

УПРАВЛЕНИЕ НЕЛИНЕЙНЫМИ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИМИ ПРОЦЕССАМИ, РАЗВИВАЮЩИМИСЯ В УСЛОВИЯХ ОГРАНИЧЕНИЙ. МЕТОДЫ ОПРЕДЕЛЕНИЯ ПАРАМЕТРОВ ДИНАМИКИ ЭПИДЕМИИ COVID 19

Полунин Ю.А.

*Институт проблем управления им. В.А. Трапезникова РАН,
Россия, г. Москва, ул. Профсоюзная д.65*

yplnn@yandex.ru

На примере статистического анализа динамики заражения COVID 19 рассматриваются вопросы определения параметров процесса, развивающегося в условиях ограничений, которые необходимы для оценки эффективности управления. Предложена модель заражений определяющая структуру регрессионных моделей, которая позволяет интерпретировать коэффициенты регрессионных моделей в категориях заражений.

Ключевые слова: нелинейный процесс, ограничения процесса, параметры заражений, регрессионные модели.

Введение

Несмотря на то, что вопросы управления социально-экономическими процессами, развитие которых происходит в условиях ограничений, широко обсуждаются [1] эпидемия COVID 19 сделала актуальной новую задачу управления при малой априорной информации о механизмах заражений. Это порождает ситуацию, когда выдвигаются много противоречивых объяснений, основанных на частных фактах. Рациональным подходом является использование статистических методов для анализа динамики заражений. Наиболее пригодны для описания динамики заражений методы регрессионного анализа, изложенные в классических работах [2], [3]. Для использования в задачах управления необходимы и регрессионные модели результаты которых должны позволять интерпретацию в категориях механизмов заражений. Ниже рассматриваются подходы, основанные на ежедневной статистике числа выявленных больных [4], позволяющие интерпретировать полученные статистические результаты. Для этого, исходя из самых общих предположений, предлагается модель заражений с учетом влияющих на процесс ограничений, с использованием категорий «потенциал» и «интенсивность заражений». Исходя из структуры модели заражений, выбираются связанные с ней предикторы регрессионных моделей.

1 Модель заражений

Пока, в условиях неоконченной пандемии, достоверных знаний о механизмах заражения недостаточно, более того уже ясно видно, что механизмы заражений меняют свои свойства, что наиболее ярко проявляется в виде волн заражений. Поэтому для статистического анализа необходимы статистические модели, базирующиеся на самых общих представлениях о механизмах заражений и позволяющие сравнивать динамику заражений на разных участках в сопоставимых категориях. Будем рассматривать динамику эпидемии на уровне страны или достаточно однородной по механизму заражения территории. Примем самые общие допущения о механизмах заражения.

- Новые заражения происходят за счет передачи инфекции двумя типами людей - выявленными в диагностике больными, плюс не выявленными на данный момент больными. Ко второму типу заражающих относятся больные, прибывающие на данную территорию извне или не выявленные за предыдущие периоды.
- Больные люди являются заразными ограниченное время – в пределах цикла заражения; переболевшие люди не заражают и не подвергаются повторному заражению.
- Механизмы заражения в разных странах или разных моментах времени в одной стране могут отличаться.

Для создания модели заражений не хватает априорной информации о средней продолжительности цикла заражения и характеристик заражения внутри цикла. Следует подчеркнуть, что речь идет о среднестатистической информации по всему процессу заражений, а не об отдельных, пусть даже и многочисленных, случаях протекания болезни.

Используем допущения для создания модели заражений, позволяющей выявить объясняющие переменные (предикторы) для использования в регрессионном анализе. Рассмотрим в качестве первого приближения ситуацию, описываемую средними характеристиками показателей механизма заражений. Все среднестатистические оценки показателей зависят от продолжительности цикла заражения. Обозначим « q » среднестатистическое число дней, когда возможно заражение больными

людьми. Параметры моделей заражений будут иметь индекс «q», подчеркивающий, что оценки параметров получены при определенной продолжительности цикла заражений. В этом случае число новых зараженных X_{n+1} на следующий $n + 1$ день эпидемии, можно представить в виде суммы двух слагаемых. Первое слагаемое, отражает заражения выявленными больными, оно равно произведению $Q_{q,n}A_{q,n}R_{q,n}$, где: $Q_{q,n}$ - число выявленных заражающих; $A_{q,n}$ - интенсивность заражения; $R_{q,n}$ - число здоровых людей, которые потенциально могут быть заражены. Второе слагаемое, назовем его «аддитивной составляющей», определяется числом не выявленных на предыдущем этапе больных, обозначим его $C_{q,n}$. С учетом этих двух составляющих можно записать, что общее число зафиксированных зараженных на следующий день определяется как:

$$X_{n+1} = Q_{q,n}A_{q,n}R_{q,n} + C_{q,n}, \quad (1)$$

При таком обозначении число заражающих людей $Q_{q,n}$ будет равняться сумме числа больных в каждый предшествующий день, рассматриваемого цикла, когда от них возможно заражение. Число заражающих равно:

$$Q_{q,n} = \sum_{i=n-q}^n X_i,$$

где X_i число выявленных заразных больных в каждый из q дней цикла. Как будет показано ниже, оценки параметров моделей при разных средних продолжительностях циклов заражений оказываются взаимосвязанными. Выбор циклов заражений в регрессионных моделях будет определяться целями анализа процессов заражений.

Число здоровых людей $R_{q,n}$, которые потенциально могут оказаться заражёнными, представим с использованием понятия «потенциал заражения $L_{q,n}$ ». Будем понимать под потенциалом заражения $L_{q,n}$ оценку максимального числа людей, которые могут быть заражены в рассматриваемом цикле заражения. Число здоровых людей в каждый момент времени можно представить как разность между потенциалом заражения $L_{q,n}$ и числом уже зараженных людей:

$$R_{q,n} = L_{q,n} - \sum_1^n X_i.$$

Значение потенциала заражения $L_{q,n}$ может меняться вместе со сменой доминирующих механизмов заражения, которые сейчас обозначают как «волны заражений». Среднестатистическая оценка интенсивности заражения будет так же зависеть от выбранного цикла заражения - $A_{q,n}$. Тогда выражение (1) можно записать:

$$X_{n+1} = A_{q,n} \sum_{i=n-q}^n X_i (L_{q,n} - \sum_1^n X_i) + C_{q,n}; \text{ или, раскрыв скобки:}$$

$$X_{n+1} = A_{q,n}L_{q,n} \sum_{i=n-q}^n X_i - A_{q,n} \sum_{i=n-q}^n X_i \sum_1^n X_i + C_{q,n}. \quad (2)$$

Произведение $A_{q,n}L_{q,n}$ дает оценку числа людей, которых мог бы заразить один больной за один день при условии, что число здоровых равно потенциалу заражений $L_{q,n}$. Но реально часть людей уже заражена. Поэтому второе слагаемое – поправка, корректирующая число заражений с учетом числа больных в рассматриваемой популяции. Третье слагаемое $C_{q,n}$ отражает при данном цикле заражений q аддитивную составляющую, не связанную с числом выявленных заражающих.

Чтобы вычислить на следующем шаге $n + 1$ число новых зараженных людей необходимо знать следующие параметры: q – сколько дней мы рассматриваем больных людей, как заразных; с какой интенсивностью они заражают $A_{q,n}$ при потенциале заражения $L_{q,n}$. Эти параметры нам неизвестны и у нас нет априорной информации для их однозначного определения. Кроме того, изначально неизвестен вклад аддитивной составляющей в число новых больных. Но в выражении (2) в двух первых слагаемых есть два параметра, рассчитываемые по статистике заражений, которые могут

являться объясняющими переменными (предикторами) в регрессионных моделях. Рассмотрим возможные регрессионные модели с применением таких предикторов.

1 Модели регрессионного анализа динамики заражений

Мы не имеем достоверных данных о типе цикла заражения, а кроме того, допускается изменение типа заражений, поэтому мы не можем выделить единый, универсальный тип регрессионной модели, и задать ее параметры. Эта ситуация порождает два подхода к выбору моделей регрессионного анализа динамики заражений: во-первых, можно поставить задачу наиболее точной аппроксимации имеющихся данных, отдавая себе отчет, что при этом крайне трудно интерпретировать результаты; во-вторых, можно поставить задачу интерпретации полученных коэффициентов модели в категориях механизмов заражений, разумно уменьшая точность аппроксимации.

Для решения первой задачи включим на первом этапе в регрессионную модель набор альтернативных циклов заражений с продолжительностью от минимального (1) до максимального (q), такую модель можно представить как:

$$X_{q,n+1} = \sum_1^q b_{1,q} \sum_{n-q}^n X_i + \sum_1^q b_{2,q} \sum_{n-q}^n X_i \sum_1^n X_i + Const + \xi, \quad (3)$$

где $b_{1,q}, b_{2,q}$ коэффициенты регрессионной модели, которые соответствуют циклам по которому они рассчитаны. Индексация коэффициентов регрессионных моделей, выполнена следующим образом. Коэффициент $b_{1,q}$ при предикторе « $\sum_{n-q}^n X_i$ » имеет первую цифру в индексе «1», вторая цифра в индексе равна циклу заражения « q ». Коэффициенты модели $b_{2,q}$ при предикторах типа « $\sum_{n-q}^n X_i \sum_1^n X_i$ » имеют первый индекс «2», затем следует значение цикла заражения « q ». Поскольку в итоговую модель могут входить слагаемые, представляющие несколько циклов заражений, то константа - $Const$ и невязки модели - ξ не имеют индекса.

Слагаемые и, соответственно, коэффициенты этой регрессионной модели определяются в статистическом пакете SPSS процедурой пошагового включения переменных. На первом шаге при помощи дисперсионного анализа проверяется нулевая гипотеза об отсутствии линейной зависимости в модели. Для предлагаемой структуры модели на данных о заражениях любых стран дисперсионный анализ всегда показывает малую вероятность нулевой гипотезы, все линейные модели данной структуры корректны. Поэтому не будем приводить результаты расчетов по проверке этой нулевой гипотезы – исходные данные соответствуют линейной модели.

Еще необходимо оценить гипотезу о равенстве нулю константы в модели. Что бы различать модели с константой и без константы в случае выполнения условия о ненулевой константе введем дополнительный индекс «с» в коэффициенты модели.

Сама процедура пошагового отбора обеспечивает ненулевые коэффициенты отобранных переменных.

Таблица 1. Коэффициенты регрессионной модели (3) с константой и значимость гипотез о равенстве нулю коэффициентов.

	Нестандартизованные коэффициенты	Стандартизованные коэффициенты	Значимость нулевой гипотезы
$Const$	28,969		,004
$b_{1,1}$	1,086	1,087	,000
$b_{2,1}$	2,219E-05	2,033	,000
$b_{2,2}$	-1,082E-05	-1,925	,000
$b_{1,2}$	-,086	-,166	,003

Рассмотрим в качестве примера применение модели с константой (3) для данных Швеции с 9 марта по 27 октября 2020 года. Скорректированный R-квадрат такой модели равен 0,947. Выбор страны и диапазона используемых данных определяется следующими причинами. Швеция относится к странам, где долго не предпринимались попытки управления эпидемией, следовательно, это динамика эпидемии «в чистом виде». Ну а выбор данных определялся тем, что после 27 октября 2020 года в данных Швеции начали появляться пропуски и суммарные за несколько дней значения,

построить предлагаемые модели по таким данным невозможно. Наряду с коэффициентами модели приведены и стандартизированные коэффициенты (бета-коэффициенты), они позволяют сравнивать значимости слагаемых регрессионной модели.

Расчеты числа заражений по приведенной модели и реальные данные представлены на рисунке 1.



Рис. 1. Исходные данные и аппроксимация по модели.

В регрессионной модели оказалась значимая константа, это говорит о том, что в Швеции за рассматриваемый период ежедневно были в среднем по двадцать девять необнаруженных заражений.

Модели без константы имеют структуру аналогичную (3), но для разделения коэффициенты моделей без константы обозначены « d » вместо « b » с одинаковой индексацией, приведенной выше.

Расчеты по модели (3) без константы дают коэффициенты, представленные в таблице 2.

Таблица 2. Коэффициенты регрессионной модели (3) без константы и значимость гипотез о равенстве нулю коэффициентов.

	Нестандартизованные коэффициенты	Стандартизованные коэффициенты	Значимость нулевой гипотезы
$d_{2,1}$	2,093E-05	1,494	,000
$d_{2,2}$	-1,102E-05	-1,547	,000
d_{17}	,308	1,958	,000
$d_{1,3}$	-,340	-,983	,000
$d_{2,4}$	4,064E-06	1,107	,000
$d_{2,8}$	-3,678E-06	-1,852	,000
$d_{1,12}$	-,091	-,961	,000
$d_{2,13}$	1,094E-06	,854	,000

Такая сложная модель получилась в соответствии с критерием максимальной точности описания. Но увеличивая сложность модели, прибавляя новые предикторы, мы не всегда существенно повышаем точность. Это хорошо видно по тому, как меняются значения скорректированного R-квадрат при добавлении новых, уточняющих предикторов, эти изменения представлены на рисунке 2.

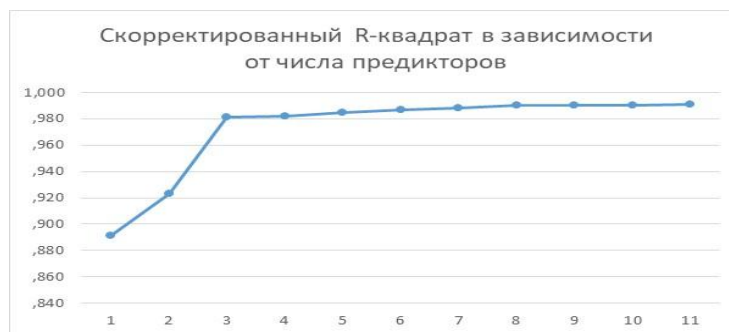


Рис. 2. Изменение значений R-квадрат в зависимости от сложности модели

Как видно существенное увеличение точности происходит при использовании трех оптимальных предикторов, затем скорректированный R-квадрат увеличивается лишь на 0,01. Но даже если мы, с целью упростить модель, возьмем только три первых предиктора, интерпретировать коэффициенты модели в категориях механизмов заражения очень сложно, и, тем более, это не удастся в полной модели, приведенной в таблице.

Можно сделать следующие заключения по использованию моделей (3) для описания динамики эпидемий.

- Модели структуры (3) с константой и без константы могут использоваться для анализа, а их результаты в плане точности описания динамики заражений можно рассматривать как близкие.
- Модели позволяют аппроксимировать динамику заражений с учетом изменений в разные дни недели.
- Главная проблема моделей структуры (3) заключается в невозможности интерпретировать значения их коэффициентов в категориях механизмов заражений.
- Целесообразно использовать менее точные, но более простые по структуре модели.

Для решения второй задачи – получения интерпретируемых в категориях заражений коэффициентов регрессионной модели воспользуемся следующей структурой модели:

$$X_{n+1} = b_{1,q^*} \sum_{n-q^*}^n X_i + b_{2,q^*} \sum_{n-q^*}^n X_i \sum_1^n X_i + \text{Const}_{q^*} + \xi_{q^*}, \quad (4)$$

где « q^* » среднее значение продолжительности единственного цикла, выбираемого базовым. Это приведет к некоторой потере точности аппроксимации, поскольку динамика будет описываться более сглажено, в среднем, без сильных изменений по дням недели. Коэффициенты данной структуры модели (4) определяются в статистическом пакете SPSS процедурой принудительного включения, но затем для каждого из них оценивается значимость нулевой гипотезы. В дальнейшем используются только модели с коэффициентами, значимость нулевой гипотезы которых не более 0,05.

Как и для обозначений в модели (3) в модели (4) случае значимой нулевой гипотезы о равенстве нулю константы воспользуемся другими обозначениями коэффициентов:

$$X_{n+1} = d_{1,q^*} \sum_{n-q^*}^n X_i + d_{2,q^*} \sum_{n-q^*}^n X_i \sum_1^n X_i + \mu_{q^*}. \quad (5)$$

где: d_{1,q^*} и d_{2,q^*} - коэффициенты моделей, аналогичные коэффициентам моделей с константами, имеющим такие же индексы; μ - невязки моделей.

Рассмотрим, какую информацию можно извлечь из коэффициентов таких регрессионных моделей. Как и в модели (3) значения константы Const_{q^*} являются оценками числа заражений не связанных с выявленными больными, это «аддитивная составляющая заражений» $C_{q^*} = \text{Const}_{q^*}$.

При отрицательном значении коэффициентов b_{2,q^*} в моделях (4), или d_{2,q^*} в (5) на основании значений их коэффициентов можно получить оценки параметров моделей заражений, то есть интерпретировать полученные результаты в категориях заражений. Поскольку используются два варианта регрессионных моделей, то параметрам механизмов заражений оцениваемых на основании моделей с константой (4) присваивается дополнительный индекс «с». Сравнивая параметры моделей заражений (2) и коэффициенты регрессионных моделей (4) можно выявить их соответствие $A_{q^*,c} L_{q^*,c} = b_{1,q^*}$, а оценка интенсивности заражения будет равняться $A_{q^*,c} = -b_{2,q^*}$. Оценка потенциала заражения по модели (4) будет равна:

$$L_{q^*,c} = \frac{b_{1,q^*}}{-b_{2,q^*}} = \frac{A_{q^*,c} L_{q^*,c}}{A_{q^*,c}}. \quad (6)$$

Получаемая оценка значения потенциала зависит от величин аддитивной составляющей, которые при нормальном развитии ситуации могут уменьшаться, становясь даже статистически незначимыми, поэтому часто более последовательные по времени результаты получаются на основании моделей (5) без константы.

Оценка потенциала заражения по модели (5) будет равна:

$$L_{q^*} = \frac{d_{1,q^*}}{-d_{2,q^*}} = \frac{A_{q^*} L_{q^*}}{A_{q^*}}. \quad (7)$$

Значение коэффициента $d_{1,q^*} = A_{q^*} L_{q^*}$ можно трактовать как среднестатистическую оценку скорости заражения одним больным за один день при условии, что вся совокупность, описываемая потенциалом заражения, состоит из здоровых людей. Тогда произведение $q^* d_{1,q^*} = q^* A_{q^*} L_{q^*}$ будет обобщенной характеристикой, отражающей специфику заражений в анализируемой динамике – средней гипотетической оценкой числа людей, зараженных одним больным за весь цикл заражений.

Вне зависимости от типа модели (с константой или без) знак второго коэффициента b_{2,q^*} или d_{2,q^*} характеризует особенность динамики. Отрицательное значение коэффициента свидетельствует, что число зараженных возрастает, стремясь к значению потенциала заражений, это говорит о том, что эпидемия развивается или относительно стабильно, или замедляется. Положительное значение коэффициента означает, что эпидемия усиливается и идет переходный процесс к большему потенциалу заражений. Механизм, приводящий к увеличению потенциала заражений, заключается в мутациях вируса с увеличением заражающей способности. Вероятность таких мутаций зависит от числа зараженных людей, как это видно во втором слагаемом.

Представление об основных закономерностях изменения оценок параметров механизмов заражений, в зависимости от выбираемой продолжительности цикла заражений q^* дают расчеты по коэффициентам моделей (5) с разными продолжительностями циклов заражений, приведенные в таблице 3. Оценки получены по данным о числе заражений в России с 15 марта 2020 года по 11 апреля 2021 года.

Таблица 3. Точность моделей и оценки параметров механизмов заражений, полученные на основании моделей разных циклов заражений.

Средний цикл заражений	Значения скорректированных R-квадратов для моделей средних циклов	Оценки потенциала заражений в зависимости от продолжительности среднего цикла	Произведение оценок интенсивности на число дней в цикле
3	,997	71 496 491	1,04
4	,997	55 173 540	1,05
5	,997	45 361 673	1,06
6	,997	38 642 437	1,07
7	,996	33 956 970	1,08
8	,996	30 558 838	1,09
9	,995	27 834 225	1,10
10	,994	25 557 888	1,12
11	,994	23 650 831	1,13
12	,993	22 018 082	1,14
13	,992	20 618 806	1,15
14	,991	19 423 483	1,16
15	,991	18 412 526	1,17
16	,990	17 522 414	1,18
17	,989	16 725 900	1,19
18	,988	16 007 950	1,20
19	,987	15 360 319	1,21
20	,986	14 774 297	1,23
21	,985	14 247 010	1,24
22	,984	13 777 899	1,25

Для рассмотрения связей параметров заражений, получаемых на основании моделей разных циклов заражений, были рассчитаны регрессионные модели типа (5) для циклов заражений в диапазоне от 3 до 22 дней. По формуле (7) были рассчитаны оценки потенциалов для циклов заражений. Результаты расчетов приведены на рисунке 3, представляется вероятным, что значения оценок соответствует степенному закону. Для проверки этого предположения была рассмотрена зависимость логарифмов оценок потенциалов от логарифмов длительности циклов заражений, она приведена на рисунке 4. При степенном законе распределения между этими переменными должна существовать линейная зависимость. Как видно на рисунке 4 эта зависимость выполняется с высокой точностью, что свидетельствует о существовании степенной зависимости между продолжительностью циклов и оценками потенциалов заражений.

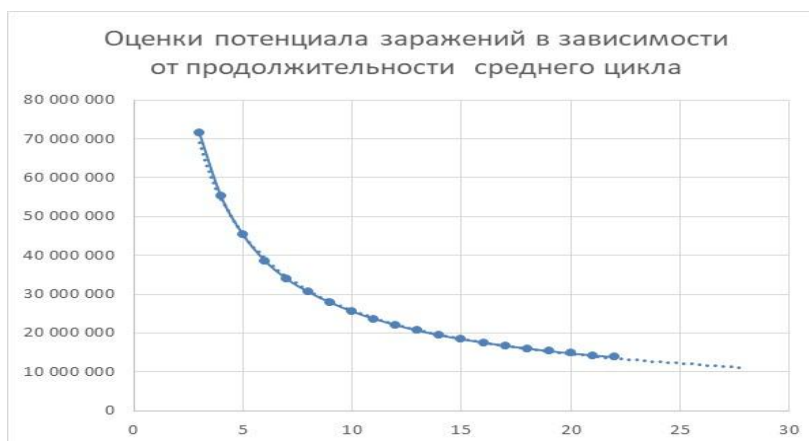


Рис. 3. Зависимость оценок потенциала от продолжительности циклов заражений

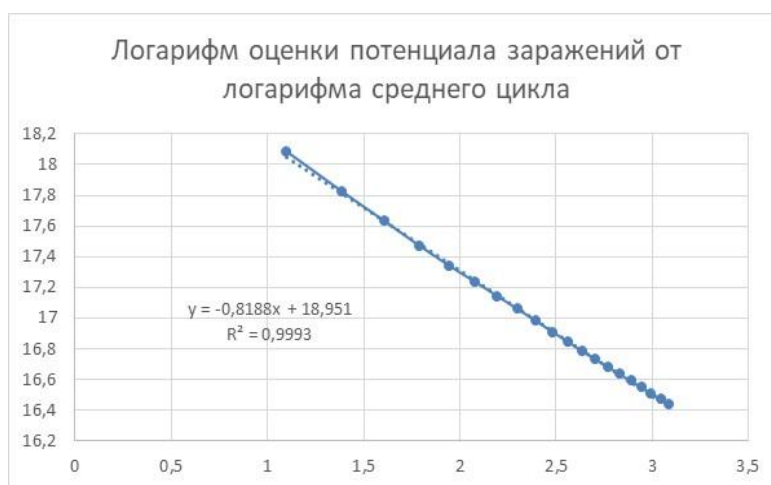


Рис. 4. Зависимость логарифмов оценок заражений от логарифмов продолжительности циклов

Особенно близки результаты к степенной зависимости, начиная с модели семидневного цикла заражения. При допущении, что перед нами степенной закон распределения потенциалов заражений, можно сделать вывод, что должна проявляться масштабная инвариантность по числу дней циклов заражений. Это значит, что можно выбирать для исходных расчетов любое значение продолжительности цикла заражений и полученные в модели результаты окажутся инвариантными и для других циклов. Можно выбирая одинаковый цикл заражений получать возможность сравнивать потенциалы заражений разных стран между собой. Кроме того, можно проводить сравнительный анализ динамики заражений в разных странах по значениям z , при достаточном приближении к степенной зависимости.

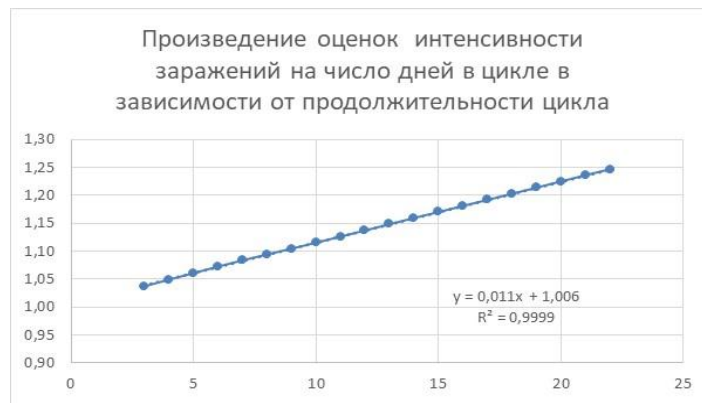


Рис. 5. Зависимость средних оценок возможного числа заражений от продолжительности цикла

Значения оценок другой характеристики механизмов заражений $q^*A_q^*L_q^*$ – гипотетической средней оценки числа людей, заражаемых одним больным за весь цикл заражений, при моделях разной продолжительности заражения оказываются также связаны. На рисунке 5 представлены расчеты этого параметра, они хорошо соответствуют линейной зависимости.

Оценки параметров механизмов заражений связаны со средними днями циклов функциональными зависимостями с высокой точностью.

Модель семидневного цикла заражений является в некоем роде компромиссом между точностью описания динамики и выявления ее основных закономерностей, сглаживающих изменения числа заражений в разные дни недели. Результаты аппроксимации динамики заражений в России регрессионной моделью (5) с семидневным циклом заражения приведены на рисунке 6.



Рис. 6. Динамика заражений в РФ и ее аппроксимация моделью семидневного среднего цикла

Как видно на рисунке 6 модель аппроксимирует динамику «в среднем», описывая наиболее точно динамику последнего времени. Более высокая точность описания динамики по последним результатам, по сравнению с началом эпидемии, определяется структурой регрессионной модели, ибо последние данные оказываются более «весомыми».

Заключение

1. Предложенные параметры механизмов заражений «интенсивность», «потенциал» и «продолжительность цикла заражений» позволяют на основании самых общих предположений выявить предикторы (объясняющие переменные) и структуру моделей регрессионного анализа, обеспечивающие описание динамики заражений с высокой точностью.
2. Предлагаемая модель механизмов заражений позволяет в дальнейшем доработать ее для описания динамики с учетом динамики числа людей, получивших вакцину.
3. На базе предложенного описания механизмов заражений реализуются два типа моделей регрессионного анализа – с константой и без константы, что позволяет рассмотреть разные типы эпидемических ситуаций. Оба варианта регрессионных моделей оказываются взаимосвязанными и описывают динамику с высокой точностью. Применяемая структура

моделей регрессионного анализа динамики COVID 19 позволяет реализовать на ее базе как модели, описывающие различия заражений в разные дни недели, так и модели, отражающие динамику «в среднем».

4. Структуры регрессионных моделей позволяют на основании значений коэффициентов выявлять переходные процессы, порождаемые изменением механизмов заражений.
5. На основании коэффициентов регрессионных моделей можно получать взаимосвязанные оценки потенциалов и интенсивностей заражений для различных предположений о продолжительности цикла заражений. Принимая допущения об одинаковой продолжительности циклов заражений можно сравнивать механизмы заражений разных стран в категориях потенциалов и интенсивностей. Подобные оценки могут быть использованы и для моделирования общих закономерностей эпидемий.
6. Регрессионные модели позволяют сравнивать динамику эпидемий разных стран, как в целом, так и по отдельным волнам заражений.

Литература

1. *Шрагенхайм* Управленческие дилеммы: Теория ограничений в действии. Пер. с англ. – М.: Альпина Бизнес Букс 2017
2. *Дрейпер Н., Смит Г.* Прикладной регрессионный анализ. Книга 1. В 2-х кн. М.: Финансы и статистика, 1986. — 366 с.
3. *Вапник В.Н., Глазкова Т.Г., Коцеев В.А. и др.,* под ред. В.Н.Вапника Алгоритмы и программы восстановления зависимостей - Москва : Наука, 1984. - 815 с.
4. *Thomas Hale , Noam Angrist , Rafael Goldszmidt, Beatriz Kira , Anna Petherick , Toby Phillips, Samuel Webster, Emily Cameron-Blake , Laura Hallas, Saptarshi Majumdar, and Helen Tatlow.* (2021). “A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker).” *Nature Human Behaviour.*