МИНОБРНАУКИ РОССИИ ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ «ВОРОНЕЖСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ» (ФГБОУ ВО «ВГУ»)

Математический факультет Кафедра математического моделирования

Анализ математических моделей инфодемии

Бакалаврская работа

Направление 01.03.04 Прикладная математика

Профиль Применение математических методов к решению инженерных и экономических задач

Зав. кафедрой	 д.фм.н., доц.	М.Ш. Бурлуцкая	04.06. 2024 г
Обучающийся		Ю.М. Кудашов	
Руководитель	 к.фм.н.	С.Л. Царев	

Содержание

Введени	1е	3	
Глава 1	Необходимые теоретические сведения	4	
1.1.	. Краткий обзор математических моделей эпидемий. Про-		
	стейшая модель SIR	4	
1.2.	. Модель SEIR	6	
1.3	. Модель SEIRD	7	
Глава 2	Построение математической модели инфодемии	9	
2.1.	. Сходства и отличия инфодемии и эпидемии	9	
2.2.	. Модель SEIR для инфодемии	10	
2.3	. Модель SEIRD для инфодемии	11	
Глава 3	Численное моделирование	13	
Заключ	ение	20	
Список литературы			
Приложение 1			
Приложение 2			

Введение

Настоящая работа посвящена применению математических моделей в такой сфере, как инфодемия. Дезинформация опасна, так как оказывает негативное влияние на: экономику, политику и психологическое состояние человека. За счет высокой скорости передачи информации через интернет и провокационному содержанию фейка, а также распространении дезинформации от человека к человеку. Можно сказать, что распространение схоже с вирусом. В итоге в информационном пространстве появляется новая «болезнь», названная инфодемией. Этот феномен представляет собой стремительное и неконтролируемое распространение в медиа необоснованной и ложной информации о различных событиях, усиливающее общественное беспокойство.

Как пример можно вспомнить недавнюю эпидемию — COVID-19. В связи с дезинформацией правительству стран стало сложнее проводить мероприятия для решения проблемы, так как ложная информация вызвала волну мнений и публикаций, ставящих под сомнение положительные стороны вакцинации в целом. Предпосылками к зарождению инфодемии коронавируса можно считать самоизоляцию, когда интернет и социальные сети стали основным источником новостей. Нехватка информации привела к аномальному доверию к анонимным источникам, в которых проверке фактов на достоверность авторы и пользователи уделяли минимальное время. При этом модерация информации, которая велась в большинстве соцсетей, могла осуществляться необъективно, защищая экономические и политические интересы некоторых стран и конкретных людей.

Изоляция, страх, значительное количество фейков, недостаток верной информации, все это в совокупности указывает на необходимость эффективных средств контроля над инфодемией. В связи с этим требуются модели, позволяющие прогнозировать развитие событий, а также ущерб.

Будем считать, что инфодемия это болезнь, и рассматривать ее распространение в пространстве, за распространение некоторого заболевания. Поэтому в качестве основы для модели распространения инфодемии были взяты модели распространения эпидемий, такие как SEIR и SEIRD. Подробное описание этих моделей будет дано в первой главе.

Глава 1 Необходимые теоретические сведения

1.1 Краткий обзор математических моделей эпидемий. Простейшая модель SIR

Распространение эпидемиологических заболеваний - сложное явление с множеством взаимодействующих факторов. Ключевая роль математической эпидемиологии заключается в создании моделей распространения патогенов.

Важным шагом в моделировании инфекций стала работа Андерсона Маккендрика и Уильяма Кермака, опубликованная в 1927 году [4]. Они создали базовую математическую модель, получившую название SIR-модель.

В классической модели SIR (от англ. Susceptible-Infected-Removed; по рус. Восприимчивый-Инфицированный-Ремонтированный) популяция разбивается на три группы, которые выделяются по признаку восприимчивости человека к болезни:

- \bullet восприимчивые (S) индивидуумы не инфицированы, но предрасположены к болезни;
- инфицированные (I) обозначает группу инфицированных индивидуумов, по совместительству разносчиков болезни;
- \bullet невосприимчивые (R) обозначает группу выздоровевших индивидуумов или умерших в результате болезни.

Численность групп восприимчивые, инфицированные, невосприимчивые обозначим соответственно $S(t),\,I(t),\,R(t).$ Существенно, что эти величины зависят от времени t. Изменения в зависимости от времени этих величин и составляют суть обсуждаемой математической модели, при этом общая численность популяции $N=S\left(t\right)+I\left(t\right)+R(t)$ принимается постоянной, то есть популяция рассматривается как замкнутая система. Считается, что переболевшие индивиды приобретают иммунитет и не могут быть заражены вторично. Схема перехода из одной группы в другую показана на рис. 1.1.

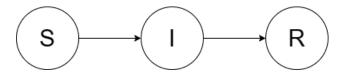


Рисунок 1.1 - Схема модели SIR

Модель выражается системой дифференциальных уравнений:

$$\begin{cases}
\frac{dS(t)}{dt} = -\frac{\beta S(t)I(t)}{N}, \\
\frac{dI(t)}{dt} = \frac{\beta S(t)I(t)}{N} - \gamma I(t), \\
\frac{dR(t)}{dt} = \gamma I(t), \\
S(t) + I(t) + R(t) = N.
\end{cases}$$
(1.1)

Кроме уже введенных нами величин S(t), I(t), R(t) в эту систему входят параметры β (отвечает за интенсивность распространения инфекции) и γ (выражает интенсивность выздоровления). Параметр γ представляет собой обратную величину к типичному времени болезни.

Кратко опишем смысл уравнений. Первое уравнение выражает тот факт, что количество восприимчивых индивидуумов со временем убывает пропорционально числу контактов с инфицированными, за число контактов восприимчивых с инфицированными отвечает произведение S(t)I(t), значение параметра бета зависит как от природы заболевания (разные вирусы имеют разную степень заразности), так и от социальных факторов (вероятность встречи здорового человека с больным может отличаться в разных популяциях).

Второе уравнение — количество заболевших индивидуумов со временем меняется пропорционально числу контактов восприимчивых с инфицированными и уменьшается по мере выздоровления заболевших.

Третье уравнение — количество выздоровевших или умерших пропорционально количеству заболевших.

Мы рассмотрели простейшую модель, которая не учитывает многие факторы такие как: изменение популяции (рождаемость, смертность, индивид может умереть в любой группе, независимо от болезни), неоднородность популяции (например, различную плотность населения в разных районах), миграция индивидов, разные пути передачи инфекции и т.д. Идеи, заложенные в SIR-модели, стали основой для дальнейших исследований

различных ученых. Постепенно на основе SIR-модели была построена целая иерархия математических моделей распространения инфекционных заболеваний. В этом разделе работы мы опишем основные эпидемиологические модели, а именно: SEIR, SEIRD.

1.2 Модель SEIR

Существует множество способов улучшения модели SIR. Один из таких способов — введение дополнительной группы.

Модель SEIR (от англ. Susceptible-Exposed-Infected-Removed; по рус. Восприимчивый-Латентный-Инфицированный-Невосприимчивый) является одной из ключевых моделей для анализа и прогнозирования распространения инфекционных болезней. Эта модель применялась для прогнозирования распространения COVID-19 в Китае и в европейских странах.

Модель системы:

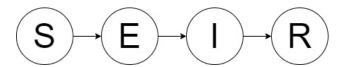


Рисунок 1.2 - Схема модели SEIR

Модель описывается системой дифференциальных уравнений:

$$\begin{cases} \frac{dS(t)}{dt} = -\frac{\beta S(t)I(t)}{N}, \\ \frac{dE(t)}{dt} = \frac{\beta S(t)I(t)}{N} - \sigma E(t), \\ \frac{dI(t)}{dt} = \sigma E(t) - \gamma I(t), \\ \frac{dR(t)}{dt} = \gamma I(t), \\ S(t) + E(t) + I(t) + R(t) = N. \end{cases}$$

$$(1.2)$$

Общий смысл уравнений остается таким же, как и в системе (1.1). Опишем изменения и новые параметры:

Появляется новая группа — E(t), представляющая количество индивидуумов, находящихся в инкубационном периоде на момент времени t. Обозначение происходит от слова "Exposed" (латентные).

Параметр σ описывает интенсивность инкубации и выражается как обратная величина средней продолжительности инкубационного периода.

Второе уравнение в системе вводит временную задержку при переходе от состояния контактного к состоянию инфицированного.

Общая численность популяции N остается постоянной и равна сумме численностей всех групп: N = S(t) + E(t) + I(t) + R(t).

Таким образом, вводятся изменения, учитывающие латентный период заболевания и задержку в переходе состояний, сохраняя общую структуру модели SIR.

1.3 Модель SEIRD

Модель SEIRD (от англ. Susceptible-Exposed-Infected-Removed-Dead; по рус. Восприимчивый-Латентный-Инфицированный-Невосприимчивый-Умерший) разделяет популяцию на пять групп. Группы S, E, I имеют то же значение, что и в модели SEIR. Добавляется группа D (умершие), которая разделяет группу R (невосприимчивые) на выздоровевшие и умершие. Данная модель была использована для описания распространения COVID-19 в России.

Модель системы:

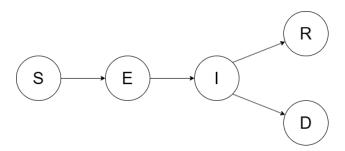


Рисунок 1.3 - Схема модели SEIRD

Модель описывается системой дифференциальных уравнений:

$$\begin{cases} \frac{dS(t)}{dt} = -\frac{\beta S(t)I(t)}{N}, \\ \frac{dE(t)}{dt} = \frac{\beta S(t)I(t)}{N} - \sigma E(t), \\ \frac{dI(t)}{dt} = \sigma E(t) - (\mu + \gamma)I(t), \\ \frac{dR(t)}{dt} = \gamma I(t), \\ \frac{dD(t)}{dt} = \mu I(t), \\ S(t) + E(t) + I(t) + R(t) + D(t) = N, \end{cases}$$

$$(1.3)$$

где обозначения $S,\,I,\,R,$ и т. д. имеют тот же смысл, что и в системах (1.1), (1.2). Величина D(t) выражает количество умерших в момент времени t (в моделях SIR, SEIR, эта группа относилась к категории R). В данной системе μ — коэффициент смертности. Данная модель более информативна потому, что учитывает больше факторов, однако это не предел.

В этом разделе мы рассмотрели базовые модели эпидемий, на основании которых мы будем конструировать модели инфодемии.

Глава 2 Построение математической модели инфодемии

2.1 Сходства и отличия инфодемии и эпидемии

Эпидемиологическое распространение болезни очень похоже на распространение ложной информации в интернет-пространстве, поэтому за основу мы возьмем математические модели эпидемий (1.2), (1.3).

Одно из отличий эпидемии от инфодемии состоит в том, что у фейка есть две стадии развития:

- 1. Невидимая стадия информация отправляется друзьям и близким в личных сообщениях [3].
- 2. Видимая стадия(публичная) новостные источники популяризируют информацию [3].

Стадии могут идти в обратном порядке. Новостные источники распространяют дезинформацию, а люди не намеренно популяризируют её обсуждениями и комментариями.

Ещё одно отличие состоит в том, что при эпидемии, чтобы заболеть требуется контакт с инфицированным. В случае инфодемии расстояние учитываться не будет. Так же стоит учесть, что патогены не обладают самосознанием, следовательно, они не выбирают кого заражать в отличие от дезинформации, ведь слухи распространяют люди, но не всегда люди, получившие дезинформацию, верят в неё.

Итак, мы будем строить 2 математических модели дезинформации, основываясь на ранее изученных моделях, а именно SEIR, SEIRD. Для начала обозначим группы:

Индивидуумы, которые могут поддаться дезинформации или те кто еще не знает эту информацию — S.

Индивидуумы, узнавшие новость и обдумывающие ee - E.

«Инфицированные» индивидуумы, распространяющие дезинформацию — I.

Невосприимчивые, не распространяют дезинформацию — R. Индивидуумы, потерявшие интерес к новости — D.

Состав каждой группы меняется со временем. Индивиды могут переходить из одной группы в другую в зависимости от получаемых в дальнейшем новостей или при угасании интереса к теме или т. п. Договоримся буквами $S,\,E,\,I,\,R,\,D$ обозначать не сами группы, а их численность, поскольку она зависит от времени мы будем иметь дело с функциями $S(t),\,E(t),\,I(t),\,R(t),\,D(t)$

Переход каждого отдельного человека в группу E происходит не автоматически поскольку человек может не поверить новости или игнорировать ее. Поэтому естественно ввести параметр p, который будет описывать вероятность «заражения».

2.2 Модель SEIR для инфодемии

Учитывая вышесказанное, построим модель:

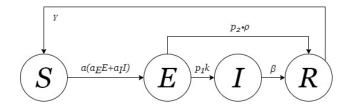


Рисунок 2.1 - Схема модели инфодемии SEIR

Модель описывается системой дифференциальных уравнений:

$$\begin{cases} \frac{dS(t)}{dt} = -a(t-\tau) \left(\frac{a_I S(t) I(t)}{N} + \frac{a_E S(t) E(t)}{N} \right) + \gamma R(t), \\ \frac{dE(t)}{dt} = a(t-\tau) \left(\frac{a_I S(t) I(t)}{N} + \frac{a_E S(t) E(t)}{N} \right) - (pk + (1-p)\rho) E(t), \\ \frac{dI(t)}{dt} = pk E(t) - \beta I(t), \\ \frac{dR(t)}{dt} = \beta I(t) + (1-p)\rho E(t) - \gamma R(t), \end{cases}$$

$$(2.1)$$

где:

- t время (0, T);
- τ «невидимая» стадия 24 часа;
- $a(t-\tau)$ параметр аналогичен параметру социальной изоляции в эпидемиях (0,1);
- ρ интенсивность потери интереса (0, 1);

- β интенсивность борьбы со слухом или скорость переубеждения (0, 1);
- γ Интенсивность перехода из состояния невосприимчивых в состояние восприимчивых. Это может означать, что информация была забыта, либо появилась новая, более правдоподобная дезинформация, и т.д. (0, 1);
- a_I параметр распространения слухов между инфицированными индивидуумами (0, 1);
- a_E параметр заражения между «дозревающими» зараженными индивидуумами (0, 1);
- k интенсивность принятия новости (0, 1).

2.3 Модель SEIRD для инфодемии

В этом разделе мы рассмотрим модификацию модели описанной в [3]. Построим модель:

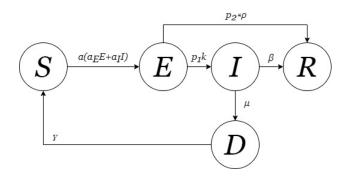


Рисунок 2.2 - Схема модели инфодемии SEIRD

Модель описывается системой дифференциальных уравнений:

$$\begin{cases} \frac{dS(t)}{dt} = -a(t-\tau) \left(\frac{a_I S(t)I(t)}{N} + \frac{a_E S(t)E(t)}{N} \right) + \gamma D(t), \\ \frac{dE(t)}{dt} = a(t-\tau) \left(\frac{a_I S(t)I(t)}{N} + \frac{a_E S(t)E(t)}{N} \right) - (pk + (1-p)\rho)E(t), \\ \frac{dI(t)}{dt} = pkE(t) - (\beta + \mu)I(t), \\ \frac{dR(t)}{dt} = \beta I(t) + (1-p)\rho E(t), \\ \frac{dD(t)}{dt} = \mu I(t) - \gamma D(t). \end{cases}$$

$$(2.2)$$

Смысл параметров в этой системе уравнений такой же как и в модели SEIR (2.1). Новым является параметр μ , который отвечает за интенсивность потери интереса к слуху.

Глава 3 Численное моделирование

Системы уравнений (2.1), (2.2) не имеют аналитического решения [4], поэтому мы используем компьютерное моделирование, чтобы получить численное решение. Для написания кода в Jupyter Notebook на Python мы использовали следующие библиотеки и расширения: numpy, matplotlib, pandas, ipywidgets, scipy, matplotlib.animation и IPython.display. Для обработки задачи применялся модуль scipy.integrate.solve_ivp. Часть кода программ представлена в приложении.

Ввиду ограниченной осведомленности относительно инфодемии и отсутствия соответствующих прецедентов ее распространения, наше численное моделирование предполагает использование адаптивных параметров. Таким образом, каждый из параметров модели подлежит гибкому изменению в процессе исследования.

Общий размер популяции равен N=200000, рассмотрим промежуток в 150 дней. Результаты отражены на иллюстрациях приведенных на следующих страницах.

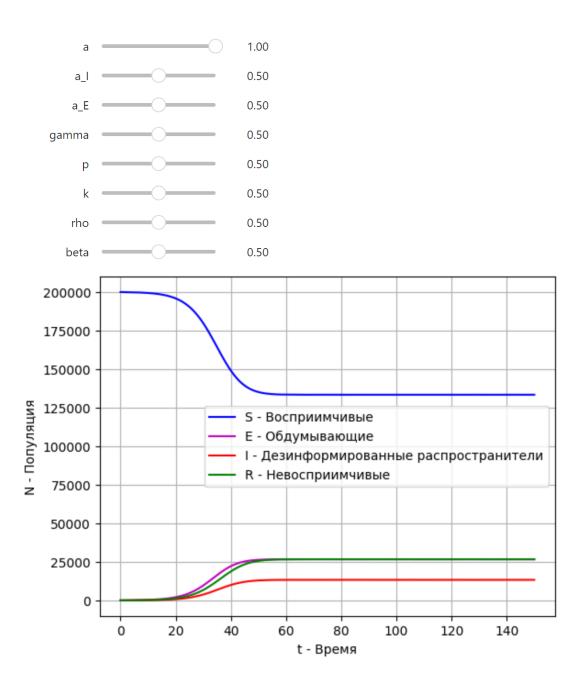


Рисунок 3.1 - Результаты численного моделирования системы SEIR при значениях параметров: $a=1, a_I=0.5, a_E=0.5; \gamma=0.5; p=0.5; k=0.5; \rho=0.5; \beta=0.5$

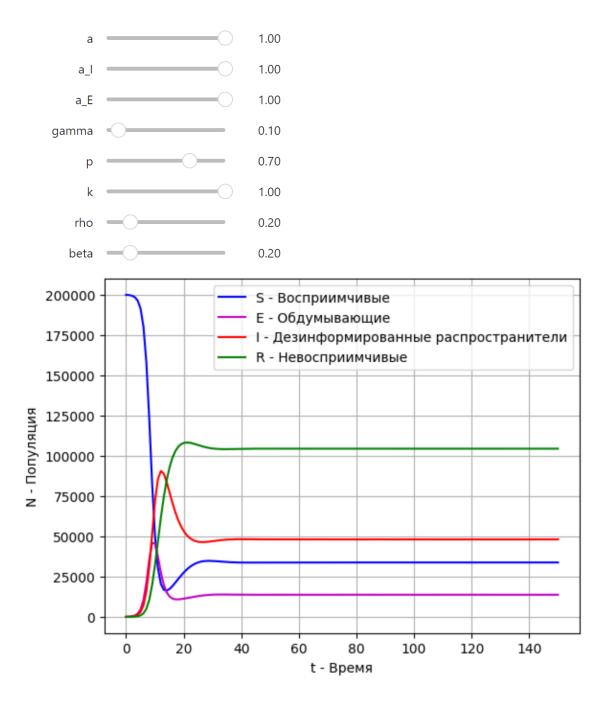


Рисунок 3.2 - Результаты численного моделирования системы SEIR при значениях параметров: $a=1, a_I=1, a_E=1, \gamma=0.1, p=0.7, k=1, \rho=0.2, \beta=0.2$

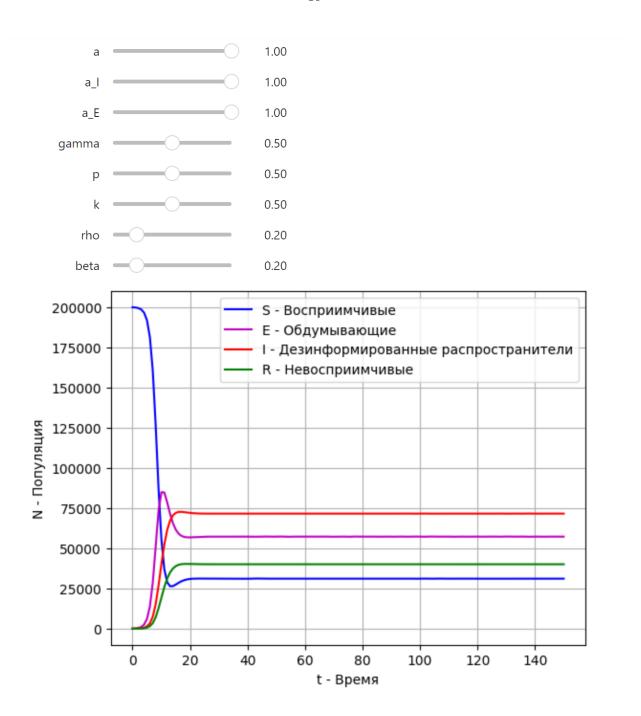


Рисунок 3.3 - Результаты численного моделирования системы SEIR при значениях параметров: $a=1, a_I=1, a_E=1, \gamma=0.5, p=0.5, k=0.5, \rho=0.2, \beta=0.2$

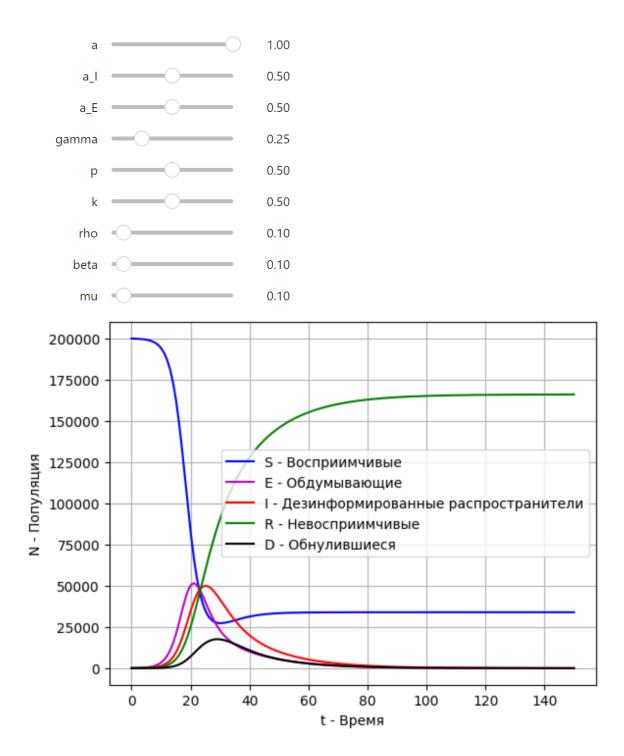


Рисунок 3.4 - Результаты численного моделирования системы SEIRD при значениях параметров: $a=1, a_I=0.5, a_E=0.5, \gamma=0.25, p=0.5, k=0.5, \rho=0.1, \beta=0.1, \mu=0.1$

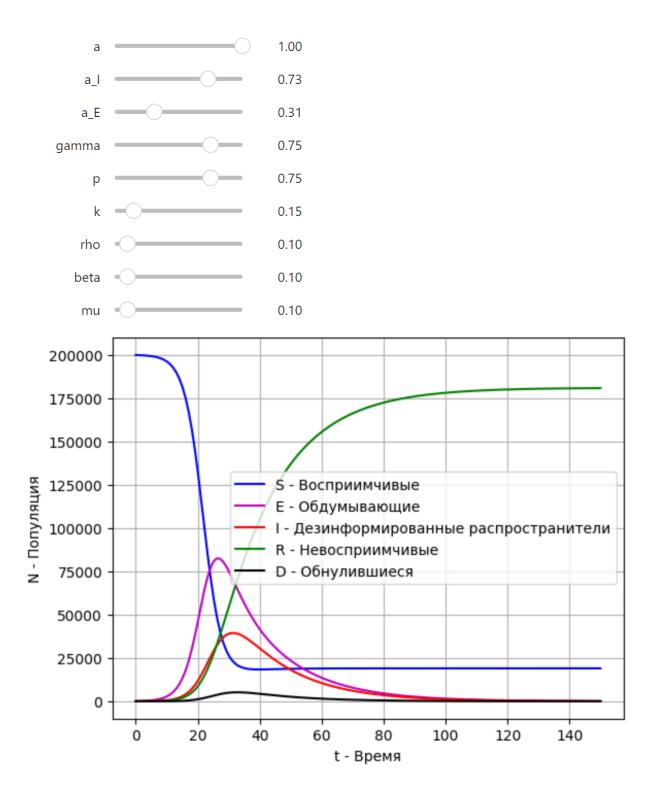


Рисунок 3.5 - Результаты численного моделирования системы SEIRD при значениях параметров: $a=1, a_I=0.73, a_E=0.31, \gamma=0.75, p=0.75, k=0.15, \rho=0.1, \beta=0.1, \mu=0.1$

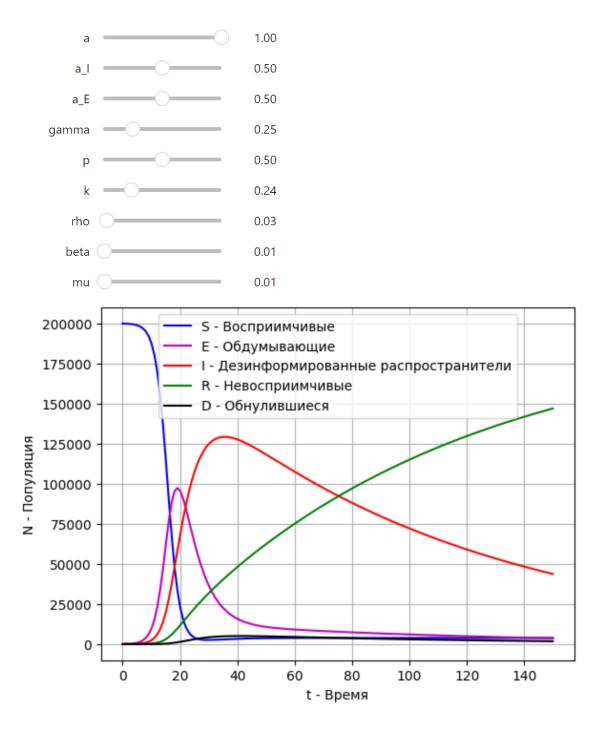


Рисунок 3.6 - Результаты численного моделирования системы SEIRD при значениях параметров: $a=1, a_I=0.5, a_E=0.5, \gamma=0.25, p=0.5, k=0.24, \rho=0.03, \beta=0.01, \mu=0.01$

Заключение

В данной работе мы представили описания эпидемиологических моделей и соответствующих им моделей инфодемии, были проведены численные эксперименты.

Дальнейшие исследования применимости моделей требуют обращения к реальным данным и их анализа для оценки точности предложенных моделей. Этим вопросом мы планируем глубже заняться в рамках нашей будущей исследовательской работы.

Список литературы

- Гришунина Ю. Б. Моделирование эпидемической ситуации с учетом внешних рисков / Ю.Б. Гришунина, Н.А. Контаров, Г.В. Архарова, Н.В. Юминова // Эпидемиология и вакцинопрофилактика. 2014. № 5 (78). С. 61-66.
- 2. *Еремеева Н. И.* Построение модификации SEIRD-модели распространения эпидемии, учитывающей особенности COVID-19 / Н.И. Еремеева // Вестник ТвГУ. Серия: Прикладная математика, 2020, выпуск 4, 14-27.
- 3. Щетинина И. В. Математическая модель COVID-инфодемии / И.В. Щетинина, Е.А. Москалева, М.Е. Волкова, В.К. Власов // Информация и безопасность. 2021. \mathbb{N}° 4(-). С. 521-530.
- 4. Яушева О. А. Математическая модель эпидемии лихорадки Эбола [Выпускная квалификационная работа бакалавра] / О. А. Яушева. Снкт-Петербург: СПбГУ, 2016. 39 с.

Приложение 1

```
[1]: # Connecting libraries
import numpy as np
import matplotlib.pyplot as plt
from scipy import integrate
from ipywidgets import interact
```

```
[2]: # Creating functions
     def seird_model(t, y, N, a, a_I, a_E, gamma, p, k, rho, beta):
         S, E, I, R = y
         dS_dt = - a *(a_I * S * I / N + a_E * S * E / N) + gamma * R
         dE_dt = a *(a_I * S * I / N + a_E * S * E / N) - (p * k + (1 - p) * rho) * E
         dI_dt = p * k * E - beta * I
         dR_dt = beta * I + (1 - p) * rho * E - gamma * R
         return dS_dt, dE_dt, dI_dt, dR_dt
     def interact_sir(a=1, a_I=.5, a_E=.5, gamma=.5, p=.5, k=.5, rho=.5, beta=.5):
         T = 365
         sol = integrate.solve_ivp(lambda t, y: seird_model(t, y, N, a, a_I, a_E, __
     ⇔gamma,
                                                            p, k, rho, beta),
                                   [0, T], [S0, E0, I0, R0], dense_output=True)
         t = np.linspace(0, T, T)
         y = sol.sol(t)
         S, E, I, R = y[0, :], y[1, :], y[2, :], y[3, :]
         plt.plot(t, S, color='b', label='S - Восприимчивые')
         plt.plot(t, E, color='m', label='E - Обдумывающие')
         plt.plot(t, I, color='r', label='I - Дезинформированные распространители')
         plt.plot(t, R, color='g', label='R - Невосприимчивые')
         plt.xlabel('t - Bpems')
         plt.ylabel('N - Популяция')
         plt.legend()
         plt.grid()
         plt.show()
```

```
[3]: # Starting parameters
N = 200000
p = .5
a = 1
rho = .5
beta = .5
gamma = .5
a_I = .5
a_E = .5
k = .5
E0 = 0
I0 = 100
R0 = 0
S0 = N - E0 - I0 - R0
```

interactive(children=(FloatSlider(value=1.0, description='a', max=1.0, step=0.01),
 →FloatSlider(value=0.5, desc...

[4]: <function __main__.interact_sir(a=1, a_I=0.5, a_E=0.5, gamma=0.5, p=0.5, k=0.5, rho=0.5, beta=0.5)>

Приложение 2

```
[1]: # Connecting libraries
import numpy as np
import matplotlib.pyplot as plt
from scipy import integrate
from ipywidgets import interact
```

```
[2]: # Creating functions
     def seird_model(t, y, N, a, a_I, a_E, gamma, p, k, rho, beta, mu):
         S, E, I, R, D = y
         dS_dt = - a * (a_I * S * I / N + a_E * S * E / N) + gamma * D
         dE_dt = a *(a_I * S * I / N + a_E * S * E / N) - (p * k + (1 - p) * rho) * E
         dI_dt = p * k * E - beta * I - mu * I
         dR_dt = beta * I + (1 - p) * rho * E
         dD_dt = mu * I - gamma * D
         return dS_dt, dE_dt, dI_dt, dR_dt, dD_dt
     def interact_sir(a=1, a_I=.5, a_E=.5, gamma=.25, p=.5,
                      k=.5, rho=.1, beta=.1, mu=.1):
         T = 365
         sol = integrate.solve_ivp(lambda t, y: seird_model(t, y, N, a, a_I, a_E,
                                                            gamma, p, k, rho, beta,
     →mu),
                                   [0, T], [S0, E0, I0, R0, D0], dense_output=True)
         t = np.linspace(0, T, T)
         y = sol.sol(t)
         S, E, I, R, D = y[0, :], y[1, :], y[2, :], y[3, :], y[4, :]
         plt.plot(t, S, color='b', label='S - Восприимчивые')
         plt.plot(t, E, color='m', label='E - Обдумывающие')
         plt.plot(t, I, color='r', label='I - Дезинформированные распространители')
         plt.plot(t, R, color='g', label='R - Невосприимчивые')
         plt.plot(t, D, color='k', label='D - Обнулившиеся')
         plt.xlabel('t - Bpems')
         plt.ylabel('N - Популяция')
         plt.legend()
         plt.grid()
         plt.show()
```

```
[3]: # Starting parameters
     N = 200000
    p = .5
     a = 1
    rho = .1
     beta = .1
     gamma = .25
    mu = .1
     a_I = .5
     a_E = .5
    k = .5
     EO = 0
     IO = 100
     RO = 0
     DO = 0
     SO = N - EO - IO - RO - DO
```

interactive(children=(FloatSlider(value=1.0, description='a', max=1.0, step=0.01),
 →FloatSlider(value=0.5, desc...

[4]: <function __main__.interact_sir(a=1, a_I=0.5, a_E=0.5, gamma=0.25, p=0.5, k=0.5, rho=0.1, beta=0.1, mu=0.1)>