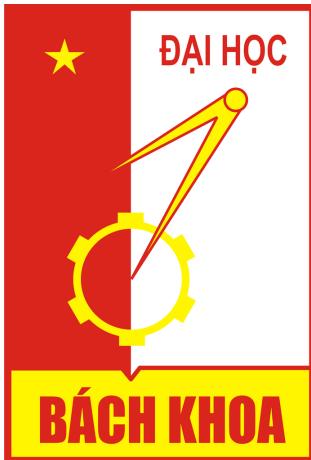


Đại học Bách Khoa Hà Nội

Viện Toán Ứng Dụng và Tin học



Các mô hình ngẫu nhiên và Ứng dụng

Mô hình ngẫu nhiên dự báo và kiểm soát
dịch bệnh COVID-19 ở Trung Quốc

Sinh viên: Nguyễn Trọng Tuấn – 20176898

Trần Đình Hoàng – 20172572

Nguyễn Công Đoàn – 20173496

Giảng viên hướng dẫn: TS. Nguyễn Thị Ngọc Anh

Hà Nội, 2020

Mục lục

1	Đặt vấn đề	2
2	Mô hình	2
3	Dữ liệu	5
4	Ước lượng tham số	5
5	Dự báo và kiến nghị	7
5.1	Hệ số lây nhiễm	7
5.2	Dự báo	7
5.3	Kiến nghị	8
6	Bàn luận	9

Danh sách hình vẽ

1	Biểu đồ dòng miêu tả sự lây nhiễm của dịch bệnh COVID-19	4
2	Dữ liệu về các ca nhiễm mới được xác nhận theo từng ngày và tổng số lượng các ca nhiễm được xác nhận của đại dịch COVID-19 từ ngày 11 tháng 1 năm 2020 đến ngày 13 tháng 2 năm 2020. Mô phỏng được thực hiện 100 lần. Các ca nhiễm mới được xác nhận theo từng ngày được biểu diễn ở hình (a) và tổng số ca nhiễm đã được xác nhận ở hình (b). Đường màu xanh đậm là dữ liệu thu thập được, các đường màu xanh nhạt biểu diễn các mô phỏng ngẫu nhiên	6
3	Dữ liệu về các ca phục hồi và tử vong mới của đại dịch COVID-19 từ ngày 11 tháng 1 năm 2020 đến ngày 13 tháng 2 năm 2020. Mô phỏng được thực hiện 100 lần. Các ca phục hồi mới theo từng ngày được biểu diễn ở hình (a), các ca tử vong mới ở hình (b). Đường màu xanh đậm là dữ liệu thu thập được, các đường màu xanh nhạt biểu diễn các mô phỏng ngẫu nhiên	6
4	Hệ số lây nhiễm hiệu quả $R_c(t)$ từ ngày 11 tháng 1 đến ngày 14 tháng 2 năm 2020. Khoảng màu xám là độ lệch chuẩn của hệ số lây nhiễm hiệu quả.	7
5	Kết quả dự báo các ca nhiễm mới được xác nhận theo từng ngày được biểu diễn ở hình (a) và tổng tất cả các ca nhiễm được xác nhận theo từng ngày được biểu diễn ở hình (b). Các dự báo được giả định rằng các biện pháp phòng ngừa dịch vẫn được áp dụng sau 350 ngày kể từ ngày 11 tháng 1 năm 2020.	8
6	Kết quả dự đoán các ca nhiễm mới theo từng ngày biểu diễn ở hình (a) và tổng số ca nhiễm theo từng ngày biểu diễn ở hình (b) với 4 giả định đã đặt ra và được mô phỏng trong 350 ngày kể từ ngày 11 tháng 1 năm 2020. Kết quả được biểu diễn là trung bình của 50 lần mô phỏng	9

Lời nói đầu

Dịch bệnh COVID-19 đã và đang tiếp tục lây lan một cách mạnh mẽ trên toàn thế giới kể từ lần đầu tiên xuất hiện ở Trung Quốc vào tháng 12 năm 2019. Sự lây lan của dịch bệnh đã ảnh hưởng đến nền kinh tế và cuộc sống của người dân của rất nhiều quốc gia trên toàn thế giới. Ở trong báo cáo này, chúng tôi sẽ trình bày và giới thiệu một mô hình dự báo sự lây lan của dịch bệnh ở Trung Quốc với sự lan truyền được cho là tuân theo phân phối nhị thức. Tham số của mô hình sẽ được học từ dữ liệu dịch bệnh của Trung Quốc từ ngày 11 tháng 1 đến ngày 13 tháng 2 năm 2020. Mô hình sẽ cho thấy rằng dịch bệnh sẽ đạt đỉnh ở cuối tháng 2 năm 2020 do đã được áp dụng các biện pháp cách ly và giãn cách xã hội. Ngoài ra, báo cáo cũng sẽ tính toán thời gian nào là hợp lý để người dân ở Trung Quốc có thể trở lại với cuộc sống bình thường.

1 Đặt vấn đề

Dịch bệnh COVID-19 do virus corona chủng mới gây ra được phát hiện lần đầu tiên ở thành phố Vũ Hán, Trung Quốc từ tháng 12 năm 2019 [2]. Dịch bệnh sau đó đã lây ra toàn Trung Quốc và khắp nơi trên thế giới [3, 4]. Một người bình thường có thể nhiễm bệnh nếu có tiếp xúc gần với người bị bệnh. Các dấu hiệu và triệu chứng được báo cáo gồm sốt, mệt mỏi, ho khan, khó thở và suy hô hấp. Hầu hết các bệnh nhân đều có các triệu chứng nhẹ và có tiên lượng tốt. Các ca bệnh dẫn tới tử vong thường là người đã có độ tuổi cao và có các bệnh lý nền như tim mạch và đái tháo đường [2]. Cho đến ngày 15 tháng 2 năm 2020 đã có 68500 trường hợp nhiễm bệnh và 1596 ca tử vong được báo cáo ở Trung Quốc [8].

Chính phủ Trung Quốc đã sử dụng nhiều biện pháp cách ly và giãn cách xã hội để giảm sự lan truyền của dịch bệnh trong cộng đồng. Diễn hình như là tìm kiếm, cách ly và theo dõi 14 ngày những người có tiếp xúc trực tiếp với người bệnh. Từ ngày 23 tháng 1 năm 2020, những nơi dịch bệnh bùng phát như Vũ Hán được phong tỏa hoàn toàn. Các trường học và doanh nghiệp được tạm ngừng để phòng tránh dịch bệnh lây lan. Người dân ở một vài quốc gia không được cấp VISA để nhập cảnh vào Trung Quốc. Nhiều chuyến bay trong và ngoài nước cũng bị hủy do dịch bệnh hoành hành.

Ở thập kỷ trước, thế giới cũng đã trải qua một số đại dịch như dịch SARS năm 2003, dịch H1N1 năm 2009, dịch H7N9 năm 2009. Do đó mà đã tạo ra nhu cầu về các dự báo sớm và sự nguy hiểm của các dịch bệnh. Các mô hình toán học nhờ đó đã được phát triển để nghiên cứu và phân tích sự lan truyền của dịch bệnh và đề ra các giải pháp phòng chống dịch bệnh. Ví dụ như Zhou và Ma [11] đã tạo ra một mô hình toán học rời rạc để nghiên cứu sự lan truyền của dịch bệnh SARS. Kết quả mô phỏng của họ cho thấy việc cách ly sớm là yếu tố quan trọng trong việc ngăn chặn dịch bệnh. Chowell [5] và Lekone [6] đã tạo ra mô hình xác định sử dụng các phương trình vi phân, và mô hình xác định SEIR để nghiên cứu sự lan truyền của dịch bệnh và sự hiệu quả của các biện pháp phòng dịch. Mô hình của họ được kiểm nghiệm dựa trên dữ liệu của đợt bùng phát dịch bệnh Ebola ở Cộng hòa Congo năm 1995.

Trong báo cáo này, chúng tôi sẽ sử dụng mô hình ngẫu nhiên rời rạc để dự báo sự lan truyền của dịch COVID-19 ở Trung Quốc. Mô hình sẽ mô phỏng sự lan truyền của dịch bệnh và sự ảnh hưởng của các biện pháp ngăn chặn dịch bệnh của chính phủ Trung Quốc. Các dữ liệu báo cáo về dịch bệnh sẽ được dùng để ước lượng tham số của mô hình. Từ đó, mô hình sẽ được áp dụng để dự báo sự lan truyền của dịch bệnh trong khoảng thời gian kế tiếp. Ngoài ra, mô hình cũng được sử dụng để tính toán thời gian hợp lý để trở lại cuộc sống bình thường của người dân Trung Quốc.

2 Mô hình

Dựa vào sự phát triển và đặc điểm của dịch bệnh COVID-19, ta sẽ sử dụng mô hình SEIR để nghiên cứu sự lan truyền của dịch. Giả sử rằng N là kích cỡ dân số của Trung Quốc (xấp xỉ 1,4 tỷ người). Tập dân số sẽ được chia thành các nhóm rời rạc như mẫn cảm $S(t)$, phơi nhiễm $E(t)$, nhiễm bệnh $I(t)$,

điều trị $H(t)$, phục hồi $R(t)$. Các trường hợp tiếp xúc với người bệnh được phát hiện sẽ được cách ly và theo dõi là $S_q(t)$ và $E_q(t)$. Trong đó, $S_q(t)$ là những người không bị nhiễm bệnh và hoàn thành thời gian cách ly; $E_q(t)$ là những người bị nhiễm bệnh trong thời gian cách ly. Ở đây, biến t là biến thời gian rời rạc để miêu tả sự tiến triển của dịch bệnh qua từng ngày. Bước thời gian là một ngày $h = 1$. Tại một thời điểm, số lượng cá thể của mỗi nhóm sẽ phụ thuộc vào số lượng của ngày hôm trước và sự thay đổi giữa các nhóm trong ngày hôm đó. Đặt $B_{ij}(t)$ là số lượng các cá thể dịch chuyển giữa các nhóm tại thời điểm t . Ta định nghĩa chi tiết như sau:

- $B_{11}(t)$ là số người mãn cảm mới bị mắc bệnh
- $B_{12}(t)$ là số người bị cách ly do tiếp xúc với người bệnh nhưng không bị nhiễm
- $B_{21}(t)$ là số người bị mắc bệnh với các triệu chứng rõ ràng
- $B_{31}(t)$ là số người được xác minh nhiễm bệnh và được đưa vào điều trị
- $B_{32}(t)$ là số ca tử vong từ những người bị nhiễm bệnh (chưa qua điều trị)
- $B_{33}(t)$ là số người bị bệnh nhưng có khả năng tự phục hồi
- $B_{41}(t)$ là số người đã hoàn thành cách ly và được thả về
- $B_{51}(t)$ là số người được đưa vào điều trị từ khu cách ly
- $B_{61}(t)$ là số người phục hồi sau quá trình điều trị
- $B_{61}(t)$ là số ca tử vong trong quá trình điều trị

Sự dịch chuyển của một người từ nhóm này sang nhóm kia được coi là một quá trình ngẫu nhiên. Khoảng thời gian một người ở trong một nhóm nhất định được coi là tuân theo phân phối mũ. Nếu ta giả sử rằng tham số của phân phối mũ là $\lambda(t)$, thì xác suất rời khỏi nhóm hiện tại trong khoảng thời gian h là $1 - \exp(-\lambda(t)h)$. Mặt khác, số lượng dòng vào và ra của một nhóm hay là $B_{ij}(t)$ được cho là tuân theo phân phối nhị thức $Bin(n, p)$. Trong đó n bằng số lượng người trong nhóm hiện tại. Sự lan truyền của dịch bệnh diễn ra thông qua tiếp xúc với người bệnh. Giả sử xác suất lan truyền là β và tỷ lệ tiếp xúc là $c(t)$ thì $\beta c(t)I(t)/N$ là tham số của phân phối mũ của xác suất các cá thể có thể bị nhiễm bệnh khi tiếp xúc với người bệnh. Giả sử q là tỷ lệ những trường hợp nhiễm bệnh bị cách ly. Dựa vào những giả định trên và mô hình SEIR đã được phát triển [6], ta có mô hình như sau:

$$\begin{aligned}
 S(t+h) &= S(t) - B_{11}(t) - B_{12}(t) + B_{41}(t) \\
 E(t+h) &= E(t) + (1-q)B_{11}(t) - B_{21}(t) \\
 I(t+h) &= I(t) + B_{21}(t) - B_{31}(t) - B_{32}(t) - B_{33}(t) \\
 S_q(t+h) &= S_q(t) + B_{12}(t) - B_{41}(t) \\
 E_q(t+h) &= E_q(t) + qB_{11}(t) - B_{51}(t) \\
 H(t+h) &= H(t) + B_{31}(t) + B_{51}(t) - B_{61}(t) - B_{62}(t) \\
 R(t+h) &= R(t) + B_{33}(t) + B_{61}(t)
 \end{aligned}$$

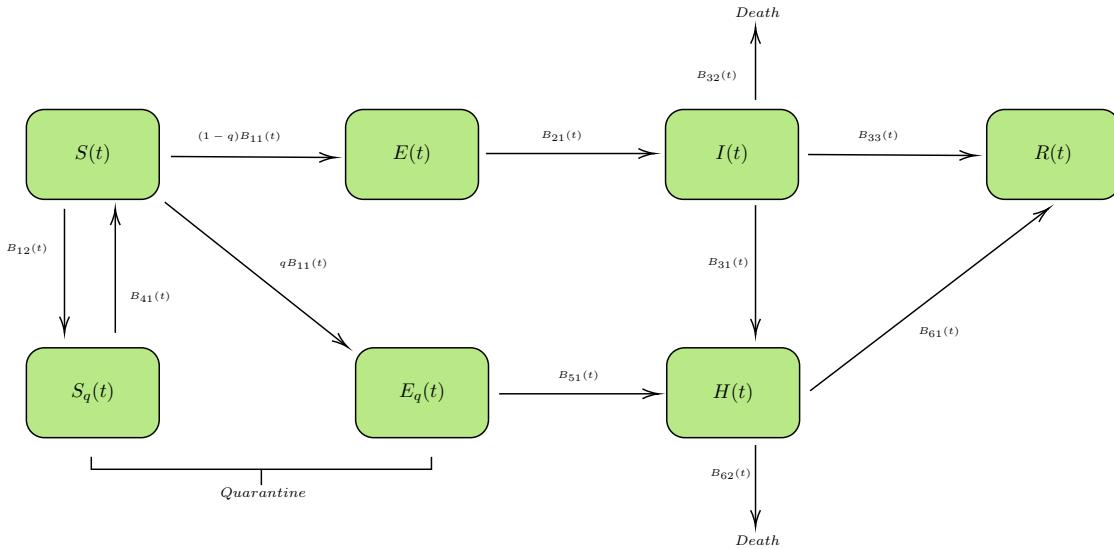
trong đó

$$\begin{aligned}
 B_{11}(t) &\sim Poi(S(t) * P_{11}(t)), \quad B_{12}(t) \sim Poi(S(t) * P_{12}(t)), \\
 B_{21}(t) &\sim Bin(E(t), P_{21}), \quad B_{31}(t) \sim Bin(I(t), P_{31}), \\
 B_{32}(t) &\sim Bin(I(t), P_{32}), \quad B_{33}(t) \sim Bin(I(t), P_{33}), \\
 B_{41}(t) &\sim Bin(S_q(t), P_{41}), \quad B_{51}(t) \sim Bin(E_q(t), P_{51}), \\
 B_{61}(t) &\sim Bin(H(t), P_{61}), \quad B_{62}(t) \sim Bin(H(t), P_{32})
 \end{aligned}$$

Các xác suất trong các phân phối nhị thức trên được mô tả chi tiết như sau:

$$\begin{aligned}
 P_{11}(t) &= 1 - \exp(-\beta c(t)I(t)h/N), & P_{12}(t) &= 1 - \exp(-c(t)q(1 - \beta)I(t)h/N), \\
 P_{21} &= 1 - \exp(-\sigma h), & P_{31} &= 1 - \exp(-\delta_I h), \\
 P_{32} &= 1 - \exp(-\alpha h), & P_{33} &= 1 - \exp(-\gamma_I h), \\
 P_{41} &= 1 - \exp(-\lambda h), & P_{51} &= 1 - \exp(-\delta_q h), & P_{61} &= 1 - \exp(-\gamma_H h)
 \end{aligned}$$

Lưu ý rằng số lượng người ở $S(t)$ xấp xỉ bằng tổng dân số N ở Trung Quốc. Do giới hạn của phân phối nhị thức là phân phối Poisson và số lượng người ở $S(t)$ là lớn. Vì vậy mà $B_{11}(t)$ và $B_{12}(t)$ tuân theo phân phối Poisson. Tổng quát của mô hình được miêu tả bằng biểu đồ 1. Các tham số của mô hình được tổng kết ở bảng 1.



Hình 1. Biểu đồ dòng miêu tả sự lây nhiễm của dịch bệnh COVID-19

Hạn chế tiếp xúc là một biện pháp phòng chống dịch bệnh hiệu quả để kiểm soát sự lây lan của dịch bệnh. Nâng cao nhận thức của cộng đồng về phòng ngừa dịch bệnh và hạn chế di chuyển hoặc đeo khẩu trang khi ở nơi công cộng. Trong mô hình này, tỷ lệ tiếp xúc $c(t)$ (tức là số lượng tiếp xúc trung bình của một người trên một đơn vị thời gian) được coi là một hàm thành phần. Nó là hằng số trước thời điểm t^* hay là trước thời điểm các biện pháp giãn cách xã hội được thực hiện. Ta giả định rằng nó sẽ giảm dần từ c_0 đến c_u . Khi đó ta có hàm $c(t)$:

$$c(t) = \begin{cases} c_0 & \text{nếu } t \leq t^* \\ (c_0 - c_u)e^{-k(t-t^*)} + c_u & \text{nếu } t > t^* \end{cases}$$

Các biện pháp cách ly được áp dụng ở Trung Quốc từ ngày 23 tháng 1 năm 2020. Dữ liệu được thu thập từ ngày 11 tháng 1 năm 2020. Do đó, $t^* = 12$.

Bảng 1: Các tham số đã được ước lượng và ban đầu của mô hình

Parameters	Definition	Baseline values	Range	Means value	(estimated)Std	Source
c_0	Tỷ lệ tiếp xúc ban đầu	31	(0, 50]	34.037	0.389	<i>MCMC</i>
c_u	Tỷ lệ tiếp xúc thấp nhất khi đã áp dụng các biện pháp giãn cách xã hội	1	(0, 50]	0.933	0.0037	<i>MCMC</i>
k	Tỷ lệ giảm của tỷ lệ tiếp xúc theo thời gian	0.1	[0, 1]	0.144	0.0035	<i>MCMC</i>
β	Xác suất lan truyền của dịch bệnh	0.095	[0, 1]	0.111	0.0015	<i>MCMC</i>
q	Tỷ lệ cách ly những trường hợp bị phơi nhiễm	0.4	[0, 1]	0.415	0.016	<i>MCMC</i>
σ	Tỷ lệ chuyển từ trạng thái E sang I	1/7	[0, 1]	1/7	—	[9]
λ	Tỷ lệ chuyển từ S_q sang S (được thả từ cách ly)	1/14	[0, 1]	1/14	—	[9]
δ_I	Tỷ lệ chuyển từ I sang H	0.1	[0, 1]	0.304	0.001	<i>MCMC</i>
δ_q	Tỷ lệ chuyển từ E_q sang H	0.42	[0, 1]	0.413	0.0116	<i>MCMC</i>
γ_I	Tỷ lệ chuyển từ I sang R (tỷ lệ những trường hợp có khả năng tự hồi phục)	0.008	[0, 1]	0.0085	0.0001	<i>MCMC</i>
γ_H	Tỷ lệ chuyển từ H sang R (tỷ lệ hồi phục sau khi được chăm sóc y tế)	0.017	[0, 1]	0.018	0.0003	<i>MCMC</i>
α	Tỷ lệ chết	0.0027	[0, 1]	0.0027	0.0001	<i>MCMC</i>
N	Toàn bộ dân số (xấp xỉ S)	—	—	1.4×10^{10}	—	<i>Data</i>
$E(0)$	Các trường hợp phơi nhiễm ban đầu	105	[1, 200]	52	0.64	<i>MCMC</i>
$I(0)$	Các trường hợp nhiễm bệnh ban đầu	54	[1, 100]	31	0.346	<i>MCMC</i>
$S_q(0)$	Các trường hợp không bị phơi nhiễm được cách ly ban đầu	—	—	734	—	<i>Data</i>
$E_q(0)$	Các trường hợp bị phơi nhiễm được cách ly ban đầu	2	[0, 100]	5	0.24	<i>MCMC</i>
$H(0)$	Các trường hợp nhiễm bệnh được chăm sóc y tế ban đầu	—	—	41	—	<i>Data</i>
$R(0)$	Các trường hợp hồi phục ban đầu	—	—	6	—	<i>Data</i>

3 Dữ liệu

Dữ liệu được thu thập từ các báo cáo của Ủy ban Y tế Quốc gia Cộng hòa Nhân dân Trung Quốc [8]. Dữ liệu sẽ được sử dụng để ước tính tham số của mô hình, chẳng hạn như số lượng các ca nhiễm mới được xác nhận, số lượng các ca phục hồi mới, số lượng tử vong mới, và số lượng người đã hoàn thành cách ly. Toàn bộ dân số N được coi bằng với dân số của Trung Quốc, $N \approx 1,4 \times 10^9$. Tổng số lượng tích lũy các ca nhiễm được xác nhận được sử dụng để so sánh với sự mô phỏng của mô hình. Dữ liệu sẽ được thu thập bắt đầu từ ngày 11 tháng 1 năm 2020 đến ngày 13 tháng 2 năm 2020.

4 Ước lượng tham số

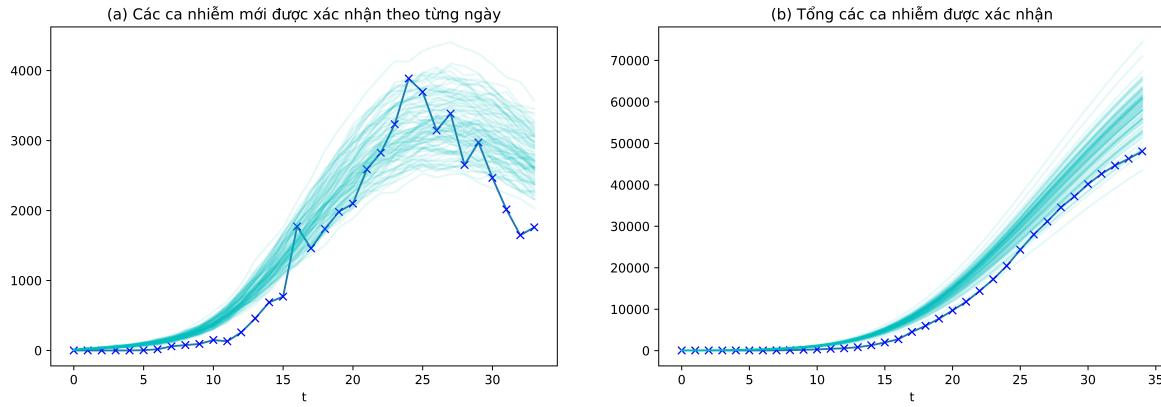
Một cách tổng quát, ta có thể sử dụng ước lượng hợp lý cực đại để ước lượng các tham số chưa biết của một mô hình ngẫu nhiên nào đó [6]. Bởi vì các biến ngẫu nhiên B_{ij} độc lập với nhau, hàm hợp lý có thể được biểu diễn như sau:

$$L(B_{11}(t), B_{12}(t), B_{21}(t), B_{31}(t), B_{32}(t), B_{33}(t), B_{41}(t), B_{51}(t), B_{61}(t), B_{62}(t) | \Theta) = \prod_{t=0}^{T_n} g_{ij}(B_{ij}(t) | \cdot)$$

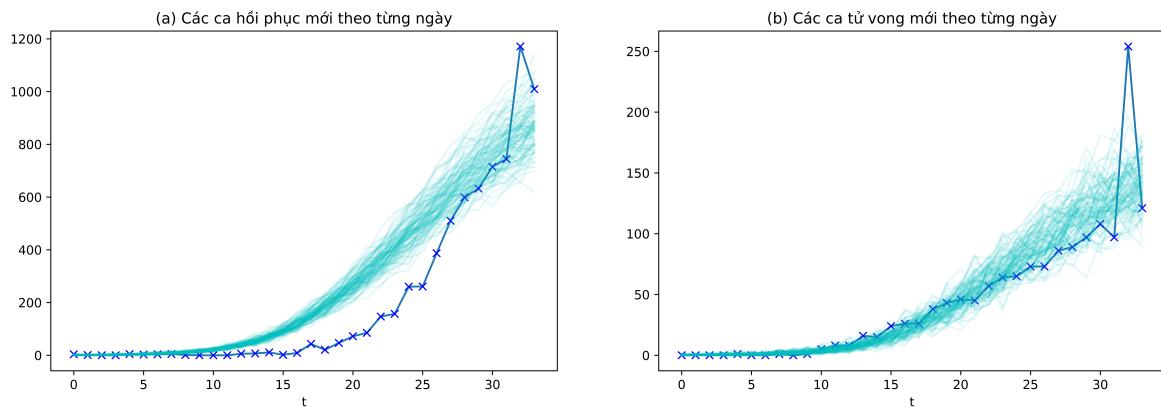
trong đó g_{ij} là hàm mô độ của phân phối nhị thức $B_{ij}(t)$. Ta có thể sử dụng ước lượng hợp lý cực đại trong trường hợp đã biết số lượng của các nhóm theo thời gian. Tuy nhiên, trong mô hình này,

ta không xác định được các trường hợp bị nhiễm bệnh như $E(t)$ và $I(t)$. Theo dữ liệu thu thập được, ta đã biết được B_{41}, B_{61}, B_{62} và số các trường hợp nhiễm bệnh mới là $B_{51} + B_{31}$ theo thời gian. Còn $B_{11}, B_{12}, B_{21}, B_{32}, B_{33}$ khó xác định. Vì vậy ta không thể sử dụng ước lượng hợp lý cực đại để ước lượng tham số của mô hình trong trường hợp này. Thay vào đó, ta sẽ sử dụng thuật toán Metropolis-Hastings (MH) để ước lượng tham số của mô hình [7, 10].

Ngoài ra, có các tham số đã biết trước như $\sigma = 1/7$ và $\lambda = 1/14$ do mỗi người đều phải cách ly 14 ngày và thời gian ủ bệnh của virus là khoảng 7 ngày [9, 1]. Từ dữ liệu thu thập được, ta biết được $H_0 = 41$, $R_0 = 6$, và tổng của S_q và E_q là 739. Ta sẽ lấy trung bình của 10000 lần lấy mẫu sau khi đã bỏ đi 5000 mẫu ban đầu. Các tham số đã được ước lượng có thể xem ở bảng 1. Hình 2 và hình 3 cho ta thấy mô hình của ta mô phỏng khá tốt về dịch bệnh COVID-19 ở Trung Quốc.



Hình 2. Dữ liệu về các ca nhiễm mới được xác nhận theo từng ngày và tổng số lượng các ca nhiễm được xác nhận của đại dịch COVID-19 từ ngày 11 tháng 1 năm 2020 đến ngày 13 tháng 2 năm 2020. Mô phỏng được thực hiện 100 lần. Các ca nhiễm mới được xác nhận theo từng ngày được biểu diễn ở hình (a) và tổng số ca nhiễm đã được xác nhận ở hình (b). Đường màu xanh đậm là dữ liệu thu thập được, các đường màu xanh nhạt biểu diễn các mô phỏng ngẫu nhiên



Hình 3. Dữ liệu về các ca phục hồi và tử vong mới của đại dịch COVID-19 từ ngày 11 tháng 1 năm 2020 đến ngày 13 tháng 2 năm 2020. Mô phỏng được thực hiện 100 lần. Các ca phục hồi mới theo từng ngày được biểu diễn ở hình (a), các ca tử vong mới ở hình (b). Đường màu xanh đậm là dữ liệu thu thập được, các đường màu xanh nhạt biểu diễn các mô phỏng ngẫu nhiên

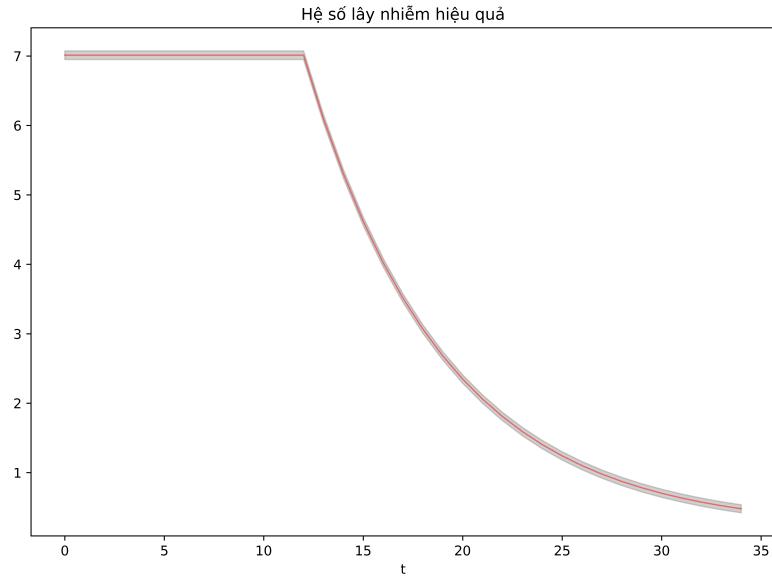
5 Dự báo và kiến nghị

5.1 Hệ số lây nhiễm

Khác với hệ số lây nhiễm cơ bản, hệ số lây nhiễm hiệu quả có thể sử dụng để tính số lượng các F_1 , tức là có tiếp xúc trực tiếp với F_0 (những người dương tính với virus corona). Hệ số này thay đổi theo thời gian do các biện pháp phòng chống dịch bệnh được áp dụng. Ta có hệ số lây nhiễm hiệu quả như sau

$$R_c(t) = \frac{\beta c(t)(1 - q)}{\delta_I + \alpha + \gamma_I}$$

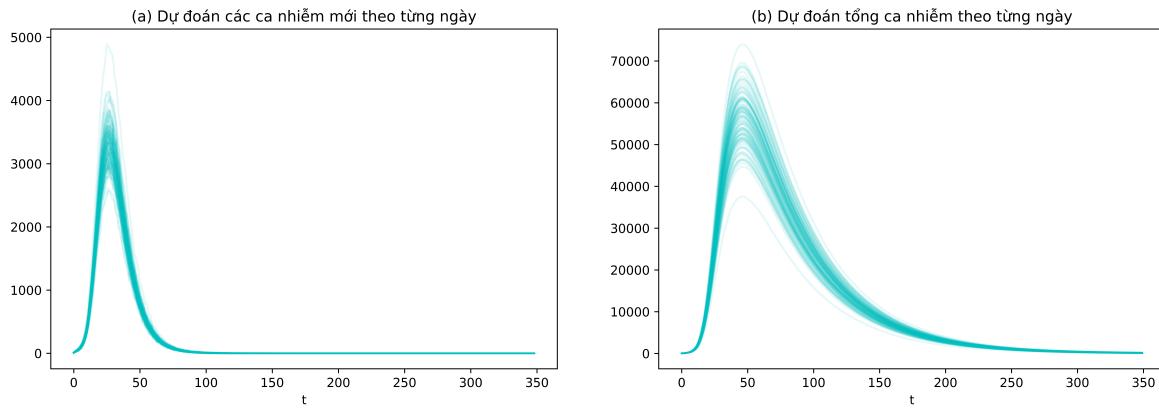
Ta thấy ở đây hệ số $c(t)$ thay đổi theo thời gian kéo theo hệ số lây nhiễm hiệu quả cũng thay đổi theo. Các tham số trong công thức trên đã được ước lượng bằng phương pháp Markov Chain Monte Carlo (MCMC). Kết quả tính toán hệ số $R_c(t)$ theo thời gian được biểu diễn bằng hình 4.



Hình 4. Hệ số lây nhiễm hiệu quả $R_c(t)$ từ ngày 11 tháng 1 đến ngày 14 tháng 2 năm 2020. Khoảng màu xám là độ lệch chuẩn của hệ số lây nhiễm hiệu quả.

5.2 Dự báo

Ta đã giả định rằng tỷ lệ tiếp xúc sẽ giảm dần theo thời gian nhờ việc áp dụng các biện pháp phòng ngừa dịch bệnh. Việc ước lượng các tham số của mô hình cho ta thấy tỷ lệ tiếp xúc ban đầu là 34.03, và tỷ lệ tiếp xúc khi đã áp dụng các biện pháp giãn cách xã hội sẽ giảm dần theo thời gian xuống còn 0.93. Điều này cho thấy rằng các biện pháp giãn cách xã hội có tác dụng phòng chống dịch bệnh rất hiệu quả. Từ các tham số đã được ước lượng và các giả định của mô hình, ta có thể mô phỏng sự lây lan của dịch bệnh trong tương lai. Trong hình 5 (a) và (b), các mô phỏng cho thấy sự thay đổi của các số liệu theo từng ngày trong vòng 350 ngày kể từ ngày 11 tháng 1 năm 2020. Số lượng ca nhiễm mới theo từng ngày có dấu hiệu giảm dần và tiệm cận về không trong khoảng 102 – 119 ngày kể từ ngày 11 tháng 1. Cùng với đó, tổng số lượng các ca nhiễm theo từng ngày sẽ đạt đỉnh vào khoảng ngày thứ 49 – 52 kể từ ngày 11 tháng 1.



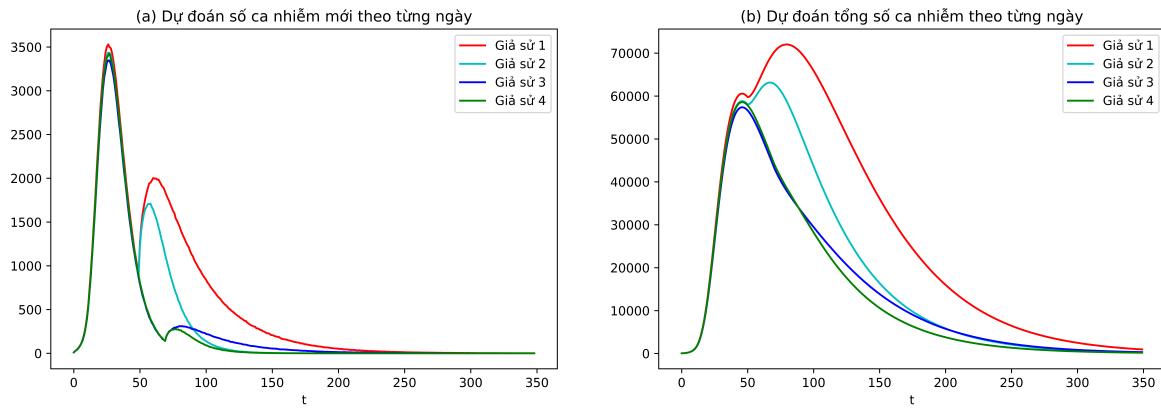
Hình 5. Kết quả dự báo các ca nhiễm mới được xác nhận theo từng ngày được biểu diễn ở hình (a) và tổng tất cả các ca nhiễm được xác nhận theo từng ngày được biểu diễn ở hình (b). Các dự báo được giả định rằng các biện pháp phòng ngừa dịch vẫn được áp dụng sau 350 ngày kể từ ngày 11 tháng 1 năm 2020.

5.3 Kiến nghị

Một câu hỏi được người dân và chính phủ rất quan tâm đó là việc áp dụng các biện pháp phòng ngừa dịch bệnh sẽ diễn ra trong bao lâu? Ta không thể phủ nhận rằng các biện pháp này đã ảnh hưởng rất nhiều đến cuộc sống và kinh tế của toàn xã hội. Nếu mọi người trở lại cuộc sống thường nhật, tỷ lệ tiếp xúc sẽ tăng, kéo theo đó là hệ số lây nhiễm hiệu quả cũng tăng theo. Điều đó dẫn tới nguy cơ bùng phát dịch bệnh lần thứ hai. Ta sẽ xem xét bốn trường hợp sau đây:

- 1) Giả sử thời điểm nối lỏng các biện pháp cách ly là ngày 1 tháng 3 năm 2020, tức là 50 ngày sau khi dịch bệnh bùng phát. Tham số $c(t) = 3$ khi $t > 50$.
- 2) Giả sử thời điểm nối lỏng các biện pháp cách ly là ngày 1 tháng 3 năm 2020, tức là 50 ngày sau khi dịch bệnh bùng phát. Tham số $c(t) = 1.5$ khi $t > 50$.
- 3) Giả sử thời điểm nối lỏng các biện pháp cách ly là ngày 20 tháng 3 năm 2020, tức là 70 ngày sau khi dịch bệnh bùng phát. Tham số $c(t) = 3$ khi $t > 70$.
- 4) Giả sử thời điểm nối lỏng các biện pháp cách ly là ngày 20 tháng 3 năm 2020, tức là 70 ngày sau khi dịch bệnh bùng phát. Tham số $c(t) = 1.5$ khi $t > 70$.

Ta áp dụng mô hình đã xây dựng với các tham số đã được ước lượng trên bộ dữ liệu đã được thu thập để mô phỏng bốn trường hợp trên. Kết quả của mô phỏng được thể hiện trong Hình 6, bao gồm dự đoán tổng các ca nhiễm mới theo từng ngày và tổng các ca nhiễm đang được điều trị theo từng ngày. Từ kết quả của mô phỏng, ta thấy rằng các biện pháp phòng ngừa dịch bệnh như mỗi người cách nhau hai mét là cần thiết để phòng tránh sự bùng phát của dịch bệnh. Tức là, khi trở lại với cuộc sống thường nhật, người dân vẫn cần có sự tự phòng bị với môi trường xung quanh để tránh bị nhiễm bệnh. Ta cũng thấy rằng, việc nối lỏng các biện pháp cách ly vào ngày 20 tháng 3 năm 2020 là tốt hơn với số ca nhiễm mới chỉ tăng một lượng nhỏ, nhưng sau đó sẽ giảm xuống một cách nhanh chóng. Do đó, việc kéo dài thời gian áp dụng các biện pháp cách ly sẽ làm giảm xác suất để một đợt bùng phát dịch thứ hai có thể xảy ra.



Hình 6. Kết quả dự đoán các ca nhiễm mới theo từng ngày biểu diễn ở hình (a) và tổng số ca nhiễm theo từng ngày biểu diễn ở hình (b) với 4 giả định đã đặt ra và được mô phỏng trong 350 ngày kể từ ngày 11 tháng 1 năm 2020. Kết quả được biểu diễn là trung bình của 50 lần mô phỏng

6 Bàn luận

Bằng việc xây dựng một mô hình ngẫu nhiên rời rạc để nghiên cứu sự lây truyền của dịch bệnh COVID-19, ta đã có thể quan sát được tính bất định trong sự lây lan của dịch bệnh. Trong mô hình này, chúng tôi đã giả định rằng tỷ lệ tiếp xúc với người bị nhiễm bệnh sẽ giảm theo dần theo thời gian nhờ vào việc áp dụng các biện pháp phòng dịch hiệu quả. Dữ liệu để phục vụ cho việc ước lượng tham số của mô hình được chúng tôi thu thập từ ngày 11 tháng 1 đến ngày 13 tháng 2 năm 2020. Ngoài ra, để ước lượng được tham số của mô hình, chúng tôi đã sử dụng phương pháp Markov Chain Monte Carlo (MCMC).

Giá trị ban đầu của tỷ lệ tiếp xúc khi chưa được áp dụng các biện pháp phòng dịch là 34.02 và sau đó giảm dần theo thời gian xuống còn 0.93. Điều này cho thấy các biện pháp phòng chống dịch bệnh là hiệu quả. Hệ số lây nhiễm hiệu quả được cho là rất cao trong những giai đoạn đầu của dịch, tương ứng cho việc dịch bệnh COVID-19 có thể bùng phát một cách nhanh chóng. Tuy nhiên, nó cũng giảm dần theo thời gian xuống dưới 1, tức là sự bùng phát của dịch bệnh sẽ được dập tắt.

Sự mô phỏng của mô hình cho thấy nó không phù hợp với các số liệu đã được báo cáo trong giai đoạn đầu của dịch bệnh, nhưng ở những giai đoạn sau nó mô phỏng khá tốt với các số liệu mô phỏng phù hợp với các số liệu ở thực tế. Một minh chứng thực tế là dịch bệnh COVID-19 ở Trung Quốc đã được kiểm soát và số ca nhiễm còn lại đã giảm tiệm cận về không trong những ngày cuối tháng 4 và đầu tháng 5 năm 2020, điều mà đã được mô hình dự báo trước đó. Để xem xét thời điểm thích hợp để người dân có thể trở lại cuộc sống bình thường, chúng tôi đã xem xét bốn trường hợp khác nhau. Theo đó, chúng tôi giả sử tỷ lệ tiếp xúc $c(t) = 3$ nếu như không có các biện pháp phù hợp để bảo vệ người dân khỏi dịch bệnh khi trở lại cuộc sống thường nhật. Trong trường hợp này, mô phỏng cho thấy rằng dịch bệnh sẽ có một đợt bùng phát thứ hai nếu người dân trở lại cuộc sống bình thường vào ngày 1 tháng 3 năm 2020. Với cùng tỷ lệ tiếp xúc đó, tình hình dịch bệnh sẽ tốt hơn nếu người dân trở lại cuộc sống bình thường vào ngày 20 tháng 3 năm 2020.

Tóm lại, mô phỏng của chúng tôi cho thấy rằng tỷ lệ tiếp xúc là yếu tố chính trong việc kiểm soát dịch bệnh COVID-19. Việc không kiểm soát được tỷ lệ tiếp xúc như ở Mỹ và các nước châu Âu đã làm cho dịch bệnh bùng phát mạnh mẽ và chưa có điểm dừng cho tới thời điểm hiện tại. Ngoài ra, các biện pháp bảo vệ cá nhân như đeo khẩu trang và hạn chế di chuyển khi có dịch bệnh là một cách hiệu quả để giảm nguy cơ lây nhiễm cộng đồng. Thực tế cho thấy rằng có nhiều khả năng gây ra các đợt bùng phát tiếp theo mà mô hình của chúng tôi chưa thể dự báo trước được, ví dụ như các biến thể mới của virus corona. Vì vậy mà mọi người nên cảnh giác trước sự lây lan của dịch bệnh này.

Tài liệu

- [1] World Health Organization (WHO). URL: [https://www.who.int/news-room/detail/23-01-2020-statement-on-the-meeting-of-the-international-health-regulations-\(2005\)-emergency-committee-regarding-the-outbreak-of-novel-coronavirus-\(2019-ncov\)](https://www.who.int/news-room/detail/23-01-2020-statement-on-the-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-outbreak-of-novel-coronavirus-(2019-ncov)).
- [2] World Health Organization (WHO). *Coronavirus*. URL: <https://www.who.int/health-topics/coronavirus>.
- [3] World Health Organization (WHO). *Disease Outbreak News*. URL: https://www.who.int/csr/don/archive/disease/novel_coronavirus/en/.
- [4] World Health Organization (WHO). *Situation reports*. URL: <https://covid19.who.int/>.
- [5] G. Chowell et al. “Model Parameters and Outbreak Control for SARS”. In: *Emerg. Infect. Dis.* 10 (2004), pp. 1258–1263.
- [6] P. Lekone and B. Finkenstädt. “Statistical Inference in a Stochastic Epidemic SEIR Model with Control Intervention: Ebola as a Case Study”. In: *Biometrics* 62 (2006), pp. 1170–1177.
- [7] A. Morton and B. Finkenstädt. “Discrete time modelling of disease incidence time series by using Markov chain Monte Carlo methods”. In: *J. R. Stat. Soc* 54 (2005), pp. 575–594.
- [8] National Health Commission of the People’s Republic of China. URL: http://www.nhc.gov.cn/xcs/yqtb/list_gzbd.shtml.
- [9] B. Tang et al. “Estimation of the transmission risk of the 2019-nCoV and its implication for public health interventions”. In: *J. Clin. Med.* 9 (2020), p. 462.
- [10] S. Tang et al. “Community-based measures for mitigating the 2009 H1N1 pandemic in China”. In: *PLoS One* 5 (2010), e10911.
- [11] Y. Zhou, Z. Ma, and F. Brauer. “A Discrete Epidemic Model for SARS Transmission and Control in China”. In: *Math. Comput. Model.* 40 (2004), pp. 1491–1506.