

Математическое моделирование динамики развития инфекционных заболеваний с использованием библиотек Python Numpy и Matplotlib

О. И. Канищева, email: oleka_olesya@mail.ru^{1,2}

В. С. Боровков, email: mirrox2@yandex.ru¹

¹ ВУНЦ ВВС «ВВА им. проф. Н.Е. Жуковского и Ю.А. Гагарина»
(г. Воронеж)

² Воронежский государственный университет

***Аннотация.** В данной работе рассмотрены классические математические модели распространения эпидемий SIR, представляющие собой системы дифференциальных уравнений, фазовые переменные которых являются долями некоторой однородной популяции, а сами уравнения отображают динамику перемещения особей популяции между группами. Представлены результаты численного моделирования развития заболеваний для SEIR модели с использованием языка программирования Python, проиллюстрирована динамика развития инфекционного заболевания.*

***Ключевые слова:** математическое моделирование, SEIR модель, SIRS модель, SEIR модель, MSEIR модель, моделирование эпидемий, популяция, эпидемия, Python.*

Введение

В медицине под эпидемией понимают прогрессирующее распространение инфекционного заболевания среди людей, значительно превышающее обычно регистрируемый на данной территории уровень заболеваемости и способное стать причиной чрезвычайной ситуации [1]. В случае распространения эпидемии на большие территории или территории многих стран говорят о пандемии.

Изучение механизмов развития и распространения эпидемий является важным способом борьбы с заболеваниями наряду с поиском новых лекарственных средств, вакцинацией и профилактическими мерами.

Первую попытку использовать математический аппарат для исследования механизмов распространения заболеваний предпринял Д. Бернулли. Следующий шаг был сделан У. Фарром, применившим нормальное распределение для анализа смертности людей от оспы.

Основываясь на работах предшественников, британские ученые А. Кермак и У. Маккендрик разработали широко используемую на сегодняшний день модель SIR [2]. Эта аббревиатура происходит от английских слов Susceptible – Infected – Recovered, буквально означающих «восприимчивые – инфицированные – выздоровевшие». Под «восприимчивыми» подразумеваются еще не инфицированные организмы.

В рамках этой модели с помощью систем дифференциальных уравнений (при условии непрерывности времени и популяции большого объема) или разностных уравнений (при дискретном времени и ограниченной популяции) описывается динамика распространения инфекционного заболевания.

1. Математические модели развития эпидемиологической ситуации

Модель SIR

SIR–модель получила популярность в силу простоты построения и использования. Ее применение позволяет точно моделировать эпидемии гриппа и других инфекционных заболеваний в больших городах, вводить новые параметры и анализировать разные сценарии развития болезни [3].

Система уравнений SIR имеет вид:

$$\begin{cases} \frac{dS(t)}{dt} = -\frac{\beta \cdot I(t) \cdot S(t)}{N}, \\ \frac{dI(t)}{dt} = \frac{\beta \cdot I(t) \cdot S(t)}{N} - \gamma \cdot I(t), \\ \frac{dR(t)}{dt} = \gamma \cdot I(t), \end{cases} \quad (1)$$

где $S(t)$ – численность восприимчивых индивидов в момент времени t ; $I(t)$ – численность инфицированных индивидов в момент времени t ; $R(t)$ – численность переболевших индивидов в момент времени t ; β – коэффициент интенсивности контактов индивидов с последующим инфицированием; γ – коэффициент интенсивности выздоровления инфицированных индивидов; N – общая численность популяции.

Первое уравнение системы (1) означает, что изменение числа здоровых индивидов (и при этом восприимчивых к заболеванию) уменьшается со временем пропорционально числу контактов с

инфицированными. После контакта происходит заражение, восприимчивый переходит в состояние инфицированного.

Второе уравнение системы (1) показывает, что скорость увеличения числа заразившихся растет пропорционально числу контактов здоровых и инфицированных людей и уменьшается по мере выздоровления последних.

Третье уравнение системы (1) демонстрирует, что число выздоровевших в единицу времени пропорционально числу инфицированных. Иначе говоря, каждый заболевший через некоторое время должен выздороветь.

Таким образом, мы видим, что заболевание в модели SIR развивается по схеме «восприимчивые становятся инфицированными, потом выздоравливают». Условие

$$\frac{dS(t)}{dt} + \frac{dI(t)}{dt} + \frac{dR(t)}{dt} = 0 \quad (2)$$

описывает неизменность численности популяции (и не учитывает случаи смерти от заболевания).

Количество одномоментно болеющих, определяется параметром $R_0 = \frac{\beta}{\gamma}$. Эта величина называется «базовым коэффициентом воспроизведения».

Что касается задания начальных условий для системы (1), можно заметить, что часто изначально в популяции отсутствуют особи с иммунитетом к заболеванию, т. е. $R(0) = 0$. Систему уравнений (1) также называют системой Кермака-МакКендрика [4].

SIR-модель перестает работать в случае необходимости учитывать неоднородность популяции (например, различную плотность населения в разных регионах), разные пути передачи инфекции и факторы случайности, значимые в малых популяциях и на начальной фазе распространения заболевания [5].

Развитием модели SIR стали, в частности, следующие модели:

– SIRS – «восприимчивые – инфицированные – выздоровевшие – восприимчивые»: модель описания динамики заболеваний с временным иммунитетом (выздоровевшие индивиды со временем снова становятся восприимчивыми);

– SEIR – «восприимчивые – контактные (Exposed) – инфицированные – выздоровевшие»: модель для описания распространения заболеваний с инкубационным периодом;

– SIS – «восприимчивые – инфицированные – восприимчивые»: модель для распространения заболевания, к которому не вырабатывается иммунитет;

– MSEIR – «наделенные иммунитетом от рождения (Maternally derived immunity) – восприимчивые – контактные – инфицированные – выздоровевшие»: модель, учитывающая иммунитет детей, приобретенный внутриутробно.

Модель SEIR

Именно по этой модели развиваются самые опасные эпидемии, поскольку длительный инкубационный период может препятствовать своевременному обнаружению заболевания. В этом случае есть риск, что заболевание охватит значительное число индивидуумов в популяции.

Инфекция развивается по схеме «восприимчивые – контактные – инфицированные – выздоровевшие» и описывается системой уравнений:

$$\left\{ \begin{array}{l} \frac{dS(t)}{dt} = \mu \cdot N - \mu \cdot S(t) - \beta \cdot \frac{I(t)}{N} \cdot S(t), \\ \frac{dE(t)}{dt} = \beta \cdot \frac{I(t)}{N} \cdot S(t) - (\mu + \alpha) \cdot E(t), \\ \frac{dI(t)}{dt} = \alpha \cdot E(t) - (\gamma + \mu) \cdot I(t), \\ \frac{dR(t)}{dt} = \gamma \cdot I(t) - \mu \cdot R(t), \end{array} \right. \quad (3)$$

где μ – уровень смертности; α – величина, обратная среднему инкубационному периоду заболевания; $E(t)$ – численность индивидов – носителей заболевания в момент времени t .

Как и в модели SIR (1), первое уравнение системы (3) означает, что изменение числа здоровых (и при этом восприимчивых к заболеванию) индивидуумов уменьшается со временем пропорционально числу контактов с инфицированными. После заражения здоровый индивид переходит в состояние контактного по данному заболеванию, или носителя инфекции.

Второе уравнение системы (3) вносит задержку по времени при переходе из состояния контактного в состояние инфицированного (больного). Это происходит через некоторое время, равное инкубационному периоду болезни.

Третье уравнение системы (3) описывает переход из состояния «контактный» в состояние «инфицированный».

Четвертое уравнение системы (3) демонстрирует, что число выздоровевших в единицу времени пропорционально числу инфицированных. При этом в каждом состоянии индивидум может погибнуть, что учитывает коэффициент μ в каждом уравнении.

Иначе говоря, в каждый момент времени каждый индивидум с определенной вероятностью может заразиться, через некоторое время – заболеть, а затем выздороветь либо погибнуть.

Численность популяции $N = S(t) + E(t) + I(t) + R(t)$ при этом не является постоянной с течением времени. Интенсивность эпидемии описывает базовый коэффициент воспроизведения: $R_0 = \frac{\alpha}{\mu + \alpha} \cdot \frac{\beta}{\mu + \gamma}$.

Модель SIS

Модель «восприимчивые – инфицированные – восприимчивые» применима при анализе распространения заболеваний, к которым не вырабатывается иммунитет, например гриппа и ОРВИ. Она описывается следующей системой уравнений:

$$\begin{cases} \frac{dS(t)}{dt} = -\frac{\beta \cdot S(t) \cdot I(t)}{N} + \gamma \cdot I(t), \\ \frac{dI(t)}{dt} = \frac{\beta \cdot S(t) \cdot I(t)}{N} - \gamma \cdot I(t), \end{cases} \quad (4)$$

Вместе первое и второе уравнение системы (4) означают, что число здоровых и больных в сумме не меняется, а число заражений пропорционально числу контактов здоровых и больных.

Второе уравнение системы (4) описывает изменение числа заболевших в единицу времени, которое пропорционально числу заражений (числу контактов здоровых и инфицированных индивидуумов) за вычетом числа выздоровлений.

Модель MSEIR

Эта модель, построенная для заболевания с инкубационным периодом и учитывающая иммунитет детей, приобретенный внутриутробно, – одна из самых сложных для анализа в силу наличия большого числа независимых параметров. Система уравнений для нее имеет вид:

$$\begin{cases} \frac{dM(t)}{dt} = B(t) - \delta \cdot M(t) - \mu \cdot M(t), \\ \frac{dS(t)}{dt} = \delta \cdot M(t) - \beta \cdot S(t) \cdot I(t) - \mu \cdot S(t), \\ \frac{dE(t)}{dt} = \beta \cdot S(t) \cdot I(t) - (\varepsilon + \mu) \cdot E(t), \\ \frac{dI(t)}{dt} = \varepsilon \cdot E(t) - (\gamma + \mu) \cdot I(t), \\ \frac{dR(t)}{dt} = \gamma \cdot I(t) - \mu \cdot R(t), \end{cases} \quad (5)$$

где $M(t)$ – численность индивидов с приобретенным внутриутробно иммунитетом.

От ранее рассмотренных моделей эта система (5) отличается тем, что учитывает рождение детей, вероятность заражения которых растет со временем по мере утраты ими иммунитета, приобретенного внутриутробно. Эти зависимости описаны в первых двух уравнениях системы (5).

Приобретенный внутриутробно иммунитет может быть не у всех появившихся на свет детей, но вакцинацией можно охватить сто процентов рожденных младенцев. Введение в математическую модель этого параметра приводит к качественному изменению картины развития эпидемий.

Система дифференциальных уравнений для этой модели будет следующей:

$$\begin{cases} \frac{dS(t)}{dt} = \mu \cdot N \cdot (1 - P) - \mu \cdot S(t) - \beta \cdot \frac{I(t)}{N} \cdot S(t), \\ \frac{dI(t)}{dt} = \beta \cdot \frac{I(t)}{N} \cdot S(t) - (\gamma + \mu) \cdot I(t), \\ \frac{dV(t)}{dt} = \mu \cdot N \cdot P - \mu \cdot V(t), \end{cases} \quad (6)$$

где P – доля привитых младенцев ($0 < P < 1$).

Первые два уравнения системы (6) повторяют модель SIR с учетом того, что вероятность заражения привитых детей равна нулю, а значит, вероятность заражения равна вероятности, что ребенок не привит, и, в свою очередь, равна $(1 - P)$.

Последнее уравнение системы (6) учитывает смертность от других причин и позволяет рассчитать полную численность популяции.

Существует ещё много различных модификаций SIR моделей. Все они, включая саму SIR, являются представителями целого класса моделей, которые называют «компарментальными эпидемиологическими моделями». Сложность таких моделей не ограничена разделением популяции на три или четыре группы. Эти модели могут учитывать самые разные сценарии: введение карантина (SIQR, добавляется группа Q – quarantine), потеря иммунитета (SIRS, переход с некоторой вероятностью из R обратно в S), различные варианты течения болезни (несколько групп I: I1, I2, ..., в каждую из которых своя вероятность попадания восприимчивых особей) и т.д. [6].

Исследование динамики передачи любого инфекционного заболевания зависит от характера данных и разработки модели, которая наилучшим образом описывает сценарий вспышки. Математическое и имитационное моделирование исследования, основанные на эпидемиологических данных, могут помочь оценить эффективность мер контроля и могут быть использованы для оценки эффективности вакцины.

2. Реализация модели динамики развития инфекционного заболевания с помощью Python

Реализация модели развития инфекционного заболевания проводилась на примере SEIR модели (рис. 1), которая учитывает инкубационный период (E – exposed, индивиды болеют, но не заразны и со временем полностью заболеют). В такой модели заражение восприимчивых происходит таким же способом как в модели SIR, но попадают такие особи не в группу I, а в группу E. А из E с определённой вероятностью (α – число обратное длительности инкубационного периода) происходит переход уже в I.

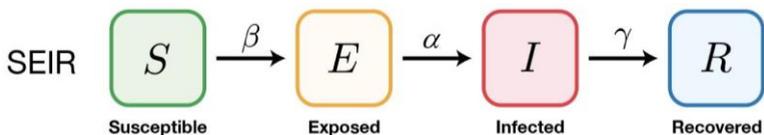


Рис. 1. Схема принципа работы SEIR модели

Для реализации модели на Python использовались следующие библиотеки: Numpy – для создания одномерного массива для хранения и расчета данных о количестве зараженных пациентов в день и Matplotlib – для графического представления количества инфицированных пациентов с количеством дней.

На рис. 2 показан импорт библиотеки Numpy и Matplotlib в Python для определения и решения системы дифференциальных уравнений (3), зависящей от времени. Эти библиотеки часто используются в различных научных вычислениях.

```
1 import numpy as np
2 import matplotlib as plt
```

Рис. 2. Подключение библиотеки Numpy и Matplotlib

На рис. 3 приведен листинг программы, реализующей SEIR модель на языке программирования Python.

```
3 # численность популяции
4 N = 1e7 + 10 + 5
5 # время (дни)
6 T = 170
7 # восприимчивые
8 s = np.zeros([T])
9 # находящиеся в инкубационном периоде
10 e = np.zeros([T])
11 # инфицированные
12 i = np.zeros([T])
13 # выздоровевшие
14 r = np.zeros([T])
15 # скорость распространения заболевания
16 lamda = 0.5
17 # скорость восстановления
18 gamma = 0.0821
19 # открытый период
20 sigma = 1 / 4
21 # первоначально инфицированные
22 i[0] = 10.0 / N
23 s[0] = 1e7 / N
24 e[0] = 40.0 / N
25 for t in range(T-1):
26     s[t + 1] = s[t] - lamda * s[t]*i[t]
27     e[t + 1] = e[t] + lamda * s[t]*i[t] - sigma * e[t]
28     i[t + 1] = i[t] + sigma * e[t] - gamma * i[t]
29     r[t + 1] = r[t] + gamma * i[t]
30 from matplotlib import pyplot as plt
31 fig, ax = plt.subplots(figsize=(10,6))
32 ax.plot(s, c='b', lw=2, label='Восприимчивые (S)')
33 ax.plot(e, c='orange', lw=2, label='Находящиеся в инкубационном периоде (E)')
34 ax.plot(i, c='r', lw=2, label='Инфицированные (I)')
35 ax.plot(r, c='g', lw=2, label='Выздоровевшие (R)')
36 ax.set_xlabel('Дни', fontsize=20)
37 ax.set_ylabel('Динамика распространения болезни', fontsize=20)
38 ax.grid(1)
39 plt.xticks(fontsize=20)
40 plt.yticks(fontsize=20)
41 plt.legend();
42 plt.show()
```

Рис. 3. Листинг программы на Python для SEIR модели

На рис. 4 приведены графики решения для SEIR модели, иллюстрирующие динамику развития инфекционного заболевания.

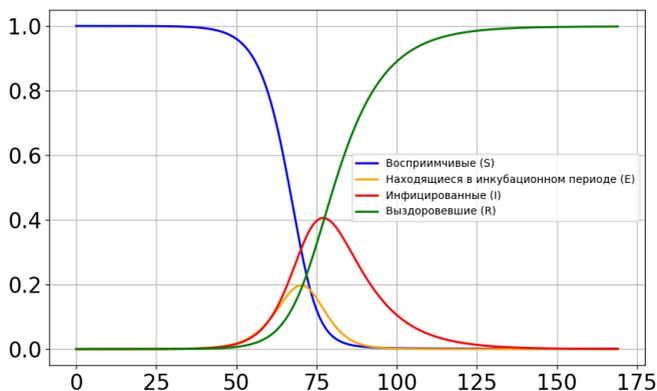


Рис. 4. Динамика изменения доли восприимчивых, контактных, инфицированных и выздоровевших по SEIR модели

Из полученных результатов видно, что чем меньше величина параметров γ , β , тем дольше будет длиться эпидемия. При этом, если оба параметра имеют небольшую величину, то эпидемия не случится вовсе.

Заключение

Проведенные в работе исследования показывают, что SEIR модель может использоваться в условиях ограниченных размеров популяций, надежности и при недостатке статистических данных. Они дают не только качественные, но и количественные прогнозы распространения инфекции. Дальнейшие исследования планируется провести для моделей, учитывающих пространственную изменчивость и неполную восприимчивость населения к инфекции.

Список литературы

1. Фрумер, А. Л. Эпидемические и эндемические болезни // Энциклопедический словарь Брокгауза и Ефрона : в 86 т. (82 т. и 4 доп.). – СПб., 1890-1907.
2. Kermack W. O., A Contribution to the Mathematical Theory of Epidemics / W. O. Kermack, A. G. McKendrick // Proc. R. Soc. Lond. Series A. – Vol. 115. – No. 772. (August 1, 1927). – pp. 700-721.

3. Kermack, W.O. Contributions to the mathematical theory of epidemics –I / W.O. Kermack, A.G. McKendrick // *Bulletin of Mathematical Biology*. – 1991. – V. 53. – P. 33-55.

4. Интегрированные математические модели, описывающие сложные биологические процессы / Е. Л. Мищенко [и др.] // *Биофизика*. – 2017. – Т. 62. – № 5. – С. 949-968.

5. Ризниченко, Г. Ю. Математическое моделирование биологических процессов. Модели в биофизике и экологии : учеб. пособие / Г. Ю. Ризниченко. – 2-е изд., перераб. и доп. – М. : Юрайт, 2019. – 181 с.

6. Brauer F. *Mathematical models in population biology and epidemiology* / F. Brauer, C. Castillo-Chavez // 2ed. – Springer, 2012. – 522 p.