Конспект лекции

К.К. Логинов

Моделирование

распространения эпидемий

***Слайд 1.*** Добрый день, я представляю лекцию 5 «Моделирование распространения эпидемий» по дисциплине «Модели больших сетевых структур и сетевые процессы». Разработчиками курса являются сотрудники Омского государственного технического университета Юдин Е.Б., Юдина М.Н., Бадрызлов В.А., а также сотрудник института математики им. С.Л. Соболева Логинов К.К., который является автором этой лекции.

# Классическая SIR (S → I → R) модель

## Описание модели

***Слайд 2.*** В данной лекции речь будет идти о болезнях, вызываемых микропаразитами, т.е. когда единичная инфекция вызывает процесс заболевания, протекающий приблизительно одинаково у всех инфицированных индивидуумов. В этом случае нет необходимости следить за динамикой паразитов, а достаточно наблюдать за динамикой здоровых и заболевших индивидуумов. На слайде представлена детерминированная модель SIR (модель Кермака–МакКендрика), название которой составлено из обозначений классов, к которым относятся индивидуумы популяции (населения в некотором регионе). К детерминированным моделям относят модели, переменные которых связаны определенной функциональной зависимостью («полная определенность в настоящем и будущем», «отсутствует случайность»). В данном случае модель описывается системой обыкновенных дифференциальных уравнений, в которых переменные (susceptible), (infected), (removed или recovered) означают, соответственно, численности восприимчивых, инфицированных и иммунных (переболевших и приобретших иммунитет) индивидуумов. Модель SIR и подобные ей модели (SIRS, SEIR, SEIRS и др.) относятся к так называемым когортным моделям, в рамках фиксированной когорты индивидуумы являются неразличимыми (не имеют индивидуальных параметров как, например, в агентных моделях). Также следует отметить отсутствие инкубационного периода инфекции (латентная стадия заболевания), т.е. он считается настолько малым, что им можно пренебречь, и заболевший восприимчивый индивидуум сразу же становится инфицированным и способен заражать других здоровых индивидуумов. Об интенсивностях , речь пойдет подробно ниже при построении стохастической модели. Эти интенсивности описывают заражение и выздоровление за бесконечно малый промежуток времени. Причем, например, время до выздоровления одного индивидуума считается распределенным экспоненциально с параметром , следовательно, обратная величина представляет собой среднюю продолжительность болезни.

***Слайд 3.*** Из уравнений следует, что размер популяции фиксирован (популяция замкнута). Кроме того, переменная является вспомогательной, т.к. не входит явно в уравнения (1), (2) и вычисляется с помощью и . Переменная введена для отражения баланса общей численности популяции. Обратим внимание на то, что скорость прироста когорты инфицированных индивидуумов пропорциональна количеству восприимчивых и инфицированных индивидуумов. Это предположение соответствует так называемому закону действующих масс (по аналогии со скоростью протекания химических реакций) для хорошо перемешанной популяции. В химической кинетике закон действующих масс означает, что скорость протекания химической реакции пропорциональна произведению концентраций реагентов. Сходным образом предполагается, что скорость изменения численности здоровых индивидуумов пропорциональна произведению плотностей (или численностей) популяций здоровых и инфицированных индивидуумов. В общем случае, конечно, индивидуумы контактируют внутри обособленных групп, которые могут быть не связаны друг с другом. Однако для большинства инфекций такая идеализация является вполне приемлемой.

***Слайд 4.*** На слайде дается некоторое представление о поведении решений модели SIR. Видно, что при ( – относительная интенсивность устранения инфекции) инфекция быстро затухает и траектории по уходят в нуль. При возникает эпидемия, и количество инфицированных сначала растет, достигая некоторого максимального значения , однако затем инфекция также затухает. Более детальное объяснение приведено на слайде.

## Финальный размер эпидемии

***Слайд 5.*** Здесь вводятся предельные численности здоровых, инфицированных и переболевших индивидуумов. Из предыдущих выкладок (см. предыдущий слайд) следует, что . Решая уравнение для , получаем выражение через . Переходя к пределу при , приходим к трансцендентному уравнению на , корень которого и является финальным размером здоровых индивидуумов, которые избежали инфекции. Несложно показать (например, графически), что для данного уравнения всегда существует корень .

***Слайд 6.*** На данном слайде вводится понятие *базового репродуктивного числа* , занимающее центральное место в математической эпидемиологии. Величина определяется как произведение – среднее число контактов между и (инфицирование) в единицу времени и – средняя продолжительность болезни у инфицированного индивидуума. Формулируется также основная пороговая теорема, которая гласит, что возникновение эпидемии среди населения возможно только при . Если же , то инфекция быстро затухает и эпидемии не возникает.

1. **Коллективный иммунитет**

***Слайд 7.*** Если известна доля индивидуумов, которая осталась здоровой, то можно сделать оценку базового репродуктивного числа для конкретной инфекции. Если известна оценка , то при вакцинации населения необходимо делать прививки определенной доли населения, чтобы избежать возможной эпидемии. Выражение для такой доли (при которой выполняется неравенство ) приводится на слайде. Таким образом, для создания коллективного иммунитета в популяции необходимо подвергнуть вакцинации только определенную долю населения. В качестве примера можно привести такие заболевания как оспа (доля составляет 70-80%), корь (доля равна 90-95%).

Сформулированная пороговая теорема справедливы не только для SIR модели, но и для широкого класса других подобных моделей (модель SIR с миграцией; SIRS модель, где индивидуумы после некоторого времени переходят в класс восприимчивых (потеря иммунитета к инфекции); SEIRS модель, где введена группа латентно-инфицированных индивидуумов и многие другие детерминированные эпидемиологические модели).

1. **СТОХАСТИЧЕСКИЙ АНАЛОГ ДЕТЕРМИНИРОВАННОЙ МОДЕЛИ SIR**
2. **Описание модели**

***Слайд 8.*** Приступим к построению стохастической (вероятностной) модели SIR. На данном слайде представлен случайный процесс , компоненты которого являются неотрицательными целочисленными (в отличие от детерминированной модели SIR, где работа идет с вещественными переменными) случайными величинами, описывающими численности когорт. Постулируются основные предположения модели, связанные с вероятностями инфицирования и выздоровления индивидуумов за бесконечно малый промежуток времени. Сразу отметим, что при таком подходе время до ближайшего контакта между и является случайной величиной с экспоненциальным распределением с параметром ; соответственно, время до ближайшего выздоровления индивидуума – случайная величина, распределенная экспоненциально с параметром .

***Слайд 9.*** На этом слайде приведены вероятности переходов процесса за бесконечно малый промежуток времени , , при фиксированных численностях всех когорт в некоторый момент времени . По построению случайный процесс является марковским (эволюция процесса при любом фиксированном моменте времени не зависит от его предыстории; «при фиксированном настоящем будущее не зависит от прошлого»). Далее показана связь между детерминированной и стохастической моделями SIR. Суть заключается в получении дифференциальных уравнений на математические ожидания изучаемых переменных , и .

***Слайд 10.*** На слайде показан пример вывода уравнения для математического ожидания величины . Используя введенные выше постулаты и стандартный аппарат для условных математических ожиданий (формулу полной вероятности для математических ожиданий), мы выражаем математическое ожидание в момент времени для некоторого малого через математическое ожидание в момент времени .

***Слайд 11.*** В итоге мы получаем дифференциальное уравнение на правостороннюю производную , которое совпадает с уравнением для в детерминированной модели SIR с точностью до слагаемого (в случае взаимодействий типа или добавилось бы слагаемое с дисперсиями соответствующих величин как, например, в модели Лотки-Вольтерра). Используя ту же технику, аналогично можно получить уравнения для и . Уравнения для , отличаются от соответствующих уравнений только слагаемым с ковариацией, поэтому можно говорить о том, что детерминированная модель в некотором смысле приближенно описывает поведение математических ожиданий , , . Следует добавить, что детерминированные модели с высокой точностью описывают изменения численностей когорт только для довольно больших численностей ( и выше). В случае же малых численностей детерминированные модели могут давать не совсем верные результаты, и предпочтительнее использовать стохастические модели.

1. **Алгоритм численного моделирования (метод Монте-Карло)**

***Слайд 12.*** Прежде чем приступить к описанию алгоритма моделирования с использованием метода Монте-Карло, еще раз напомним об экспоненциальных распределениях промежутка времени до осуществления ближайшего события, при котором случайный процесс меняет свое состояние (инфицирование или выздоравливание индивидуумов). Далее мы будем использовать известный факт о том, что минимум из нескольких исходных случайных величин с экспоненциальными распределениями будет сам распределен экспоненциально с параметром, равным сумме параметров исходных распределений.

***Слайд 13.*** На данном слайде представлен алгоритм численного моделирования (в контексте нашей темы иногда его называют алгоритмом Гиллеспи). Фиксируются конечный промежуток моделирования и начальные данные. Затем с использованием метода обратной функции распределения моделируется шаг по времени до ближайшей смены состояния процесса (см. предыдущий слайд). И, наконец, разыгрывается смена состояния (инфицирование или выздоравливание ) изучаемого случайного процесса. Затем шаги 2, 3 повторяются до тех пор, пока не произойдет «вырождения» когорты (искоренение инфекции), либо не будет пройден отрезок моделирования .

*Примечание*: метод обратной функции распределения состоит в получении выборочного значения некой случайной величины путем решения уравнения

, или ,

где – функция распределения случайной величины ;

–выборочное значение случайной величины, равномерно распределенной в интервале .

В частности, если случайная величина распределена экспоненциально с параметром :

, ,

применение метода обратной функции позволяет в явном виде получить формулу для выборочного значения .

1. **СТОХАСТИЧЕСКАЯ МОДЕЛЬ РАСПРОСТРАНЕНИЯ НОВОЙ ДЛЯ ПОПУЛЯЦИИ РЕСПИРАТОРНОЙ ВИРУСНОЙ ИНФЕКЦИИ**
2. **Описание модели**

***Слайд 14.*** Как уже говорилось выше, в рассмотренной модели SIR предполагалась однородность когорт (групп) населения (в рамках конкретной когорты индивидуумы считались неразличимыми), а также отсутствовал инкубационный период инфекции. Далее мы будем рассматривать многопараметрическую стохастическую стадия-зависимую модель распространения эпидемии, разработанную в 2021 году Перцевым Николаем Викторовичем совместно с автором этой лекции (Институт математики им. С.Л. Соболева СО РАН), учитывающей неоднородность населения по ряду критериев и позволяющей прогнозировать численности инфицированных, больных и умерших от инфекции индивидуумов.

Население региона представлено в виде двух блоков когорт индивидуумов, структурированных по иммунологическим, клиническим, эпидемиологическим и демографическим критериям. Особое внимание уделяется индивидуумам супер-распространителям инфекции, т.е. тем заразным больным индивидуумам, которые вносят основной вклад в инфицирование восприимчивых к инфекции индивидуумов. Кроме того, модель учитывает миграционные притоки восприимчивых, латентно-инфицированных и больных индивидуумов, прохождение индивидуумами нескольких стадий инфекционного заболевания и различную тяжесть заболевания.

В первый блок входят индивидуумы, считающиеся неразличимыми в рамках конкретной когорты и имеющие однотипное параметрическое описание, специфическое для каждой из когорт этого блока. Индивидуумы, входящие в каждую из когорт второго блока, также имеют однотипное параметрическое описание, но различаются между собой по времени поступления в конкретную когорту и по времени пребывания в этой когорте. Использование двух блоков когорт приводит к совмещению популяционного и индивидуум-ориентированного (агентного) подходов. Для построения модели используется непрерывно-дискретный случайный процесс, учитывающий взаимодействия (контакты) индивидуумов и прохождение индивидуумами различных стадий инфекционного заболевания. Принято, что распределение различных стадий инфекционного заболевания отлично от экспоненциального, что приводит к необходимости учета предыстории развития когорт индивидуумов, и, соответственно, к немарковскому случайному процессу.

***Слайд 15.*** На слайде представлена блок-схема модели и описание когорт индивидуумов, различающихся по возрасту, уровню контактности и иммунитета, а также по уровню заразности и форме тяжести заболевания. Предположения **H1**–**H8** описывают притоки, естественную гибель и миграцию индивидуумов, инфицирование индивидуумов, развитие заболевания и выздоровление индивидуумов (подробнее на следующих слайдах). Специфика развития заболевания у конкретного человека представлена в упрощенной форме за счет введения нескольких стадий заболевания и постоянных длительностей пребывания индивидуумов в этих стадиях. Работа органов здравоохранения региона отражена в форме проведения вакцинации населения, начиная с некоторого момента времени от начала эпидемического процесса. Главная цель состоит в построении модели, направленной на исследование возможных сценариев развития эпидемического процесса в популяции, которая впервые встречается с новой для себя респираторной вирусной инфекцией, например, COVID-19 инфекцией.

***Слайд 16–18.*** На данных слайдах подробно описаны предположения **H1**–**H8** в виде так называемых кинетических схем, подобных тем, что приводились для стохастической модели SIR.

***Слайд 19.*** На слайде приведена краткая вероятностная формализация модели (подробная формализация модели привела бы к значительному увеличению презентации; в любом случае подробное описание можно найти в работах Перцева Н.В. и Логинова К.К.). Непрерывно-дискретный случайный процесс , используемый для описания модели, содержит: ***1)*** неотрицательные целочисленные случайные переменные, задающие численности индивидуумов когорт каждого из двух блоков (вектор ); ***2)*** специальные множества – семейства уникальных типов индивидуумов, отражающих моменты поступления индивидуумов в когорты второго блока и продолжительности пребывания индивидуумов в этих когортах (множество ). Как мы видим, задача является высокоразмерной: 810 переменных, описывающих численности когорт, и 675 семейств уникальных типов. Семейство уникальных типов представляет собой упорядоченное по возрастанию множество вещественных чисел, характеризующееся моментом появления индивидуума в соответствующей когорте и продолжительностью его пребывания (без учета гибели и миграции) в этой когорте. Введение семейств уникальных типов позволяет перейти от немарковского (представленного только численностями когорт индивидуумов) к марковскому случайному процессу в расширенном пространстве состояний, включающему описание когорт индивидуумов с помощью целочисленных и вещественных переменных.

1. **Алгоритм численного моделирования (метод Монте-Карло)**

***Слайд 20.*** Здесь представлен алгоритм моделирования смены состояний изучаемого процесса , подобный описанному ранее алгоритму для стохастической модели SIR. Отличие состоит в том, что при моделировании шага по времени необходимо сравнивать момент времени, полученный из экспоненциального распределения с суммарной интенсивностью всех взаимодействий, с первым элементом каждого из специальных множеств . Помимо этого при смене состояния процесса изменения могут касаться не только численностей когорт, но элементов семейств уникальных типов.

1. **Вычислительные эксперименты**

***Слайд 21.*** На этом слайде перечислены факторы, обуславливающие варианты развития новой эпидемии, которые были выбраны при проведении вычислительных экспериментов. Приведена формула для интенсивностей контактов. Ниже представлены некоторые результаты исследования динамики численности населения в начальный период распространения респираторной инфекции, которая является новой для индивидуумов, т.е. инфекцией, с которой население региона ранее не встречалось, например, COVID-19. Исследовалась динамика численности когорт для наборов параметров, отражающих различные варианты передачи инфекции между индивидуумами, а также сопоставление динамики переменных модели при вакцинации населения и появлении нового, более агрессивного штамма вируса.

***Слайд 22.*** Начальная численность соответствует городу-миллионнику. На рисунке приведены условные доли каждой возрастной группы от всего населения. Внутри каждой возрастной группы начальная численность распределена равными долями по уровню иммунитета , и составляет 80%, 10%, 10% соответственно для уровня контактности (низкий, средний, высокий); это означает, что в начальный момент времени доля супер-распространителей достаточно мала. Оценивалась численность всех выявленных больных изолированных индивидуумов трудоспособного возраста, численность всех невыявленных больных (которые вносят основной вклад в инфицирование здоровых индивидуумов), а также численность всех умерших от инфекции индивидуумов. Особо следует отметить, что переменные , могут быть оценены (или точно подсчитаны) по реальным данным, тогда как переменные , являются скрытыми и могут быть оценены по реальным данным только косвенно. Построенная модель позволяет исследовать динамику , и сопоставлять эти переменные с переменными, значения которых могут быть получены по реальным данным.

***Слайд 23.*** На данном слайде приведены точечные оценки средней численности всех больных изолированных индивидуумов трудоспособного возраста, всех латентно-инфицированных и невыявленных больных индивидуумов, а также средней численности всех умерших от инфекции индивидуумов. Цифрами обозначены различные сценарии развития эпидемического процесса, включая вакцинацию населения и появление нового штамма вируса. Приведенные вычисления демонстрируют возможный механизм формирования второй волны эпидемического процесса за счет появления нового штамма вируса. Так, в частности, снижение среднего значения переменной на промежутке времени от 120 до 130 суток сопровождается резким увеличением среднего значения «скрытой» переменной . Далее на определенных промежутках времени происходит рост и снижение средних значений указанных переменных. Особо следует отметить, что при выбранных параметрах модели вторая волна эпидемического процесса приводит к существенному росту средних значений переменной .

В заключение отметим, что разработанная модель является легко расширяемой с точки зрения введения новых факторов и после соответствующей калибровки может служить полезным инструментом прогнозирования динамики эпидемического процесса как для медицинских служб, так и для региональных органов власти.