

**Н. П. Горидько  
Р. М. Нижегородцев**

## **ФАКТОРНЫЕ МОДЕЛИ РАСПРОСТРАНЕНИЯ КОРОНАВИРУСА И ЛОГИКА ПРОТИВОДЕЙСТВИЯ: ОПЫТ США**

**Аннотация.** Построены адекватные и значимые факторные регрессионные модели, выражающие связь между показателями заболеваемости и смертности по данным выборки 50 штатов США и объясняющими переменными, построенными на основе параметров численности населения, площади территории, наличия выхода к водоему и других физико-географических (в том числе климатических) факторов. На основе полученных результатов делаются выводы об адекватности противоэпидемиологических мер, принятых для борьбы с распространением новой коронавирусной инфекции.

**Ключевые слова:** коронакризис, факторные регрессионные модели, заболеваемость, противоэпидемиологические меры, распространение коронавируса.

### **Постановка задачи и исходные данные**

Задача заключается в том, чтобы выявить факторы, влияющие на распространение коронавируса. Эта задача решалась на примере различных штатов США по официальным статистическим данным.

Для корректного проведения расчетов по панельным данным важно обеспечить однородность условий, при которых эти

данные получены, и сопоставимость процессов, протекающих в наблюдаемых элементах выборки. Поэтому выбор объекта, на примере которого проводилось исследование, обусловлен прежде всего тем, что между разными штатами США нет различий в методике подсчета численности заболевших коронавирусом и умерших от него (включая осложнения, возникшие по причине заражения коронавирусом). В том числе важно, что одинакова точка отсчета – момент, когда начался официальный подсчет числа заболевших COVID-19 и умерших от него. В то же время, территория США достаточно велика по площади, что обеспечивает достаточное для построения значимых моделей разнообразие условий распространения вируса, в том числе природно-климатических.

Статистические данные о заболеваемости и смертности от коронавируса в реальном времени за весь период пандемии по 02.06.2020 включительно взяты с сайта [1]. Данные о численности населения штатов содержатся в [2]. Данные о среднегодовой температуре на территории штатов приведены в источнике [3]. Сведения о количестве снежных дней в году и о суммарной годовой высоте снежного покрова взяты с сайтов [4] и [5]. Кроме того, в качестве объясняющих переменных использовались данные о географической широте наиболее крупного (по численности населения) города каждого штата и о наличии выхода к водоему (приморское или приозерное расположение), которое моделировалось при помощи дамми-переменных.

Также была добавлена дамми-переменная, маркирующая Нью-Йорк, показавший аномально высокий уровень заболеваемости. Вероятно, этот факт обусловлен тем, что различные факторы, влияющие на распространение вируса, в частности, близость к водоемам и высокая плотность населения, наложились друг на друга в этом штате, и это привело к негативной синергии, вызвав всплески заболеваемости и смертности в течение рассматриваемого периода, намного превосходящие действие простой суммы факторов, повлиявших на итоговые значения объясняемых показателей.

Все модели были построены по данным 50 штатов США за один и тот же период времени. В качестве программного аппарата для построения регрессионных моделей был использован MS Excel.

В результате были построены регрессионные модели, адекватно объясняющие вариацию следующих переменных:

Y1 – количество заражений в абсолютном выражении на территории штата с начала пандемии до 02.06.2020;

Y2 – смертность нарастающим итогом от COVID-19 в абсолютном выражении на территории штата с начала пандемии до 02.06.2020;

Y3 – количество заражений на 1 тыс. жителей штата с начала пандемии до 02.06.2020;

Y4 – смертность нарастающим итогом от COVID-19 на 1 тыс. жителей штата с начала пандемии до 02.06.2020.

В качестве объясняющих переменных в составе итоговых моделей, в каждой из которых все регрессоры оказались значимыми на уровне доверия не менее чем 95%, были использованы следующие величины:

X1 – плотность населения (количество жителей на 1 кв. м площади) штата по состоянию на 01.07.2019 г.;

X2 – нормированная по максимальному значению величина – численность населения штата;

X3 – нормированная по максимальному значению величина – численность населения в квадрате, деленная на площадь штата;

D1 – наличие озера на территории штата (D1=1 – озеро есть и D1=0 – озера нет);

D2 – наличие выхода к водоему на территории штата (D2=1 – озеро, D2=0,5 – выход к морю, D2=1,5 – есть озеро и выход к морю);

D3 – фиктивная переменная для штата Нью-Йорк (D3=1 для Нью-Йорка и D3=0 для других штатов).

## **Основные результаты расчетов**

Ни одна из построенных моделей не показала значимой связи между заболеваемостью коронавирусом и климатическими параметрами: количеством снежных дней в году, высотой снежного покрова и географической широтой крупнейших городов, нормированной по отношению к наиболее густонаселенному городу самого северного штата – Аляски.

Также незначимой оказалась дамми-переменная, характеризующая приморское положение штата. Зато введение переменной D1, отражающей наличие озера, улучшало объясняющую способность большинства моделей, и этот регрессор почти неизменно оказывался значимым. Иллюстрацией данного факта может служить сопоставление характеристик моделей (1) и (2) (см. таблицы 1 и 2), объясняющих вариацию удельной заболеваемости населения штата через величину плотности населения.

Как видим из приведенных таблиц, характеристики модели (2) несколько лучше, нежели модели (1). Эта закономерность, связанная с улучшением качества модели при включении переменной D1, наблюдается для большинства построенных нами моделей.

$$Y_3 = 2,11 + 0,043 * X_1 \quad (1)$$

**Таблица 1. Эконометрические характеристики модели (1)**

<i>Регрессионная статистика</i>					
Множественный R	0,793604				
R-квадрат	0,629808				
Нормированный R-квадрат	0,622096				
Стандартная ошибка	2,656263				
Наблюдения	50				
<i>Дисперсионный анализ</i>					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Значимость F</i>
Регрессия	1	576,1884	576,1884	81,66246	6,26E-12
Остаток	48	338,6751	7,055732		
Итого	49	914,8636			
	<i>Коэффициенты</i>	<i>Стандартная ошибка</i>	<i>t-статистика</i>	<i>P-Значение</i>	
Y3-пересечение	2,109001	0,488132	4,320555	7,79E-05	
X1	0,042721	0,004727	9,036728	6,26E-12	

$$Y_3 = 1,811 + 0,042 * X_1 + 2,058 * D_1 \quad (2)$$

**Таблица 2. Эконометрические характеристики модели (2)**

<i>Регрессионная статистика</i>	
Множественный R	0,812911862
R-квадрат	0,660825695
Нормированный R-квадрат	0,646392746
Стандартная ошибка	2,569453186
Наблюдения	50

  

<i>Дисперсионный анализ</i>					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Значимость F</i>
Регрессия	2	604,5653521	302,2826761	45,78590884	9,22E-12
Остаток	47	310,2982147	6,602089675		
Итого	49	914,8635669			

  

	<i>Коэффициенты</i>	<i>Стандартная ошибка</i>	<i>t-статистика</i>	<i>P-Значение</i>
Y3-пересечение	1,811291933	0,493532105	3,670058974	0,000618055
X1	0,042242916	0,004578747	9,22586844	4,03721E-12
D1	2,057541969	0,992445552	2,073203881	0,043660172

Модель (3) является одной из немногих, в которых удалось оценить факторную зависимость удельной смертности населения штатов США. Наиболее адекватная модель с высокой объясняющей способностью ( $R^2=77\%$ ) получается при включении «обобщенной» дамми-переменной D2, комплексным образом выражающей положение территории штата относительно водоемов.

В моделях (4)-(6) в качестве регрессора участвует дамми-переменная D3, отражающая особенности штата Нью-Йорк, который продемонстрировал резкие выбросы в сторону возрастания показателей как заболеваемости населения, так и смертности. Обратим внимание на то, что в этих трех моделях переменная D3 является наиболее значимой (см. таблицы 4-6) по сравнению с другими регрессорами: ее *t*-статистика Стьюдента максимальна в сравнении с другими регрессорами, а *p*-

значение является микроскопически малым, что выражает крайне низкую вероятность ошибки и, следовательно, высокую степень доверия к найденному значению соответствующего коэффициента при данной переменной.

$$Y_4 = 0,003 * X_1 + 0,146 * D_2 \quad (3)$$

**Таблица 3. Эконометрические характеристики модели (3)**

<i>Регрессионная статистика</i>					
Множественный R					0,878038
R-квадрат					0,770951
Нормированный R-квадрат					0,745346
Стандартная ошибка					0,205852
Наблюдения					50
<i>Дисперсионный анализ</i>					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Значимость F</i>
Регрессия	2	6,846187	3,423093	80,78094	6,2E-16
Остаток	48	2,034001	0,042375		
Итого	50	8,880187			
	<i>Коэффициенты</i>	<i>Стандартная ошибка</i>	<i>t-статистика</i>	<i>P-Значение</i>	
Y4-пересечение					
X1	0,003059	0,000373	8,19441	1,12E-10	
D2	0,14606	0,074174	1,969163	0,054722	

В случаях, когда переменные X2 и X3 оказываются значимыми в рамках одной модели, более значимой оказывается X3, у нее *p*-значение меньше, соответственно, более высок уровень доверия к величине регрессора, посчитанного при помощи используемого программного аппарата. А X2 в этих случаях играет роль демпфирующей переменной (она выполняет корректирующую функцию) и часто влияет на итоговый результат заболеваемости с отрицательным знаком. Сопоставление эконометрических характеристик моделей (5) и (6) (см. таблицы 5 и 6) иллюстрирует приведенный выше тезис.

$$Y1=8021,001+130413,6*X3+248497,4*D3+20339,6*D1 \quad (4)$$

**Таблица 4. Эконометрические характеристики модели (4)**

<i>Регрессионная статистика</i>					
Множественный R					0,964917
R-квадрат					0,931064
Нормированный R-квадрат					0,926569
Стандартная ошибка					16137,11
Наблюдения					50
<i>Дисперсионный анализ</i>					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Значимость F</i>
Регрессия	3	1,61788E+11	53929283482	207,0968	1,02E-26
Остаток	46	11978683889	260406171,5		
Итого	49	1,73767E+11			
		<i>Коэффициенты</i>	<i>Стандартная ошибка</i>	<i>t-статистика</i>	<i>P-Значение</i>
Y1-пересечение		8021,001	2831,016394	2,833258324	0,006819
X3		130413,6	10248,93536	12,72460053	1,13E-16
D3		248497,4	18115,78391	13,71717514	7,21E-18
D1		20339,6	6613,23866	3,075587894	0,00353

$$Y1=832,248-5172,095*X2+10852,968*X3+23737,565*D3 \quad (5)$$

**Таблица 5. Эконометрические характеристики модели (5)**

<i>Регрессионная статистика</i>					
Множественный R					0,959206642
R-квадрат					0,920077381
Нормированный R-квадрат					0,914865037
Стандартная ошибка					1342,289608
Наблюдения					50
<i>Дисперсионный анализ</i>					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Значимость F</i>

Регрессия	3	954124258	318041419,3	176,5189	3,05E-25
Остаток	46	82880104	1801741,391		
Итого	49	1037004362			
		<i>Коэффициенты</i>	<i>Стандартная ошибка</i>	<i>t-статистика</i>	<i>P-Значение</i>
Y1-пересечение		832,2483146	255,961792	3,251455254	0,002152
X2		-5172,094916	1661,47194	-3,11295953	0,003182
X3		10852,96797	1331,69738	8,149725378	1,78E-10
D3		23737,56544	1445,74169	16,41895344	7,32E-21

$$Y1 = 667,582 - 6132,77 * X2 + 11189,6 * X3 + 22236,33 * D3 + 1894,044 * D1 \quad (6)$$

**Таблица 6. Эконометрические характеристики модели (6)**

<i>Регрессионная статистика</i>					
Множественный R		0,969585			
R-квадрат		0,940094			
Нормированный R-квадрат		0,934769			
Стандартная ошибка		1174,946			
Наблюдения		50			
<i>Дисперсионный анализ</i>					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Значимость F</i>
Регрессия	4	974881981,4	243720495,4	176,5454	6,89E-27
Остаток	45	62122380,5	1380497,344		
Итого	49	1037004362			
		<i>Коэффициенты</i>	<i>Стандартная ошибка</i>	<i>t-статистика</i>	<i>P-Значение</i>
Y1-пересечение		667,5821	228,0396778	2,927482295	0,005343
X2		-6132,77	1475,285973	-4,1570033	0,000143
X3		11189,6	1168,902099	9,572739983	2,01E-12
D3		22236,33	1323,394997	16,80248967	5,17E-21
D1		1894,044	488,4476106	3,877680955	0,000341



Отличие модели (6) от модели (5) в том, что добавляется дамми-переменная  $D1$ , обозначающая наличие озера на территории штата, и она является значимой в данной компании регрессоров, причем ее наличие немного увеличивает объясняющую способность модели ( $R^2$ ) и снижает сумму квадратов невязок.

В целом для объяснения данных о заболеваемости получилось больше адекватных и значимых моделей, чем для объяснения данных о смертности, что можно объяснить относительной расплывчатостью, нестрогостью критериев, на основании которых в качестве причины наступления смерти указывался коронавирус либо осложнения, наступившие вследствие заболевания им.

## **Обсуждение и основные выводы**

Выявив факторы, значимо влияющие на распространение коронавируса, становится возможным разработать эффективные меры противодействия этому процессу и оценить степень адекватности регулирующих воздействий, уже используемых правительствами различных стран.

Построенные модели подтвердили тот факт, что основным фактором, ускоряющим распространение коронавируса, является значительная плотность населения, его высокая концентрация на относительно небольшой площади территории. Также из приведенных моделей вытекает, что близость к озерам, замкнутым водоемам, со всех сторон окруженным сушей, выступает значимым фактором заражения. В отличие от морей, в которых вода постоянно обновляется и успевает очищаться, озера способствуют высокой концентрации вирусных заболеваний.

Следовательно, для противодействия распространению коронавируса необходимо, во-первых, соблюдение социально приемлемой физической дистанции между взаимодействующими индивидами (покупателем и продавцом, например), а во-вторых, соблюдение правил личной гигиены, простейшими из которых, по-видимому, являются привычки, приходя домой, переобуться (сменить уличную обувь на домашнюю) и помыть руки с мылом (либо иным способом обеспечить их дезинфекцию). Именно меры такого рода были приняты и рекомендованы в большинстве стран мира для обеспечения относительно безопасного образа жизни населения [6, 7].

При этом, хотя принятые меры нередко были избыточными и не всегда адекватными возникшим угрозам [8, 9], они принесли несомненный эффект и позволили, насколько это возможно, локализовать очаги заражения и выиграть время для того, чтобы сформировался коллективный иммунитет против уже известных штаммов COVID-19.

Что касается полной или частичной остановки хозяйственных процессов (так называемый локдаун) ради борьбы с коронавирусом, то меры такого рода однозначно следует отнести к разряду избыточных. Уже в 2022 году специалисты из Университета Джонса Хопкинса (США) провели оценки и выяснили, что локдаун улучшил сопротивляемость человеческих популяций и снизил смертность от этого вируса не более, чем на 0,2% [10], и эта цифра объективно находится в пределах допустимой погрешности проведенных измерений, она ничтожно мала в сравнении с издержками, которые понесла экономика, и с психологическими (в том числе психосоматическими) проблемами людей, испытавших шок от потери работы и от разрушения казавшихся устойчивыми социальных связей.

Поэтому одна из задач эффективной борьбы с распространением вируса состоит в том, чтобы обеспечить комплексный характер принимаемых мер и последовательную реализацию этих мер, высокий уровень гибкости хозяйственных процессов и их готовности к перенастройке на иных технических основаниях [11]. Сюда относится, в частности, необходимость перевода в онлайн-формат тех процессов взаимодействия, которые могут осуществляться дистанционно. Применение противозидемиологических мер санитарно-гигиенического характера возможно и без остановки хозяйственных процессов, вполне достаточно обеспечить соблюдение простых принципов личной гигиены значительной частью населения.

Как сказал Дмитрий Песков в июне 2022 года, «экономика стала заложницей политики, а политики многих стран во время коронавируса совершили огромное количество ошибок. Мир сейчас расплачивается за эти ошибки» [12]. Разумный подход заключается в том, чтобы выявлять эти ошибки и не повторять их в дальнейшем.

## ЛИТЕРАТУРА

1. Статистика заболеваемости коронавирусом COVID-19 в штатах в США [Электронный ресурс]. – Режим доступа: <https://ncov.blog/countries/us/>.
2. U.S. Census Bureau (web) [Электронный ресурс]. – Режим доступа: <https://www.citypopulation.de/en/usa/states/admin/>.
3. Average Annual Temperature for Each US State [Электронный ресурс]. – Режим доступа: <https://www.currentresults.com/Weather/US/average-annual-state-temperatures.php>.
4. U.S. Average Snow Days State Rank [Электронный ресурс]. – Режим доступа: <http://www.usa.com/rank/us--average-snow-days--state-rank.htm>.
5. Yearly snowfall averages [Электронный ресурс]. – Режим доступа: <https://www.currentresults.com/Weather/US/average-snowfall-by-state.php>.
6. Минздрав назвал три главных способа защиты от коронавируса / Lenta.Ru [Электронный ресурс, 13.06.2020]. – Режим доступа: <https://news.mail.ru/society/42174715/?frommail=1>.
7. Для спасения от эпидемии россиян будут разобщать [Электронный ресурс, 07.06.2020]. – Режим доступа: [https://yandex.ru/turbo/s/pnp.ru/social/dlya-spaseniya-ot-epidemii-rossiyan-budut-razobshhat.html?promo=navbar&utm\\_referrer=https%3A%2F%2Fzen.yandex.com](https://yandex.ru/turbo/s/pnp.ru/social/dlya-spaseniya-ot-epidemii-rossiyan-budut-razobshhat.html?promo=navbar&utm_referrer=https%3A%2F%2Fzen.yandex.com).
8. Нижегородцев Р. М., Горидько Н. П. Институциональные дилеммы коронакризиса: внешние эффекты, принятие решений и моральный выбор // Экономика коронакризиса: вызовы и решения: Сб. науч. трудов / Под ред. Р.М. Нижегородцева. – М.: ООО «НИПКЦ Восход-А», 2020. – С. 6-16.
9. Нижегородцев Р. М. Коронакризис: становление вирусно-цифровой экономики // Философия хозяйства. – 2020. – № 3. – С. 125-152.

10. Эксперты США заявили, что локдауны минимально повлияли на снижение смертности от COVID-19 [Электронный ресурс, 02.02.2022]. – Режим доступа: [https://news.mail.ru/society/49869222/?frommail=1&utm\\_partner\\_id=945](https://news.mail.ru/society/49869222/?frommail=1&utm_partner_id=945).
11. Нижегородцев Р. М. Тренды коронакризиса, деформация рынков и задачи экономической политики // Экономика коронакризиса: вызовы и решения: Сб. науч. трудов / Под ред. Р. М. Нижегородцева. – М.: ООО «НИПКЦ Восход-А», 2020. – С. 85-105.
12. Родин И. В Кремле анонсировали «чрезвычайно важную» речь Путина на ПМЭФ [Электронный ресурс, 13.06.2022]. – Режим доступа: [https://finance.rambler.ru/economics/48818447-v-kremle-anonsirovali-chrezvychayno-vazhnuyu-rech-putina-na-pmef/?utm\\_source=head&utm\\_campaign=self\\_promo&utm\\_medium=news&utm\\_content=news](https://finance.rambler.ru/economics/48818447-v-kremle-anonsirovali-chrezvychayno-vazhnuyu-rech-putina-na-pmef/?utm_source=head&utm_campaign=self_promo&utm_medium=news&utm_content=news).