t A) \qquad (16)$$

onde usando as regras de manipulação de probabilidades, o OU em vermelho se transforma numa soma aritmética, e os E em azul em produtos.

Terceiro, a asserção, equivalente à do modelo SI ($\otimes$)

$$P(S_{t+1}|A) = P(S_t|A) - P(S_t|A)P(I_{t^+}|S_t A) \qquad (17)$$

Multiplicando pelo número total de agentes N, podemos escrever o mapa

$$\begin{aligned}
s_{t+1} &= s_t(1 - \frac{\alpha \Delta t}{N} i_t) &(18)\\
i_{t+1} &= i_t(1 - \gamma \Delta t + \frac{\alpha \Delta t}{N} s_t)) &(19)\\
r_{t+1} &= r_t + \gamma \Delta t \, i_t &(20)
\end{aligned}$$

com condições iniciais sujeitas a $s_0 + i_0 + r_0 = N$. No limite de tempo contínuo

$$\begin{aligned}
\dot{s}(t) &= -\frac{\alpha}{N} s(t) i(t) &(21)\\
\dot{i}(t) &= (\frac{\alpha}{N} s(t) - \gamma) i(t) &(22)\\
\dot{r}(t) &= r(t) + \gamma i(t) &(23)
\end{aligned}$$

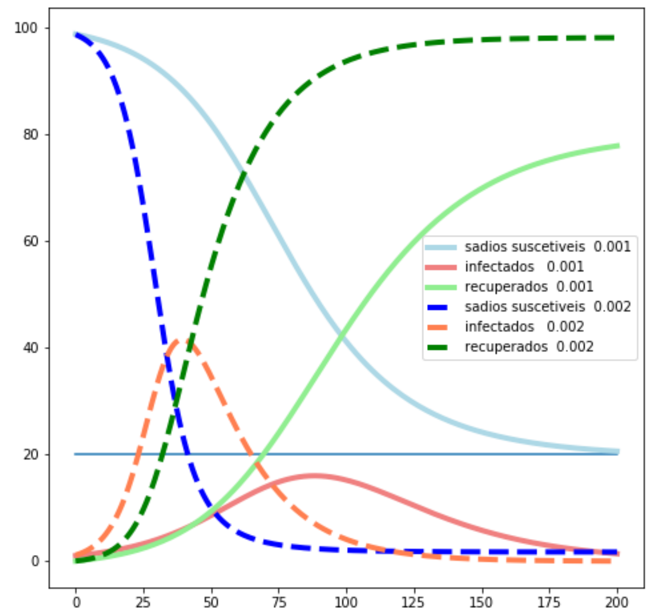

Figure 1: Modelo SIR. Distanciamento Social constante. Linha contínua α baixo, linha tracejada. α alto. A linha horizontal mostra um hipotético limiar do sistema de saúde. Se o número de infectados o ultrapssa, o sistema de saúde fica muito mais ineficiente.

Na figura 1 vemos o resultado da simulação das equações do modelo SIR (eqs 23). As curvas de infectados para dois valores do parâmetro de contágio α mostram o resultado de ter uma política de "distanciamento social". Suponha que o sistema de saúde possa lidar só com 20% da população infectada ($\otimes$). Para o valor alto de α o número de infectados passa rapidamente do valor crítico. Para α pequeno fica, o número de infectados fica abaixo do limiar da quebra do sistema de saúde.

Talvez a parte mais difícil ao usar estes modelos é ter estimativas corretas dos valores doa parâmetros. Eu não fiz nenhum esforço em usar parâmetros realistas, mas vocês podem ir atrás de dados razoáveis. Para dados realistas no Brasil talvez possam usar: `https://covid19br.github.io/`

4 $\otimes$ Taxas α e γ

Voltemos às taxas de infectação e recuperação 14.

Suponha que numa sociedade um agente tenha ν vizinhos sociais. Isto aumenta o número de variáveis para N, pois cada agente $j = 1...N$, terá seu estado descrito por X_t^j. Com probabilidade $P(I_t^j|A)$ esses vizinhos estão infectados. Queremos a probabilidade que nosso agente em foco (agente focal, f) seja contagiado por um agente j da sua vizinhança. Mostre que é razoável um modelo onde a probabilidade de "o agente focal infectou em $t^+ : t \rightarrow t+1$" é dada por

$$P(I_{t^+}^f|S_t^f A) = \sum_{j=1}^{\nu} P(I_t^j|A)P(I_{t^+}^f|I_t^j S_t^f A), \tag{24}$$

onde $P(I_{t^+}^f|I_t^j S_t A)$ é a probabilidade que tenha sido j quem contaminou f em t^+.

É claro que a probabilidade que podemos atribuir a cada agente estar contaminado depende do que sabemos sobre sua história que poderia incluir contatos anteriores e o estado do seu sistema imunológico. Se não soubermos nada que distinga um agente de outro diremos que $P(I_t^j|A)$ é independente de j. Também diremos que $P(I_{t^+}^f|I_t^j S_t^f A) = \tilde{\alpha}\Delta t$ é o mesmo para todo f e j. Discuta porque o Δt aparece. Com isso chegamos a

$$P(I_{t^+}|S_t A) \quad = \quad \alpha \Delta t P(I_t|A), \tag{25}$$

com $\alpha = \nu\tilde{\alpha}$. O valor de $\tilde{\alpha}$ depende do virus em questão. Suponha que não podemos fazer nada com ele. A única ferramenta possível é controlar o número de vizinhos ν. Sua diminuição tem recebido o nome *distanciamento social*.

O outro parâmetro γ não depende em primeira aproximação do número de vizinhos infectados, mas poderia caso ultrapasse o limiar do sistema de saúde, depender do número de agentes recuperados ou suscetíveis, pois sem eles o infectados não receberiam os mesmos cuidados.

5 Cenários

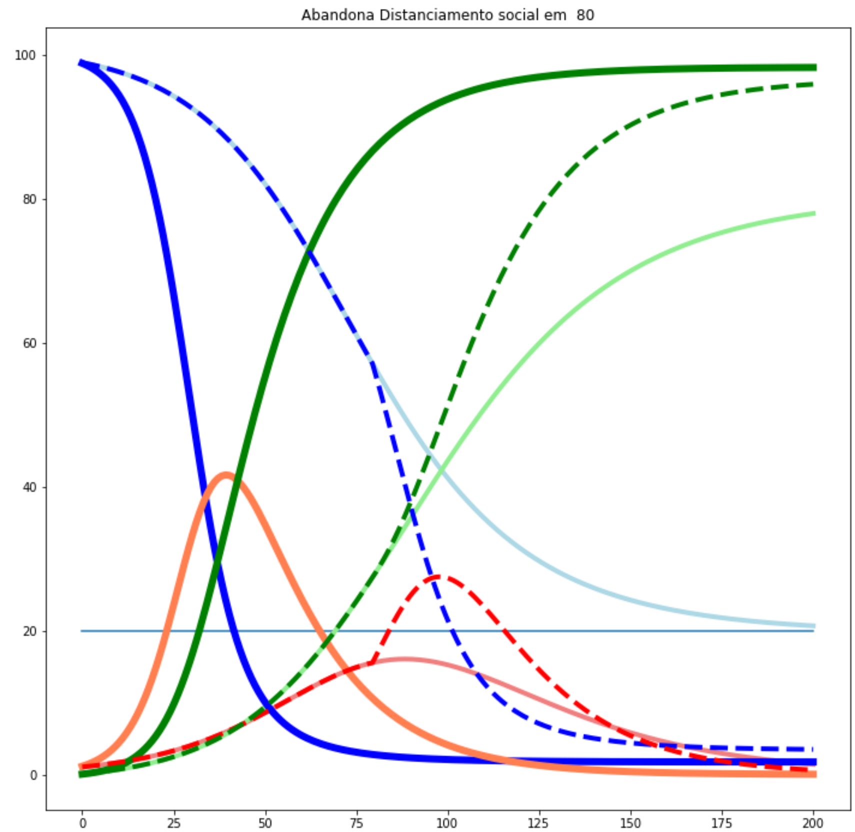

Figure 2: Modelo SIR. Estrategia mista. Abandona estrategia de distanciamento social en $t = 80$.

6 Cenários

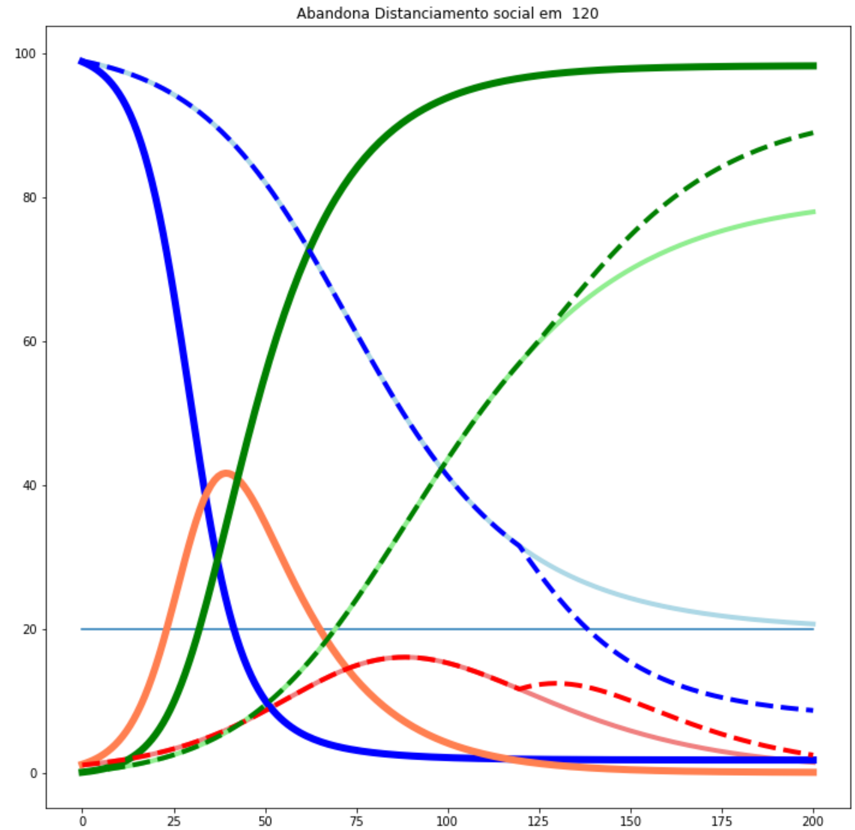

Figure 3: Modelo SIR. Estrategia mista. Abandona estrategia de distanciamento social en $t = 120$.

Suponha que você é um governante de um pais, ou seu assessor. Como todo bom estratega baseia seus conselhos e decisões em evidência. O modelo acima é muito simples mas podemos usá-los para gerar cenários.

($\otimes$) Faça um programa para simulação do modelo SIR. Coloque $N = 100$. Na figura 1 $s(0) = 99$, $i(0) = 1, r(0) = 0$. Foram usados os valores de $\alpha/N = 0.001$ e 0.002. $\gamma = .05$. Recupere estes resultados.

($\otimes$) Mude as taxas α e γ para ver os efeitos. Determine com duas casas decimais α_{20}, o maior valor de α para que não seja ultrapassado o limite de 20% de infectados. É claro que não há possibiliade de controlar α com esta precisão.

($\otimes$) Estratégia mista: Os efeitos na economia de um periodo de distanciamento social podem ser muito piores que o relaxamento de politicas de restrição de movimento dos agentes. Podemos tomar a decisão de relaxar a estratégia de distanciamento a partir de um certo instante. Encontre o valor de t_{20}, o menor tempo em que ao passar do valor alto $\alpha/N = 0.002$ para o valor baixo $\alpha/N = 0.001$, o número de infestados não passa do valor crítico.

7 Extensões

Estas generalizações são para que vocês se divirtam pensando em como melhorar os modelos. Não fiquem estressados em fazê-las. Dada as perspectivas, temo que vocês terão tempo para considerar estas variantes.

- Generalize o modelo para diferentes classes de agentes: idosos, adultos crianças
- Generalize o modelo para incluir efeitos de geografia.
- Suponha uma distribuição de probabilidades para α
- Suponha que há uma vacina disponível. Pode não ser totalmente eficiente, pode não ter o mesmo efeito para todos os agentes. Que mudança devemos fazer no modelo?
- Podemos modelar o impacto de α grande na economia para decidir quando mudar a estrategia de distanciamento social. Podemos incluir uma taxa de mortalidade devido à doença quando o sistema de saude funciona, uma taxa de mortalidade devido à doença quando o sistema não funciona, e uma taxa de mortalidade devido à ineficiência da economia.
- Suponha que tenhamos um teste para saber se uma pessoa que teve a doença assintomática está recuperada e não contamine os outros. Podemos liberar essa pessoa para trabalhar e melhorar a economia. Como seria um modelo deste tipo?
- Use dados de `https://covid19br.github.io/` para estimar os parâmetros do modelo. Você acredita que esses dados sejam realistas? Não que sejam mentira, mas todo processo de medida tem erros e certamente esses números subestimam o que deve estar occorrendo.
- Com base na análise feita, quando você acha que poderemos voltar a uma vida normal?

references

Vamos deixar as referências curtas, mas vocês podem encontrar toneladas de artigos sobre este tema:

Some discrete-time SI, SIR, and SIS epidemic models Linda J.S.Allen
https://doi.org/10.1016/0025-5564(94)90025-6