

UDC 314

**Statistical Research of Expected Life Expectancy  
in the North Western Federal District: Panel analysis**<sup>1</sup> Marina M. Kruchek<sup>2</sup> Ekaterina V. Molchanova

<sup>1</sup> Petrozavodsk state university, Russia  
Russia, Petrozavodsk, Lenina St., 33  
PhD (Physical and Mathematical), Associate Professor  
E-mail: kruchek@mail.ru

<sup>2</sup> Institute of economics of the KarRC RAS, Russia  
Petrozavodsk, A. Nevskogo St., 50  
PhD (Technical)

**Abstract.** In article possibility of application of statistical models of regression type for research of medico-demographic processes and an assessment of expected life expectancy depending on socio-economic factors for regions of the North Western federal district is considered.

**Keywords:** Medico-demographic situation; the health; expected life expectancy; regression models.

**Введение.** Современные демографические тенденции в России носят кризисный характер и вызваны разнообразными факторами (социальными, экономическими, историческими и др.). В докладе «Глобальные тенденции развития человечества до 2015 г.» говорится, что население России не только сокращается, но и становится все менее и менее здоровым, а значит, теряет способность служить движущей силой экономического возрождения страны.

Исследование социально-экономических, экологических, природно-климатических и психологических факторов, влияющих на медико-демографические тенденции и состояние здоровья населения регионов России, является актуальной задачей для современного научного сообщества. Сложность проблемы заключается в том, что она носит междисциплинарный характер и для ее разрешения необходимы усилия специалистов из разных областей знания. Особая роль в решении этой проблемы отводится математике, т.к. множество факторов, влияющих на здоровье, может оказаться столь большим, что выбор главных компонент традиционными методами оказывается невозможным. Для оценки состояния здоровья необходимо использовать специальные методы математической статистики и моделирования.

**Материалы и методы.** Целью данной работы является выявление факторов, влияющих на медико-демографические процессы в регионах Северо-Западного федерального округа (СЗФО). Исследование проводилось с помощью экономико-математических методов и моделей (регрессионный анализ панельных данных). Информационной базой служили данные Росстата («Регионы России», «Здравоохранение в России») [3, 5]. Информация по 80 субъектам Российской Федерации (из рассмотрения были исключены автономные округа) за 2005-2010 гг. была сформирована в виде специальной системы, состоящей из следующих блоков: социально-демографический состав населения, экология и природно-климатические условия, социально-экономическое развитие, доступность медицинских услуг, медико-демографические показатели, социальный стресс.

Важнейшей характеристикой человеческого потенциала и состояния здоровья можно считать показатель ожидаемой продолжительности предстоящей жизни (ОПЖ). Ожидаемая продолжительность предстоящей жизни (ОПЖ – показатель, рассчитываемый на основе таблиц дожития) в агрегированном виде характеризует, с одной стороны, уровень смертности, с другой – дает возможность адекватного сопоставления уровней смертности между любыми территориями. Ожидаемая продолжительность предстоящей жизни

рассматривается в качестве одного из базовых индикаторов социально-экономического благополучия населения [2].

В нашем исследовании мы установили взаимосвязь между ОППЖ и различными факторами из разработанной системы показателей. На первом этапе мы разбили признаки на группы таким образом, чтобы внутри группы факторов корреляция была достаточно большая, а между – маленькая. Одним из методов, который позволяет провести такую операцию, является метод корреляционных плеяд [1]. Построение корреляционных плеяд позволило выделить две основные группы признаков: социально-экономические показатели и факторы социального стресса. Остальные индикаторы представляли одиночные группы признаков. Следует отметить, что полученные плеяды не дублируют блоки разработанной информационной системы. Соответственно в регрессионную модель целесообразно включать факторы из разных плеяд или одиночные признаки, что и было выполнено в дальнейших исследованиях.

В дальнейших исследованиях мы использовали панельные данные (Panel Data), которые состоят из наблюдений одних и тех же экономических единиц или объектов (индивидуумы, домашние хозяйства, фирмы, регионы, страны и т.п.) и осуществляются в последовательные периоды времени. Для России основными источниками панельных данных служат разнообразные данные Росстата по регионам за определенные периоды времени, а также проект RLMS ([www.crc.unc.edu/rlms/project](http://www.crc.unc.edu/rlms/project)) – Российский мониторинг экономического состояния и здоровья населения. Выделяют три модели регрессии по панельным данным [4, 6]:

1. объединенная модель регрессии (**pooled model**),
2. модель регрессии с фиксированными эффектами (**fixed effect model**),
3. модель регрессии со случайными эффектами (**random effect model**).

Для исследования регрессионных моделей с панельными данными введем следующие обозначения. Пусть  $y_{it}$  – зависимая переменная для экономической единицы  $i$  в момент времени  $t$ ,  $x_{it}$  – набор объясняющих (независимых) переменных (вектор размерности  $k$ ) и  $\varepsilon_{it}$  – соответствующая ошибка,  $i=1, \dots, n$ ,  $t=1, \dots, T$ . Простейшая модель – это обычная линейная модель регрессии [4, 6]:

$$y_{it} = \alpha + x_{it}'\beta + \varepsilon_{it}, \quad (1)$$

где  $i$  – номер объекта (в данном случае региона),  $t$  – время,  $\alpha$  – свободный член,  $\beta$  – вектор коэффициентов,  $x_{it}'$  – вектор-строка матрицы объясняющих переменных, которая, по существу, не учитывает панельную структуру данных. При этом предполагается, что все ошибки  $\varepsilon_{it}$  некоррелированы между собой как по  $i$ , так и по  $t$ , и некоррелированы со всеми объясняющими переменными  $x_{it}$ . Эта модель носит название объединенной (сквозной) модели регрессии (**pooled model**).

Учесть при моделировании индивидуальные различия между экономическими единицами возможно, предположив, что случайная ошибка имеет вид:

$$\varepsilon_{it} = u_i + v_{it}, \quad (2)$$

где  $u_i$  – ненаблюдаемые индивидуальные эффекты, а  $v_{it}$  – остаточное возмущение. Большинство приложений панельных данных использует именно такую модель случайной ошибки.

Заметим, что  $u_i$  не зависят от времени и отвечают за характеристики объектов, которые не включены в регрессию непосредственно. Остаточное возмущение  $v_{it}$  меняется в зависимости от времени и объектов, и может рассматриваться как обыкновенная случайная составляющая в регрессии.

Модель с фиксированным эффектом (**fixed effect model**) описывается уравнением [4, 6]:

$$y_{it} = \alpha + x_{it}'\beta + u_i + v_{it}, \quad (3)$$

где переменные  $u_i$  – параметры, не зависящие от  $t$ , остаточные возмущения  $v_{it}$  – независимые одинаково распределенные случайные величины  $IID(0, \sigma_\varepsilon^2)$  и  $x_{it}$  – предполагаются независимыми от  $v_{it}$  для всех  $i$  и  $t$ . Эту модель предпочтительно использовать, если мы сосредотачиваемся на уникальном наборе  $N$  регионов, как в нашем случае.

Модель со случайными эффектами (random effects model) предполагает другой подход к интерпретации переменных. В модели предполагается, что индивидуальные отличия носят случайный характер. Ее можно рассматривать как компромисс между объединенной регрессией, налагающей сильное ограничение гомогенности на все коэффициенты уравнения регрессии для любых  $i$  и  $t$ , и регрессией с фиксированными эффектами, которая позволяет для каждого объекта выборки ввести свою константу и, таким образом, учесть существующую в реальности, но ненаблюдаемую гетерогенность. Модель со случайными эффектами применяется в том случае, если выбирают случайным образом  $N$  объектов из генеральной совокупности элементов, например, в случае исследований панелей домашних хозяйств, индивидуумов или мелких фирм [4, 6].

**Результаты.** В нашем исследовании именно регрессионная модель с фиксированными эффектами при учете структуры панельных данных позволила получить значимый и обоснованный вариант моделирования, который можно использовать для оценки основного демографического индикатора – ожидаемой продолжительности предстоящей жизни в регионах СЗФО в зависимости от показателей социально-экономического развития, медицинского обслуживания и фактора социального стресса. При проведении расчетов использовался статистический пакет Stata SE (табл. 1).

Таблица 1

**Регрессия с фиксированным эффектом для ОППЖ, социально-экономических факторов и показателей социального стресса для СЗФО (F(5,45)=111,28; Prob>F=0,0000; R-sq: within=0,9252; between=0,7806; overall=0,5864; corr(u\_i, Xb)=-0,9642)**

ОППЖ (лет)	Coef.	Std. Err	t	P> t	[95% Conf. Interval]	
Логарифм ВРП на душу населения (руб)	3,248012	0,5843187	5,56	0,000	2,071134	4,424891
Удельный вес городского населения (%)	0,3056367	0,1375155	2,22	0,031	0,0286662	0,5826072
Численность населения денежными доходами ниже прожиточного минимума (%)	-0,090420	0,0536784	01,68	0,099	-0,198534	0,0176931
Прерывание беременности (аборты) - на 1000 женщин 15-49 лет	-0,064890	0,0215208	-3,02	0,004	-0,108235	-0,021545
Алкоголизм и алкогольные психозы (численность больных, состоящих на учете в ЛПУ на 100 тыс.чел.)	-0,006079	0,00171	-3,56	0,001	-0,009523	-0,002635
Константа	17,30962	12,20369	1,42	0,163	-7,269865	41,88911

Sigma_u	6,0707166
Sigma_e	0,55016469
Rho	0,99185384
F test that all u_i=0: F(9,45)=21,69 Prob>F=0,0000	

Таким образом, при построении регрессии с фиксированными эффектами по панельным данным для регионов СЗФО за 2005-2010 гг. коэффициент детерминации R-sq(within) составил 0,9252, что свидетельствует об адекватности построенной модели.

**Обсуждение.** В наших исследованиях построение регрессии с фиксированными эффектами по панельным данным позволило установить, что значение ОППЖ связано с социально-экономическими факторами (ВРП на душу населения) и ростом стрессов, обусловленных переходом к рыночной экономике. Падение продолжительности жизни по регионам СЗФО напрямую связано с показателями повышения стресса, которые можно количественно измерить: численностью населения с денежными доходами ниже прожиточного минимума (%), прерыванием беременности (абортами) – на 1000 женщин 15–49 лет, алкоголизмом и алкогольными психозами (численностью больных, состоящих на учете в ЛПУ на 100 тыс. чел.).

**Выводы.** Таким образом, построенная модель показала, что состояние здоровья населения связано в большей мере с социально-экономическими проблемами, с уровнем материального благосостояния, доходами семьи, с бедностью, которые влекут за собой менее сбалансированное и качественное питание, худшие жилищные условия, более низкую мотивацию на здоровый образ жизни. Меры по улучшению демографической ситуации должны носить комплексный характер, связанный с переориентацией целевых программ на решение задач демографической политики с учетом региональной специфики. Результаты исследования позволяют выявить направления для снижения социально-экономического бремени болезни и улучшения качества жизни населения.

#### Примечания:

1. Айвазян С.А., Мхитарян В.С. Прикладная статистика и основы эконометрики. М.: ЮНИТИ, 1998. 1005 с.
2. Атлас демографического развития России / под ред. акад. РАН Г.В. Осипова и проф. С.В. Рязанцева. М.: Экономическое образование, 2009. С. 220.
3. Здоровоохранение в России 2009: стат.сб. М.: Росстат, 2009. С. 365.
4. Магнус Я.Р., Катышев П.К., Пересецкий А.А. Эконометрика. Начальный курс. 6-е изд., переаб. и допол. М.: Дело, 2004. 576 с.
5. Регионы России. Социально-экономические показатели 2010: стат. сб. М.: Росстат, 2010. С. 996.
6. Badi H. Baltagi. *Econometric Analysis of Panel Data. Third Edition.* 2005.

УДК 314

### Статистическое исследование ожидаемой продолжительности жизни в Северо-Западном Федеральном округе: панельный анализ

<sup>1</sup> Марина Марленовна Кручек

<sup>2</sup> Екатерина Владимировна Молчанова

<sup>1</sup> Петрозаводский государственный университет, Россия  
Россия, г. Петрозаводск, ул. Ленина, 33  
Кандидат физико-математических наук, доцент  
E-mail: kruchek@mail.ru

<sup>2</sup> Институт экономики КарНЦ РАН, Россия  
г. Петрозаводск, пр. А. Невского, 50  
Кандидат технических наук

**Аннотация.** В статье рассматривается возможность применения статистических моделей регрессионного типа для исследования медико-демографических процессов и оценки ожидаемой продолжительности жизни в зависимости от социально-экономических факторов для регионов Северо-Западного федерального округа.

**Ключевые слова:** медико-демографическая ситуация; здоровье; ожидаемая продолжительность жизни; регрессионные модели.