

**Прогнозирование среднемесячной заработной платы работников
административно-территориальных образований: регрессионные модели**

Хубаев Г.Н., профессор Ростовского государственного экономического университета (РИНХ), докт экон наук, профессор

Коротин Д.В., программист, отдел Информационных технологий Ростовского государственного экономического университета (РИНХ)

Сидоренко Д.С., магистрант Ростовского государственного экономического университета (РИНХ) по программе «Прикладная информатика»

Аннотация. Выявлены статистически значимые корреляции между показателями уровня жизни населения и показателями, характеризующими уровень экономического развития административно-территориального образования. Построены регрессионные модели ($R^2_{\text{корр}} > 0,9$; $F_{\text{кр}} \gg 100$; функция отклика – среднемесячная номинальная начисленная заработная плата работников). В качестве исходной информации использован массив статистических данных, сформированный Росстатом РФ.

Ключевые слова: уровень жизни населения, коэффициенты корреляции, регрессионные модели, статистическая значимость.

**Forecasting of average monthly wages of employees of administrative-territorial
education: regression models**

Khubaev G.N., Professor of the Department of Information Systems and Applied Informatics, Doctor of Economics, Professor

Korotin D.V., programmer, information technology Department Rostov state economic University (RINH),

Sidorenko D.S., Master's student of Rostov state economic University (RINH), on the program «Applied Informatics», Rostov-on-Don, Rostov State Economic University (RINH), Russia

Annotation. Statistically significant correlations between indicators of the standard of living of the population and indicators characterizing the level of economic development of an administrative-territorial unit are revealed. Constructed regression models ($R^2_{skorr} > 0.9$; $F_{kr} \gg 100$; response function – the average monthly nominal accrued wages of employees). An array of statistical data generated by the Rosstat of the Russian Federation was used as the initial information.

Keywords: wages of employees, correlation coefficients, regression models, random errors.

Постановка задачи. К основным вызовам и угрозам экономической безопасности относятся, как указано в [1], «усиление дифференциации населения по уровню доходов, ..., усиление дифференциации регионов и муниципальных образований по уровню и темпам социально-экономического развития». Действительно, в [2, 3] показано, что разница в уровне жизни населения в разных административно-территориальных образованиях (АТО) весьма велика. В такой ситуации трудно объяснить гражданам, почему в их АТО уровень жизни населения существенно ниже, чем в других АТО: заметно выше стоимость потребительской корзины, ниже величина средней заработной платы и/или оплаты за выполнение одинаковой работы и т.д. Поэтому важно определить, насколько статистически значимы взаимосвязи показателей уровня жизни с экономическими показателями, *выявить, какие экономические показатели оказывают определяющее, статистически значимое влияние на уровень жизни населения и в какой степени.

Однако в настоящее время отсутствуют публикации, содержащие *информацию об оценках взаимосвязи между социально-экономическими показателями, характеризующими эффективность деятельности АТО, и о

корреляции между показателями уровня жизни населения, и *сведения о построенных регрессионных моделях для прогнозирования среднемесячной номинальной начисленной заработной платы работников в административно-территориальных образованиях.

В статье представлены оригинальные результаты исследований, связанных с *выявлением и оценкой* взаимосвязи между показателями уровня жизни населения АТО и построением статистически значимых регрессионных моделей для прогнозирования величины одного из показателей уровня жизни населения АТО – показателя *среднемесячной заработной платы*.

Расчеты выполнены на данных РОССТАТА за период с 2011 по 2015 годы, представленных на сайте и в Статистических сборниках Росстата РФ [4]. Используются сведения о значениях 135 социально-экономических показателей по 83-м субъектам РФ.

1. Оценка взаимосвязи социально-экономических показателей.

Деятельность субъектов РФ, как известно, характеризуется множеством (>100) социально-экономических показателей. В процессе содержательного анализа, из общей совокупности показателей нами выделено несколько десятков и проведена оценка наличия корреляционной связи между выделенными показателями и величиной *среднемесячной номинальной начисленной заработной платой работников*. Проверка *наличия корреляции* осуществлялась по Р.А. Фишеру на основании t -распределения с $(n-2)$ степенями свободы при $n=83$.

В таблице 1 в первом *столбце* указан исследуемый показатель, а в первой *строке* таблицы представлены *три* производных показателя: показатель *«Наличие основных фондов /Среднегодовая численность занятых», характеризующий техническую оснащенность работающих и, возможно, общий уровень их профессиональной подготовки, и показатели, позволяющие, как нам представляется, оценивать уровень производительности труда работающих: *«Объем товаров собственного производства, выполненных работ и

услуг/Среднегодовая численность занятых» и *«ВРП/Среднегодовая численность занятых».

Таблица 1

Статистически значимые корреляции между социально-экономическими показателями

	ВРП на душу населения	Наличие основных фондов /Среднегодовая численность занятых	Объем товаров собственного производства, выполненных работ и услуг /Среднегодовая численность занятых	ВРП /Среднегодовая численность занятых
Среднемесячная номинальная начисленная заработная плата работников	0,77	0,74	0,73	0,65

Одновременно при оценке коэффициентов корреляции между показателями уровня жизни населения обнаружено наличие статистически значимой корреляционной связи. В таблице 2 представлены значения коэффициентов корреляции между показателем «Среднемесячной номинальной начисленной заработной платой работников» и показателями уровня жизни населения.

Таблица 2

Статистически значимые корреляции между показателями уровня жизни населения АТО

	Среднедушевые денежные доходы (в месяц)	Стоимость фиксированного набора потребительских товаров и услуг	Среднемесячный размер субсидий на одну семью	ВРП на душу населения/Стоимость фиксированного набора потребительских товаров и услуг
Среднемесячная номинальная начисленная заработная плата работников	0,93	0,87	0,63	0,74

В результате выполненных расчетов оказалось, что выбранный нами показатель *среднемесячной номинальной начисленной заработной платы работников* не только статистически значимо взаимосвязан с показателями уровня жизни населения, но и весьма существенно *коррелирует* с показателями, характеризующими уровень экономического развития АТО.

На следующем этапе исследованы возможности построения функциональной зависимости между показателем «*Среднемесячной номинальной начисленной заработной платой работников*» (функция отклика) и статистически с ним связанными показателями (показатели со статистически значимыми коэффициентами корреляции использованы в качестве независимых переменных). Проведение экспериментов при поиске *лучшей* по статистическим критериям регрессионной модели осуществлялось с включением в состав входных переменных различных сочетаний факторов-показателей. При этом использовался универсальный алгоритм ранжирования объектов [5], позволяющий оперативно выявлять взаимосвязи между объектами, формировать подгруппы с примерно одинаковыми величинами оценочных показателей и анализировать причины образования таких подгрупп, расширять для системы верхнего уровня возможности оптимального выбора определенного объекта, предоставляя сведения о том, *какие объекты-АТО и по каким показателям* занимают лидирующее положение, а каждому из АТО обеспечить возможность *анализа резервов* улучшения деятельности, дать первоначальную оценку его конкурентным позициям, рассчитывать статистические характеристики (математическое ожидание, дисперсию, коэффициент вариации, медиану, асимметрию, эксцесс), таблицы и гистограммы распределения АТО в зависимости от численных значений показателей (см. также [6]).

Построение регрессионных моделей выполнялось нами при использовании исходной информации, представленной в *относительных* единицах (по отношению к математическому ожиданию значений показателя).

В окончательном варианте получены *статистически значимые* регрессионные модели, *хорошо описывающие* исходную информацию, с

составом факторов, включение которых в уравнение регрессии легко обосновать содержательно.

Модель без свободного члена имеет вид:

$$Y = b_1 * X_1 + b_2 * X_2 + b_3 * X_3,$$

где:

Y – Среднемесячная номинальная начисленная заработная плата работников; X_1 - Стоимость фиксированного набора потребительских товаров и услуг;

X_2 – Отношение: Объем товаров собственного производства, выполненных работ и услуг/Среднегодовая численность занятых; X_3 – Отношение ВРП на душу населения в 2015 году к ВРП на душу населения в 2011 году.

В результате расчетов получены следующие численные значения статистических показателей качества моделей:

1) для моделей с тремя факторами-независимыми переменными:

2015 г. Значения b -коэффициентов:

Модель 1. $b_1=1,62$; $b_2=0,03$; $b_3=-0,86$;

$b_1/\sigma_{b_1}=14,77$; $b_2/\sigma_{b_2}=7,87$; $b_3/\sigma_{b_3}=-9,4$;

$R^2=0,97$; $R^2_{\text{скорр}}=0,97$; $F_{\text{кр}}=1481,2$.

Для модели со свободным членом:

Модель 2. $b_0=0,25$; $b_1=1,6$; $b_2=0,03$; $b_3=-1,1$;

$b_0/\sigma_{b_0}=0,52$; $b_1/\sigma_{b_1}=13,7$; $b_2/\sigma_{b_2}=6,8$; $b_3/\sigma_{b_3}=-2,53$;

$R^2=0,9$; $R^2_{\text{скорр}}=0,89$; $F_{\text{кр}}=213,9$.

Если изменить состав независимых переменных: X_1 - ВРП на душу населения (в 2015 году); X_2 – Стоимость фиксированного набора потребительских товаров и услуг; X_3 – ВРП на душу населения/ Стоимость фиксированного набора потребительских товаров и услуг, получим:

Модель 3. $b_0=-1,43$; $b_1=-0,56$; $b_2=2,15$; $b_3=0,92$;

$b_0/\sigma_{b_0}=-9,4$; $b_1/\sigma_{b_1}=-4,56$; $b_2/\sigma_{b_2}=15,5$; $b_3/\sigma_{b_3}=5,6$;

$R^2=0,9$; $R^2_{\text{скорр}}=0,9$; $F_{\text{кр}}=254,1$.

2) для моделей с двумя факторами-независимыми переменными:

X_1 – Стоимость фиксированного набора потребительских товаров и услуг;

X_2 – Отношение: Объем отгруженных товаров собственного производства, выполненных работ и услуг собственными силами по видам деятельности/Среднегодовая численность занятых.

Модель 4. $b_1=0,64$; $b_2=0,047$;

$b_1/\sigma_{b_1}=13,12$; $b_2/\sigma_{b_2}=9,2$;

$R^2=0,96$; $R^2_{\text{скорр}}=0,95$; $F_{\text{кр}}=1046,7$.

Для модели со свободным членом:

Модель 5. $b_0=-0,94$; $b_1=1,64$; $b_2=0,04$;

$b_0/\sigma_{b_0}=-8,7$; $b_1/\sigma_{b_1}=13,7$; $b_2/\sigma_{b_2}=9,2$;

$R^2=0,88$; $R^2_{\text{скорр}}=0,88$; $F_{\text{кр}}=297,5$.

2. О прикладной полезности построенных моделей. Оценивая *полезность* представленных *статистически значимых* уравнений регрессии, с *ничтожно малыми ошибками* описывающих достаточно *достоверную* исходную информацию, обратим внимание на то, что, **во-первых**, ЛПР на всех уровнях управления **смогут**, получив *количественно обоснованный перечень определяющих факторов* и выполнив *оптимизационные* расчеты, **выбирать** *оптимальный, требующий минимальных затрат ресурсов* вариант вложения средств для *повышения* уровня жизни населения. **Во-вторых**, население АТО получит возможность реально оценивать *качество работы* системы верхнего уровня – уровня правительства субъекта РФ или муниципалитета. Ведь включенные в модели факторы-показатели не только весьма убедительно подтвердили свою «статистическую значимость», но и их включение в модель легко обосновывается содержательно.

Представляется также важным и то, что *аналогично*, определив на начальном этапе исходный перечень показателей, *предположительно* связанных с показателями, характеризующими уровень и качество жизни населения (например, путем использования экспертного [7] или количественного анализа), можно сформировать статистически значимый перечень определяющих

факторов, *управление* которыми обеспечит рост *уровня жизни* граждан России, как, впрочем, и любой другой страны мира.

3. О достоверности исходной информации. В процессе проведения исследований нами использован достаточно большой объем исходной статистической информации: количество субъектов РФ – 83, а показателей, характеризующих их деятельность – 135, т.е. всего более 10 тысяч числовых данных только за один год. Вполне вероятно, что среди такого обилия чисел могут оказаться выбросы и случайные ошибки, недостоверные, сомнительные, аномальные наблюдения. Ведь наличие их в массиве исходной информации может *негативно* повлиять на *результаты расчета*, на *выводы*, на *качество принимаемых решений*.

Однако анализ исходной информации по методике [8] не выявил наличие аномальных наблюдений в массиве данных Росстата РФ: при анализе корреляционных матриц *по выделенным факторам* за исследуемые 2011-2015 годы и остатков при проверке качества регрессионных моделей (см., например, рис. 1 – остатки распределены достаточно равномерно вдоль оси абсцисс, выбросов нет) сомнительных данных *не обнаружено*.



Рис. 1 – График остатков у регрессионной модели 1

ВЫВОДЫ. Насколько нам известно, в результате выполненных исследований *впервые*:

1. Построены *регрессионные модели* для прогнозирования значений показателя «Среднемесячная номинальная начисленная заработная плата

работников» в административно-территориальных образованиях **лучшего качества* (при оценке по стандартным критериям статистической значимости – $R^2_{\text{скорр}} > 0,9$; $F_{\text{кр}} \gg 100$); **с лучшими прогнозными свойствами*; **с использованием открытых официальных статистических данных и одновременно с проверкой на наличие аномальных наблюдений* в массиве исходной информации, содержащем десятки тысяч числовых данных.

2. Включены в состав независимых переменных построенных регрессионных моделей факторы-показатели, **оригинальные по сформированной структуре*; **содержательно обоснованные и статистически значимые* (в большинстве случаев у b -коэффициентов отношение $b_i/\sigma_{b_i} \gg 2$).

3. Обеспечена возможность, базируясь на сформированном в процессе исследований подмножестве значимых факторов, проводить оптимизационные расчеты с целью *минимизации затрат ресурсов* на оптимальный выбор и реализацию проектов, направленных на *повышение* значений показателя «Среднемесячная номинальная начисленная заработная плата работников» в административно-территориальных образованиях.

Статья подготовлена по результатам исследований, выполненных при поддержке Российского фонда фундаментальных исследований (РФФИ) – проект 18-010-00806/18 «Уровень жизни населения административно-территориальных образований: выявление, исследование, анализ и оценка значимости определяющих факторов (для последующей оптимизации в условиях ограниченных ресурсов)»

Библиографический список

1. Указ Президента РФ №208 от 13 мая 2017 года «О Стратегии экономической безопасности Российской Федерации на период до 2030 года». URL: <http://www.garant.ru>

2. Хубаев Г.Н. Уровень бедности населения субъектов Российской Федерации: выявление, исследование и оценка статистической значимости

определяющих факторов // РИСК: Ресурсы, информация, снабжение, конкуренция. – 2018. – № 3. – с. 72-75.

3. Хубаев Г.Н. Уровень жизни населения субъектов Российской Федерации: статистическое исследование // Статистика – язык цифровой цивилизации: сборник докладов Международной научно-практической конференции «II Открытый российский статистический конгресс» (г. Ростов-на-Дону, 4-6 декабря 2018 г.): в 2 т. – Т. 1. / Российская ассоциация статистиков; Федеральная служба государственной статистики РФ, РГЭУ (РИНХ), Ростовское региональное отделение ВЭО России. – Ростов н/Д, 2018. – с. 409-414.

4. Регионы России. Основные характеристики субъектов Российской Федерации. 2017: Статистический сборник. – М., 2017. URL: http://www.gks.ru/free_doc/new_site/region_stat/sep_region.html

5. Хубаев Г.Н. Ранжирование объектов по множеству количественных показателей: универсальный алгоритм // РИСК: Ресурсы, информация, снабжение, конкуренция. – 2018. – № 1. – С. 213-217.

6. URL: <http://uroven-zhizni.ru>

7. Khubaev G. Expert review: method of intuitively agreed choice // 5th International Conference «Economy modernization: new challenges and innovative practice» (November 12, 2017, Sheffield, UK). – P. 65-80.

8. Хубаев Г.Н. Способ выявления ошибок в больших массивах числовой информации // Вопросы статистики. – 2014. – №10. – С. 20-25.

References

1. Presidential decree No. 208 of 13 may 2017 "On the Strategy of economic security of the Russian Federation for the period up to 2030". URL: <http://www.garant.ru>.

2. Khubaev G. N. The level of poverty of the population of the Russian Federation: identification, research and evaluation of the statistical significance of the determining factors // RISK: Resources, information, supply, competition. – 2018. – № 3. – p. 72-75.

3. Khubaev G.N. The standard of living of the population of the subjects of the Russian Federation: statistical study // Statistics – language of digital civilization: collection of reports of the International scientific and practical conference «II Open Russian statistical Congress» (Rostov-on-don, 4-6 December 2018): in 2 tons – Vol. 1. / Russian Association of statisticians; Federal state statistics service of the Russian Federation, RSEU (RINH), Rostov regional branch of the Russian wind farm. – Rostov n/A, 2018. – p. 409-414.

4. Region of Russia. The main characteristics of the subjects of the Russian Federation. 2017: Statistical collection. – M., 2017. URL: http://www.gks.ru/free_doc/new_site/region_stat/sep_region.html

5. Khubaev G.N. Ranking of objects on a set of quantitative indicators: universal algorithm // RISK: Resources, information, supply, competition. – 2018. – № 1. – P. 213-217.

6. URL: <http://uroven-zhizni.ru>

7. Khubaev G. Expert review: method of intuitively agreed choice // 5th International Conference «Economy modernization: new challenges and innovative practice» (November 12, 2017, Sheffield, UK). – P. 65-80.

8. Khubaev G.N. A method of detecting errors in large arrays of numerical information // Questions of statistics. – 2014. – №10. – P 20-25.