

DOI: 10.18499/2225-7357-2019-8-1-9-13

УДК 572.08

14.03.01 – анатомия человека

© Коллектив авторов, 2019

Применение метода регрессионного моделирования в оценке изменений и прогнозировании физического статуса населения России

А. Н. Тимонин¹, Е. А. Бурляева¹, Н. С. Никитин¹, С. В. Ключкова², Д. Б. Никитюк^{1*}

¹ФГБУН «Федеральный исследовательский центр питания и биотехнологии», Москва, Россия

²ФГАОУ ВО «Первый Московский государственный медицинский университет им. И.М. Сеченова» Минздрава России (Сеченовский Университет), Москва, Россия

Цель исследования – проведение регрессионного моделирования физического статуса населения России с 1917 по 2017 годы и построение прогностической модели до 2035 года.

Материал и методы. В основе методологического подхода были использованы методы корреляционно-регрессионного анализа, в частности построение линейных регрессионных моделей. Сбор данных основан на архивных исторических документах. Статистическая обработка данных проводилась с использованием SPSS и Excel. Уровень значимости принимается равным 0.05.

Результаты. В результате исследования были сформулированы основные регрессионные модели для каждого из исследуемых показателей. В соответствии с приведенными регрессионными моделями был составлен прогноз до 2035 года по всем исследуемым показателям.

Заключение. Данный метод может рассматриваться как перспективный методологический инструмент при исследовании динамического ряда соответствующих показателей при выявлении не только тренда, но и скрытой колебательной компоненты данного ряда.

Ключевые слова: регрессионное моделирование; линейный тренд; население России; прогнозирование.

Application of the Regression Modeling Method in Assessing Changes and Predicting the Physical Status of the Russian Population

© A. N. Timonin¹, E. A. Burlyeva¹, N. S. Nikitin¹, S. V. Klochkova², D. B. Nikityuk^{1*}, 2019

¹The Federal Research Centre of Biotechnology and Food Safety, Moscow, Russia

²I.M. Sechenov First Moscow State Medical University, Moscow, Russia

The aim of the study was to perform regression modeling of the physical status of the Russian population from 1917 to 2017 and to build a prognostic model until 2035.

Material and methods. The methodological approach was based on the methods of correlation and regression analysis, namely, on the construction of linear regression models. Data collection was based on the archival historical documents. Statistical data were processed using SPSS and Excel. The significance level was assumed to be 0.05.

Results. The basic regression models for each of the studied parameters were formulated in the study. In accordance with the given regression models, the authors made a forecast for all the studied parameters up to 2035.

Conclusion. This method can be considered as a promising methodological tool in the study of the dynamic series of relevant parameters in identifying not only the tendency, but also the hidden oscillatory component of the series.

Keywords: regression modeling; linear trend; population of Russia; forecasting.

*Автор для переписки:

Никитюк Дмитрий Борисович
ФГБУН «Федеральный исследовательский центр питания и биотехнологии», Устьинский проезд, 2/14, г. Москва, 109240, Российская Федерация.

E-mail: dimitrynik@mail.ru

*Corresponding author:

Dmitrii Nikityuk
The Federal Research Centre of Biotechnology and Food Safety, Ust'inskii proezd, 2/14, Moscow, 109240, Russian Federation
E-mail: dimitrynik@mail.ru

Введение

Приложение методов статистического анализа, в частности, корреляционно-регрессионного анализа в области антропометрических исследований и изучении динамики физического статуса людей представляет собой перспективное научное направление, позволяющее, с известной точностью, осуществлять прогнозирование изменения изучае-

мых показателей в определенном временном диапазоне. Уровень точности подобного рода прогноза, который сам по себе является самостоятельным научно-исследовательским продуктом, во многом зависит от успеха разработанной эмпирической прогностической модели [1–3, 11]. Таким образом, разработка эмпирической регрессионной модели является целью подобного рода исследований. Методология корреляционно-регрессионного анализа на сегодняшний день, во многом стандартизирована и автоматизирована, что позволяет использовать ее широкому кругу специалистов, работающих в самых разных сферах научно-исследовательской деятельности, в том числе – в области антропологии. Особенно следует отметить эффективность этой методологии при анализе данных, полученных не только в результате хорошо спланированного эксперимента, но и при работе с архивными данными, полученными в различное время от различных источников. В частности, корреляционно-регрессионный анализ применим при исследовании динамики антропометрических данных населения России с 1917 по 2017 годы, оценке изменения физического статуса населения России и построению прогностической модели [12]. Такое исследование позволит выявить закономерности изменения основных антропометрических показателей в течение длительного времени наблюдения, установить закономерности этих изменений, а также прогнозировать их динамику в зависимости от того, насколько благоприятны были значимые факторы воздействия в данной социально-экономической системе.

Таким образом, целью данного исследования является проведение регрессионного моделирования физического статуса населения России с 1917 по 2017 годы и построение прогностической модели до 2035 года.

Материал и методы исследования

В основе методологического подхода были использованы методы корреляционно-регрессионного анализа, в частности построение линейных регрессионных моделей.

Динамический ряд антропометрических показателей с 1917 по 2017 годы был проанализирован посредством кусочно-линейной аппроксимации с периодом в 5 лет. Рабочая гипотеза 1: тренды кусочно-линейной аппроксимации хорошо линеаризованы, а динамика изменений угловых коэффициентов прогнозируема.

Динамический ряд антропометрических показателей с 1917 по 2017 годы был проанализирован посредством следующего алгоритма: построение линейной регрессии осуществлялось на временном промежутке, полученном с помощью арифметической прогрессии, с шагом в 5 лет, то есть линейная регрессия

выстраивалась для первых 10, 15 и так до 100 лет. Полученные значения коэффициентов детерминации, углового коэффициента и коэффициента сдвига функционально анализировали. Рабочая гипотеза 2: тренды линейных регрессий хорошо линеаризованы для определенного временного отрезка, и полученная регрессия является базовой для построения прогностической модели.

При построении прогностической модели динамического ряда антропометрических данных был использован метод информационной энтропии Шеннона для количественной оценки степени неопределенности прогноза с каждым шагом в 5 лет.

При построении прогностической модели использованы три основных сценария изменения динамики антропометрических показателей: пессимистический, отражающий наихудшие прогностические ожидания в указанном временном промежутке, оптимистический, отражающий наилучшие прогностические ожидания в указанном временном промежутке и усредненный, отражающий усредненные прогностические ожидания в указанном временном промежутке.

Сбор данных основан на архивных исторических документах [4–10]. Статистическая обработка данных проводилась с использованием SPSS и Excel. Используемые критерии согласия: критерий согласия относительно углового коэффициента и коэффициента сдвига со следующими нуль гипотезами: угловой коэффициент равен нулю, коэффициент сдвига равен нулю, сравниваемые угловые коэффициенты равны и сравниваемые коэффициенты сдвига равны. Построение линейных регрессий проводилась как по экстенсивным данным, так и по стандартизированным данным анализируемого показателя. Уровень значимости принимается равным 0.05 [1–3, 7, 11].

Результаты и их обсуждение

Из объективных исторических данных для характеристики динамики антропометрических данных с 1917 по 2017 годы были выбраны рост и масса тела. Тогда индекс массы тела (ИМТ) можно рассчитать по этим данным. Такие показатели, как абсолютная масса жировой, костной и мышечной ткани в данной работе использовались с 1975 по 2015 годы. Поэтому апробация и построение прогностической модели, а также выявление статистически значимых закономерностей проводилась, в основном по росту и массе тела. В соответствии с планом первым этапом исследования стала кусочно-линейная аппроксимация трендов роста и массы тела за каждые 5 лет.

С целью проверки 1-й рабочей гипотезы рассмотрим следующие параметры: частоту

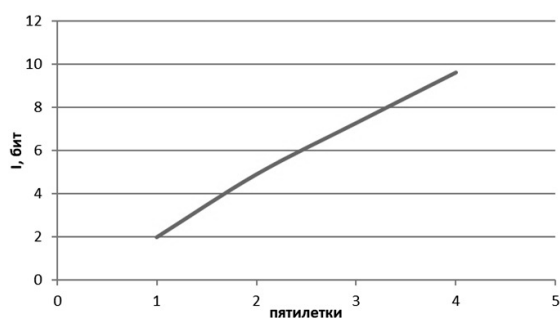


Рис. Зависимость степени неопределенности от количества пятилетних шагов.

хорошо линейаризованного тренда при кусочно-линейной аппроксимации, частоту изменений знака угловых коэффициентов, коэффициент вариации углового коэффициента и коэффициента сдвига. Если частота хорошо линейаризованного тренда (не менее 0.7 коэффициента детерминации) больше заданного уровня значимости, тогда кусочно-линейная аппроксимация значительно увеличивает вероятность нелинейного тренда в рамках предложенного анализа. Если частота изменений знака угловых коэффициентов больше заданного уровня значимости, тогда кусочно-линейная аппроксимация значительно увеличивает вероятность экстремумов первого рода в рамках предложенного анализа. Если коэффициенты вариации более 10% для углового коэффициента и коэффициента сдвига линейных аппроксимированных трендов, тогда уровень вариации коэффициентов значим.

По данным таблицы, для роста установлено, что частота хорошо линейаризованного тренда при кусочно-линейной аппроксимации составляет 0.35, что больше выбранного уровня значимости. Частота изменений знака угловых коэффициентов также составляет 0.35, что также больше уровня значимости. Коэффициент вариации для углового коэффициента составляет 241%, а для коэффициента сдвига – 2%. Исходя из полученных результатов, нельзя рассматривать метод кусочно-линейной аппроксимации адекватным инструментом при построении прогностической модели для роста, следовательно, 1-я рабочая гипотеза отклоняется.

Для массы тела установлено, что частота хорошо линейаризованного тренда при кусочно-линейной аппроксимации составляет 0.55, что больше выбранного уровня значимости. Частота изменений знака угловых коэффициентов также составляет 0.55, что также больше уровня значимости. Коэффициент вариации для углового коэффициента составляет 255%, а для коэффициента сдвига – 13%. Исходя из полученных результатов, нельзя рассматривать метод кусочно-линейной аппроксимации адекватным инструментом при построении прогностической модели для массы тела, поэтому 1-я рабочая гипотеза откло-

няется. Таким образом, кусочно-линейная аппроксимация с шагом в 5 лет является неадекватным методом в рамках данной работы.

В соответствии с планом исследования был проведен анализ данных для проверки 2-й рабочей гипотезы. Было установлено, что степень линейаризации резко увеличивается после определенного времени исследования: для роста это время составляет 35 лет, для массы тела – 50 лет. При этом линейаризация значимо не изменяется и остается на достаточно высоком уровне (коэффициент детерминации не менее 0.8). Коэффициенты вариации для угловых коэффициентов и коэффициентов сдвига составляют: рост – 8% для углового коэффициента и 0.2% – для коэффициента сдвига; масса тела – 14% для углового коэффициента и 2% – для коэффициента сдвига. Таким образом, 2-я рабочая гипотеза принимается.

На основании полученных результатов анализа были сделаны следующие выводы, используемые для построения прогностической модели динамики антропометрических показателей. В основе построения прогностической модели следует использовать рост как наиболее прогностически перспективный антропометрический показатель: коэффициент детерминации равен 0.9329, коэффициент вариации коэффициента сдвига менее 0.3%, каждая линейная регрессия, увеличенная от исходной на пять лет, достоверно не отличается от исходной линейной регрессии. Построение линейной регрессии роста, с шагом в пять лет можно сформулировать следующим образом: коэффициент сдвига остается постоянным, угловой коэффициент или остается неизменным, или увеличивается (уменьшается) на стандартную ошибку, или увеличивается (уменьшается) на стандартную ошибку, умноженную на 1.96 (так как анализ проводился на уровне значимости 0.05). Таким образом, при прогнозировании углового коэффициента для 1-й пятилетки существует 5 вариантов с исходной неопределенностью относительно вероятности реализации варианта.

Введем следующую функцию – функцию степени неопределенности, которая представляет собой среднее количество информации, необходимое для ее устранения. Тогда для 1-го шага достаточно 2 бит информации. В этом случае можно, опираясь на теорию информации, построить график зависимости уровня неопределенности от пятилетних шагов (рис.).

Из графика видно, что чем более по времени планируется сделать прогноз, тем большее количество информации (полезной информации) требуется для точного прогнозирования. В этих условиях рационально осуществить прогноз двух крайних (предельных) вариантов и вариант максимального усреднения.

Таблица 1

Основные значения регрессионных моделей

Изучаемые показатели	Угловой коэффициент	Коэффициент сдвига
Рост	0.11	166.9
Масса тела	0.25	57.0
Мышечная ткань (мужчины)	-0.30	31.7
Костная ткань (мужчины)	-0.29	15.4
Жировая ткань (мужчины)	0.50	15.8
Мышечная ткань (женщины)	-0.34	26.0
Костная ткань (женщины)	-0.20	14.1
Жировая ткань (женщины)	1.43	10.1

Таблица 2

Таблица средних прогностических величин роста, массы тела и ИМТ

Изучаемые показатели	Прогнозируемые периоды			
	2019–2023 гг.	2024–2028 гг.	2029–2032 гг.	2033–2035 гг.
Пессимистический прогноз				
Рост (см)	176.9±0.1	177.2±0.1	176.8±0.1	176.1±0.1
Масса (кг)	76.9±0.1	77.5±0.1	79.2±0.1	81.5±0.1
ИМТ	24.6±0.1	24.7±0.1	25.3±0.1	26.3±0.1
Оптимистический прогноз				
Рост (см)	178.0±0.1	179.7±0.1	180.9±0.1	181.9±0.1
Масса (кг)	78.0±0.1	76.1±0.1	74.2±0.1	72.0±0.1
ИМТ	24.6±0.1	23.6±0.1	22.7±0.1	21.8±0.1
Прогноз усредненного ожидания				
Рост (см)	176.9±0.1	177.5±0.1	178.1±0.1	178.5±0.1
Масса (кг)	76.9±0.1	77.5±0.1	78.1±0.1	78.5±0.1
ИМТ	24.6±0.1	24.6±0.1	24.6±0.1	24.6±0.1

Таблица 3

Таблица прогноза компонентного состава тела жителей РФ 2025 и 2035 годов

Изучаемые показатели	Прогнозируемые периоды			
	2025 г.		2035 г.	
	Мужчины	Женщины	Мужчины	Женщины
Абсолютное содержание жировой ткани (кг)	18.8±0.1	18.7±0.1	19.3±0.1	20.1±0.1
Абсолютное содержание мышечной ткани (кг)	29.9±0.1	23.9±0.1	29.6±0.1	23.6±0.1
Абсолютное содержание костной ткани (кг)	13.7±0.1	12.9±0.1	13.4±0.1	12.7±0.1

Таблица 4

Таблица высокоточного прогнозирования роста, массы тела и ИМТ от минимального до максимального значения

Min–Max	Рост, см	Масса тела, кг	ИМТ
	176.1–182.1	68.5–89.4	19.8–27.0

Первый вариант условно может быть обозначен как пессимистический. Алгоритм его построения: с каждой пятилеткой из исходного углового коэффициента следует вычесть стандартное отклонение углового коэффициента, умноженное на 1.96 – результат – следующая линейная регрессия.

Второй вариант условно может быть обозначен как оптимистический. Алгоритм его построения: с каждой 5-леткой к исходному угловому коэффициенту следует прибавить стандартное отклонение углового коэффициента, умноженное на 1.96 – результат – следующая линейная регрессия.

Третий вариант – вариант усредненного ожидания. Алгоритм его построения: угловой коэффициент – постоянная величина.

Важно отметить, что использование ИМТ в рамках данного исследования возможно в случае, если между угловыми коэффициентами роста и массы тела нет достоверных различий. Для проверки этого условия следует построить линейные регрессии стандартизированных данных роста и массы тела с последующей проверкой нулевой гипотезы о равенстве углового коэффициента и коэффициента сдвига. Для роста угловой коэффициент стандартизированных данных равен 0.033, а

для массы тела – 0.0314. Для роста коэффициент сдвига равен 1.68, а для массы тела этот показатель равен 1.60. Во всех случаях нулевая гипотеза принимается, что позволяет использовать рост и массы тела при вычислении ИМТ с последующим использованием расчетного параметра в прогностической модели.

Аналогичный подход был реализован для показателей абсолютной массы костной, мышечной и жировой ткани. В результате исследования были сформулированы основные регрессионные модели для каждого из исследуемых показателей (табл. 1).

В соответствии с приведенными регрессионными моделями был составлен прогноз до 2035 года (табл. 2–4).

Заключение

Цель нашего исследования была реализована с использованием регрессионного моделирования при сравнении линейных трендов, полученных во временных интервалах, рассчитанных с помощью арифметической прогрессии, а сравнение различных регрессий осуществлялось посредством использования стандартизированных данных. Данный метод может рассматриваться как перспективный методологический инструмент при исследовании динамического ряда соответствующих показателей при выявлении не только тренда, но и скрытой колебательной компоненты данного ряда.

Конфликт интересов

Авторы заявляют об отсутствии конфликта интересов.

Список литературы

1. Владимирский Б. М., Горстко А. Б., Ерусалимский Я. М. Математика. Общий курс. СПб.: Лань; 2008. 532–645 [Vladimirskii BM, Gorstko AB, Erusalimskii YM. Matematika. Obshchii kurs. Saint Petersburg: Lan'; 2008] (in Russian).
2. Еськов В. М., Хадарцев А. А., Филатова О. Е., Хадарцева К. А., Литовченко О. Г. Проблема оценки эффективности лечения на основе кинематической характеристики вектора состояния организма. Вестник новых медицинских технологий. 2015; 22(1):143–152 [Eskov VM, Khadartsev AA, Filatova OE, Khadartseva KA, Litovchenko OG. The Problem of Assessing the Effectiveness of Treatment Based on the Kinematic
- Characteristics of the State Vector of the Body. Journal of New Medical Technologies. 2015 Feb 11;22(1):143–52] (in Russian). doi: 10.12737/9096
3. Каркищенко Н. Н. Основы биомоделирования. М.: ВПК. 2004. 608 [Karkishchenko NN. Osnovy biomodelirovaniya. Moscow: VPK; 2004] (in Russian).
4. Миронов Б. Н. Историческая социология России: учебник в 2 ч. 2-е изд. М.: Юрайт; 2018; 1: 105–162 [Mironov BN. Istoricheskaya sotsiologiya Rossii: uchebnik. 2nd ed. Moscow: Yurait; 2018.1] (in Russian).
5. Миронов Б. Н. Историческая социология России: учебник в 2 ч. 2-е изд. М.: Юрайт; 2018; 2: 98–126 [Mironov BN. Istoricheskaya sotsiologiya Rossii: uchebnik. 2nd ed. Moscow: Yurait; 2018.2] (in Russian).
6. Миронов Б. Н. Модернизация России в XVIII–XX вв. как конвергенционный проект. Социологические исследования. 2018; 6(410): 91–103 [Mironov BN. Modernization of Russia in XVIII–XX centuries as a convergence project. Sociological Studies (Socis). 2018;6(410):91–103] (in Russian).
7. Миронов Б. Н. Российская империя: от традиции к модерну. Санкт Петербург: Дмитрий Буланин; 2018; 1: 126–265 [Mironov BN. Rossiiskaya imperiya: ot traditsii k modernu. Saint Petersburg: Dmitrii Bulanin; 2018.1] (in Russian).
8. Миронов Б. Н. Российская империя: от традиции к модерну. Санкт Петербург: Дмитрий Буланин; 2018; 2: 201–312 [Mironov BN. Rossiiskaya imperiya: ot traditsii k modernu. Saint Petersburg: Dmitrii Bulanin; 2018.2] (in Russian).
9. Миронов Б. Н. Российская империя: от традиции к модерну. Санкт Петербург: Дмитрий Буланин; 2018; 3: 53–112 [Mironov BN. Rossiiskaya imperiya: ot traditsii k modernu. Saint Petersburg: Dmitrii Bulanin; 2018.3] (in Russian).
10. Миронов Б. Н. Управление этническим многообразием Российской империи. Санкт Петербург. 2017 [Mironov BN. Upravlenie etnicheskim mnogoobraziem Rossiiskoi imperii. Saint Petersburg: Dmitrii Bulanin; 2017] (in Russian).
11. Реброва О. Ю. Статистический анализ медицинских данных. Применение прикладных программ STATISTICA. М.: Медиа-Сфера; 2006. 21 [Rebrova OYu. Statisticheskii analiz meditsinskikh dannykh. Primenenie prikladnykh programm STATISTICA. Moscow: Media-Sfera; 2006] (in Russian).
12. Тарушкин В. Т., Тарушкин П. В., Тарушкина Л. Т. Интервальная и нечеткая линейная регрессия для ВВП России. Успехи современного естествознания. 2007; 5: 81 [Tarushkin VT, Tarushkin PV, Tarushkina LT. Interval'naya i nechetskaya lineinaya regressiya dlya VVP Rossii. Advances in Current Natural Sciences. 2007;(5):81] (in Russian).

Поступила в редакцию 24.12.2018
Принята в печать 27.02.2019

Received 24.12.2018
Accepted 27.02.2019

Для цитирования: Тимонин А.Н., Бурляева Е.А., Никитин Н.С., Клочкова С.В., Никитюк Д.Б. Применение метода регрессионного моделирования в оценке изменений и прогнозировании физического статуса населения России. Журнал анатомии и гистопатологии. 2019; 8(1): 9–13. doi: 10.18499/2225-7357-2019-8-1-9-13.

For citation: Timonin AN., Burlyayeva EA., Nikitin NS., Klochkova SV., Nikityuk DB. Application of the regression modeling method in assessing changes and predicting the physical status of the Russian population. Journal of Anatomy and Histopathology. 2019; 8(1): 9–13. doi: 10.18499/2225-7357-2019-8-1-9-13.