

Оценка эконометрических показателей на основе регрессионного анализа (на примере потребления продуктов питания населением Архангельской области)

Целью данного исследования является выявление характерных черт и закономерностей в потреблении продуктов питания населением Архангельской области на основе статистической оценки таких эконометрических показателей, как скорость роста расходов на продукты питания с изменением доходов домохозяйств данного населения, а также эластичность расходов и эластичность доли расходов на питание по денежному доходу. Для расчета этих показателей была предложена нелинейная однофакторная регрессионная модель, описывающая взаимосвязь доходов и расходов на продукты питания населения Архангельской области. В работе значительное внимание уделено проверке выполнения условий достоверности и значимости данной регрессионной модели.

Для проведения статистического анализа были использованы микроданные выборочных обследований бюджетов домохозяйств Архангельской области за IV квартал 2012 года, проведенных Росстатом. В качестве программы для обработки данных был использован пакет статистического анализа SPSS, где методами корреляционного и регрессионного анализа, а также с помощью проверки распределений исследуемых величин на нормальность и анализа регрессионных остатков, осуществлялось решение поставленных задач. В ходе исследования была разработана дважды логарифмическая регрессионная модель и получено статистически значимое уравнение регрессии, адекватно описывающее взаимосвязь денежного дохода и расходов на продукты питания домохозяйств Архангельской области. Исходя из полученного уравнения регрессии, были рассчитаны эластичности расходов и доли расходов на продукты

питания по денежному доходу: значение эластичности первой величины показало, что при увеличении денежных доходов населения Архангельской области на 1% расходы на продукты питания возрастают в среднем на 0,5%, полученное отрицательное значение эластичности доли расходов на продукты питания свидетельствует о том, что при увеличении денежных доходов данного населения на 1% доля расходов на продукты питания убывает в среднем на 0,5%.

На основе проведенного регрессионного анализа, где, помимо прочего, особое внимание было уделено проверке соблюдения необходимых условий для его применения, и оценки соответствующих эконометрических показателей, можно сделать следующие выводы о потреблении продуктов питания населением Архангельской области: с увеличением денежного дохода расходы домохозяйств на продукты питания также увеличиваются, однако скорость роста таких расходов с увеличением дохода уменьшается; в силу малой по сравнению с единицей эластичности расходов по доходу продукты питания для населения Архангельской области являются товарами первой необходимости; наконец, потребление продуктов питания данным населением находится в согласии с экономическим законом Энгеля, определяющим зависимость доли расходов на питание в денежных доходах домохозяйств от их уровня.

Ключевые слова: потребление продуктов питания, регрессионный анализ, анализ данных в SPSS, эластичность расходов по доходу, закон Энгеля.

Sergey A. Kochkin

Northern (Arctic) Federal University named after M.V. Lomonosov
Arkhangelsk, Russia

Estimation of econometric indicators on the basis of the regression analysis (on the example of food consumption by the population of the Arkhangelsk region)

The purpose of this research is detection of characteristic features and regularities in food consumption by the population of the Arkhangelsk region on the basis of statistical estimation of such econometric indicators as the growth rate of food expenditures with change of the income of households of this population, and also income elasticity of expenditures and income elasticity of a share of food expenditures. For calculation of these indicators the nonlinear one-factorial regression model describing interrelation of the income and food expenditures of the population of the Arkhangelsk region was offered. In the paper the considerable attention is paid to check of accomplishment of conditions of reliability and importance of this regression model.

For carrying out the statistical analysis microdata of selective examinations of budgets of the households of the Arkhangelsk region for the IV quarter of 2012 conducted by Rosstat were used. For data

processing the packet of the statistical analysis of SPSS was used, and the solution of objectives was performed by methods of the correlation and regression analysis, and also by means of check of distributions of the researched quantities to a normality and the analysis of a regression residuals.

During the research double logarithmic regression model was developed and statistically significant equation of regression which adequately describes interrelation of an income and food expenditures of households of the Arkhangelsk region is received. Proceeding from the received regression equation, income elasticities of expenditures and of a share of food expenditures were calculated: value of elasticity of the first size showed that in case of increase in incomes of the population of the Arkhangelsk region by 1% food expenditures increase on average for 0,5%, the received negative value of elasticity of a share of food

expenditures demonstrates that in case of increase in incomes of this population by 1% the share of food expenditures decreases on average for 0,5%.

On the basis of the carried-out regression analysis where, in addition, special attention was paid to check of observance of necessary conditions for its application, and estimation of the corresponding econometric indicators, it is possible to draw the following conclusions on food consumption by the population of the Arkhangelsk region: with increase in income food expenditures of households also increase,

however the growth rate of such expenditures decreases; owing to income elasticity of expenditures, small in comparison with unit, food for the population of the Arkhangelsk region are essential goods; at last, food consumption by this population is in a consent with the economic Engel's law determining dependence of a share of food expenditures in volume of incomes of households on their level.

Keywords: food consumption, regression analysis, data analysis in SPSS, income elasticity of expenditures, Engel's law.

Введение

Питание является одним из важных объектов статистического наблюдения за потреблением населения в том или ином регионе, доля дохода, расходуемая на продукты питания, как известно, служит показателем уровня благосостояния рассматриваемой группы населения. Оценивая различные эконометрические показатели, отвечающие за потребление продуктов питания, можно также делать выводы о значимых особенностях потребительского поведения исследуемых домохозяйств, а проводя корреляционно-регрессионный анализ, находить взаимосвязи между потреблением продуктов питания и различными социально-экономическими явлениями или процессами. Одной из главных проблем подобного анализа является построение регрессионной модели, поскольку ее качество определяет достоверность и значимость результатов анализа прогнозов и тенденций развития тех или иных социально-экономических явлений или процессов, что также влияет на достоверность оценки исследуемых эконометрических показателей, получаемых на основе этой регрессионной модели. Не каждый исследователь уделяет пристальное внимание всем этапам подобного анализа и правильной интерпретации получаемых результатов. В настоящей работе подробно рассматриваются все шаги построения однофакторной нелинейной регрессионной модели на примере использования пакета программ статистического анализа SPSS. В качестве предмета анализа выбрано построение регрессионной модели, описывающей статистическую взаимосвязь расходов на продукты питания и денежного дохода населения Архангельской области.

На основе полученной модели далее проведены расчет и интерпретация таких эконометрических показателей, как скорость роста расходов на продукты питания с изменением денежного дохода, а также эластичности расходов и доли расходов на продукты питания по доходу. Знание указанных величин позволит выявить характерные особенности потребительского поведения домохозяйств данного населения в отношении продуктов питания, а также даст возможность проверить выполнение закона Энгеля для потребления продуктов питания населением Архангельской области.

Для проведения статистического анализа были использованы имеющиеся микроданные выборочных обследований бюджетов домохозяйств Архангельской области за IV квартал 2012 года, проведенных Росстатом [1]. Общее число различных домохозяйств в выборке – 833.

Всякий однофакторный регрессионный анализ, то есть в случаях с одной зависимой и одной независимой переменной, включает в себя решение следующих четырех задач [2]:

1) Проверка на нормальность распределений переменных и отсутствие грубых наблюдений.

2) Корреляционный анализ: проверка наличия и характера связи между переменными.

3) Регрессионная модель: проверка значимости модели и значимости коэффициентов регрессии.

4) Анализ регрессионных остатков.

Рассмотрим подробно применение регрессионного анализа для описания статистической взаимосвязи денежного дохода и расходов на продукты питания населением Архангельской области.

1. Проверка на нормальность и отсутствие грубых наблюдений

Денежные доходы домохозяйств Российской Федерации как в целом, так и разных ее субъектов не подчиняются нормальному распределению [3]. Однако нередко этим свойством обладает натуральный логарифм денежного дохода. В нашем исследовании проверка с помощью критерия Шапиро-Уилка на уровне значимости $\alpha = 0,05$ (далее в работе все нулевые гипотезы проверяются при этом уровне значимости) показала, что денежный доход населения Архангельской области имеет нормальное распределение на уровне статистической тенденции ($p = 0,098$, $n = 833$). Дополнительный анализ квантильной диаграммы для логарифмически преобразованного денежного дохода показал, что большинство его значений находится на прямой линии, что также говорит о близости фактического распределения данной величины нормальному.

Расходы на продукты питания населением Архангельской области тоже не подчиняются нормальному распределению, в отличие от натурального логарифма этой величины. Применение критерия Шапиро-Уилка привело к статистически значимому выводу о нормальном распределении логарифмированных расходов на продукты питания ($p = 0,310$, $n = 833$).

Анализ ящичковых диаграмм показал отсутствие грубых наблюдений как для логарифма денежного дохода, так и для логарифма расходов на продукты питания.

2. Корреляционный анализ

Построение скаттерграммы (рис. 1) для логарифмов денежного дохода и расходов на продукты пи-

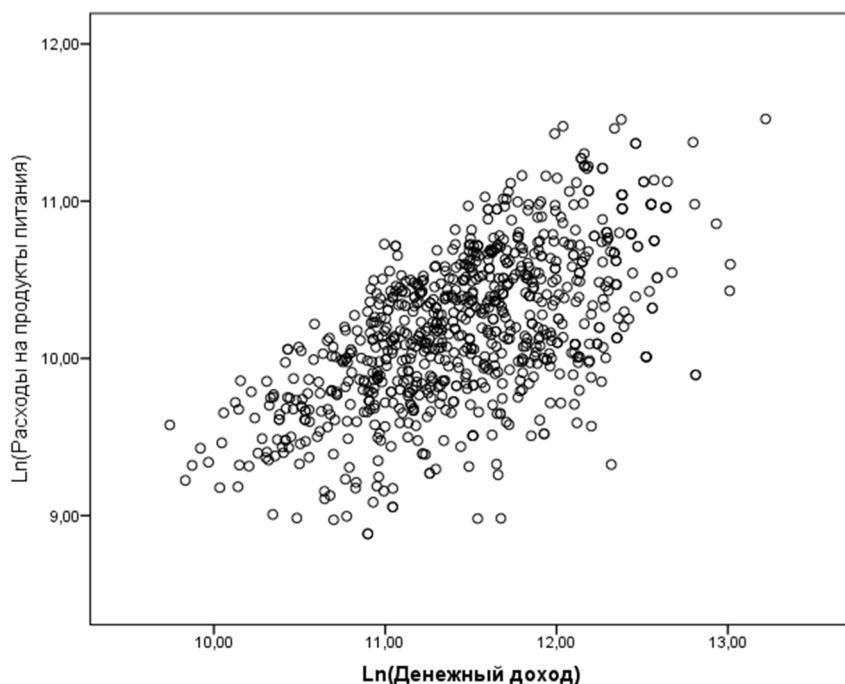


Рис. 1. Скаттерограмма для натуральных логарифмов денежного дохода и расходов на продукты питания 833 домохозяйств Архангельской области

тания показало наличие прямолинейной корреляционной связи между данными величинами. Поэтому для проведения корреляционного анализа был использован и рассчитан коэффициент корреляции Пирсона: $r = 0,611$ ($p < 0,05$, $n = 833$). На основании этого результата можно сделать вывод о том, что между натуральными логарифмами расходов на продукты питания и денежного дохода прямая и достаточно сильная связь [4].

3. Регрессионная модель

В силу нормального распределения исследуемых величин и результатов корреляционного анализа в качестве уравнения регрессии была выбрана дважды логарифмическая модель между денежным доходом и расходами на продукты питания, определяемая следующим уравнением:

$$\ln y_t = a + b \ln x, \quad (1)$$

где y_t – теоретическое (или предсказанное) значение расходов на продукты питания в рублях, x – денежный доход в рублях, a и b – коэффициенты регрессии. Очевидно, что в обозначениях логарифмически преобразованных величин денежного дохода и расходов на про-

дукты питания уравнение (1) имеет вид линейной зависимости.

Добавим, что регрессионная модель в целом имеет вид

$$\ln y = a + b \ln x + \varepsilon,$$

где ε – ошибка, характеризующая отклонения реального значения

результативной величины $\ln y$ от теоретического $\ln y_t$, найденного по уравнению регрессии (1).

Далее приведем результаты регрессионного анализа. На рис. 2 в таблице представлено значение коэффициента детерминации R^2 , определяющего долю общей вариации, которую способна объяснить данная регрессионная модель ($R^2 = 0,373$), а также значение критерия Дурбина-Уотсона, предназначенного для проверки условия независимости наблюдений. Допустимые значения для данного критерия – от 1 до 3 [2]. В нашем случае значение критерия Дурбина-Уотсона, равное 1,723, попадает в указанный интервал.

На рис. 3 приведены результаты критерия F для проверки значимости регрессионной модели: $F = 494$, $p < 0,05$. Таким образом, нулевую гипотезу об отсутствии взаимосвязи между исследуемыми величинами можно отвергнуть, и, значит, выбранная нами регрессионная модель статистически значима.

На следующем рис. 4 в таблице представлены значения коэффициентов регрессии с указанием их значимостей и 95%-ных доверительных интервалов. Критерий

Сводка для модели^b

Модель	R	R-квадрат	Скорректированный R-квадрат	Стандартная ошибка оценки	Дурбин-Уотсон
1	,611 ^a	,373	,372	,38638	1,723

a. Предикторы: (константа), Ln(Денежный доход)

b. Зависимая переменная: Ln(Расходы на продукты питания)

Рис. 2. Общие данные о регрессионной модели

ANOVA^a

Модель	Сумма квадратов	ст. св.	Средний квадрат	F	Знач.
1 Регрессия	73,714	1	73,714	493,765	3,170E-86 ^b
Остаток	124,059	831	,149		
Всего	197,773	832			

a. Зависимая переменная: Ln(Расходы на продукты питания)

b. Предикторы: (константа), Ln(Денежный доход)

Рис. 3. Результат применения критерия F

Коэффициенты^a

Модель	Нестандартизованные коэффициенты		Стандартизованные коэффициенты	t	Знач.	95,0% Доверительный интервал для B	
	B	Стандартная Ошибка	Бета			Нижняя граница	Верхняя граница
1 (Константа)	4,438	,259		17,110	1,878E-56	3,929	4,947
Ln(Денежный доход)	,502	,023	,611	22,221	3,170E-86	,457	,546

a. Зависимая переменная: Ln(Расходы на продукты питания)

Рис. 4. Таблица коэффициентов регрессии

Стьюдента, использованный для проверки гипотезы о равенстве коэффициентов регрессии нулю, показал, как видно из таблицы, статистическую значимость обоих коэффициентов ($p < 0,05$).

В результате имеем: $a = 4,438$, $b = 0,502$, поэтому уравнение регрессии (1) для взаимосвязи логарифмов денежного дохода и расходов на продукты питания можно представить в виде

$$\ln y_t = 4,438 + 0,5 \ln x. \quad (2)$$

4. Анализ регрессионных остатков

Последней и важной задачей регрессионного анализа является анализ остатков регрессии, который в свою очередь состоит из следующих четырех этапов [2]: 1) проверка на нормальность распределения остатков и отсутствие грубых наблюдений; 2) проверка равенства средней величины остатков нулю; 3) проверка независимости остатков; 4) проверка остатков на гомоскедастичность.

Следует особо подчеркнуть, что нарушение соответствующих условий в каждом из этапов анализа остатков может существенно понизить достоверность полученной регрессионной модели, и, как следствие, прогнозирование результирующей величины может привести к некорректным результатам, не говоря уже о достоверности значений эконометрических показателей, оцениваемых по данной модели.

Рассмотрим результаты, полученные на каждом из этапов анализа остатков для нашей модели.

1) Анализ квантильной диаграммы для остатков показал, что большинство его значений находится на прямой линии, что говорит о распределении остатков, близком к нормальному [2].

Анализ ящичковых диаграмм показал отсутствие грубых наблюдений в распределении остатков.

2) Применение одновыборочного t -критерия Стьюдента показало, что среднее арифметическое всех остатков статистически значимо равно нулю ($t = -2,95 \cdot 10^{14}$, $df = 832$, $p = 1$).

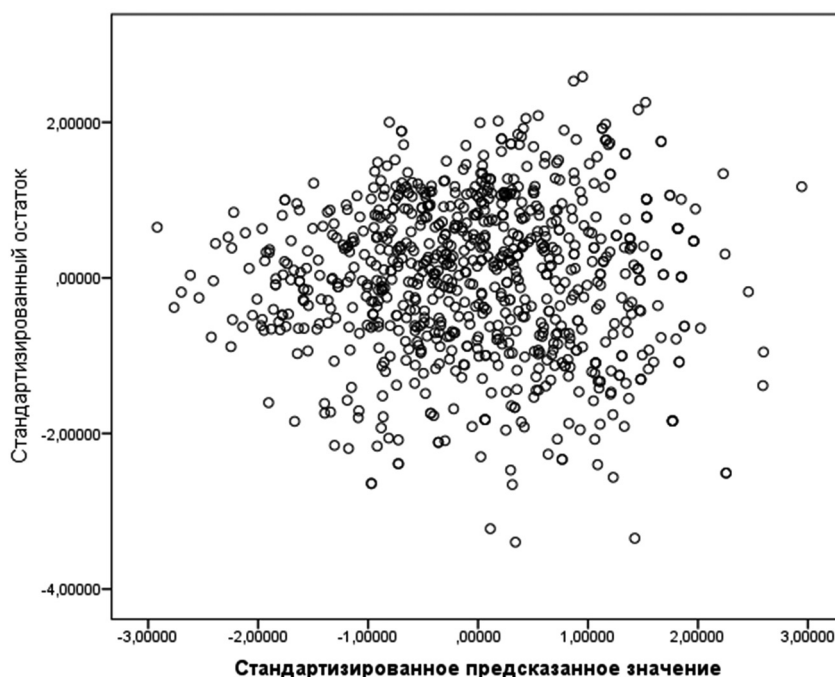


Рис. 5. Зависимость стандартизированных остатков регрессии от стандартизированных предсказанных значений расходов на продукты питания

3) Независимость остатков, как и независимость наблюдений, проверяется по значению критерия Дурбина-Уотсона, который в нашей модели равен 1,723. Как уже было сказано ранее, получилось число, попадающее в интервал допустимых для этого критерия значений. Значит, регрессионные остатки независимы.

4) Анализ построенного графика (рис. 5) зависимости стандар-

тизированных остатков от стандартизированных предсказанных значений показал, что разброс остатков приблизительно одинаков для всех значений стандартизированной предсказанной величины, что означает, что условие гомоскедастичности соблюдается.

Дополнительное вычисление коэффициента корреляции Пирсона для связи между предсказанными стандартизированными

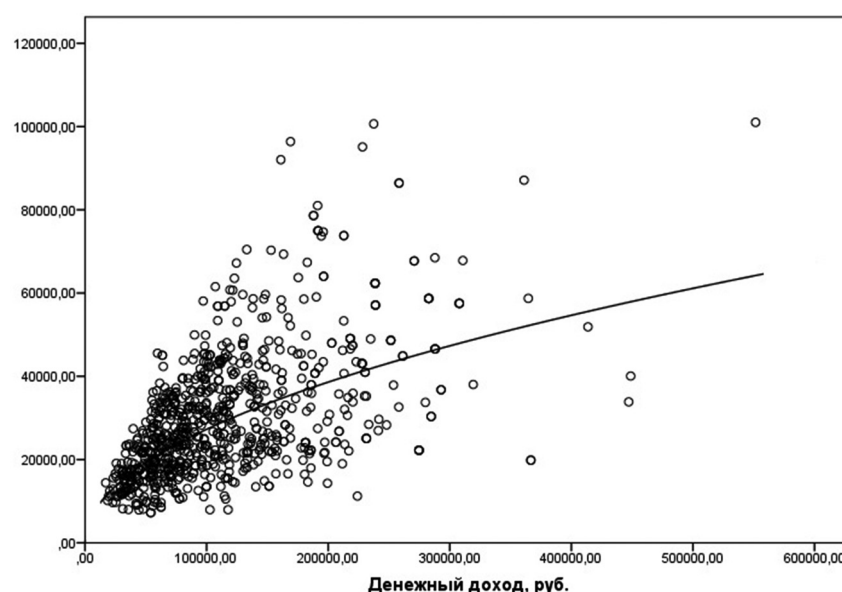


Рис. 6. Зависимость предсказанного значения расходов на продукты питания от денежного дохода (сплошная линия – график регрессионной кривой, кружки – данные 833 домохозяйств Архангельской области)

значениями расходов на продукты питания и стандартизированными остатками ($r = 1,1 \cdot 10^{15}$, $n = 833$, $p = 1$) также показало отсутствие тренда между ними.

Таким образом, проверка выполнения условий анализа остатков показала, что наша регрессионная модель достоверна и пригодна для прогнозирования данных.

Результаты

Итак, статистически значимое уравнение регрессии, адекватно описывающее статистическую взаимосвязь логарифмически преобразованных величин денежного дохода и расходов на продукты питания домохозяйств Архангельской области по выборочным данным IV квартала 2012 года, имеет вид (2). Потенцируя уравнение (2), можно получить искомое уравнение регрессии для исходных величин денежного дохода и расходов на продукты питания в следующем виде

$$y = 84,61x^{0,5}. \quad (3)$$

Отметим, что обе формулы (2) и (3) позволяют прогнозировать среднее значение расходов на продукты питания только для того диапазона денежного дохода, на основании которого было построено уравнение регрессии. Данный диапазон денежного дохода можно обнаружить на графике зависимости исследуемых величин, построенной по уравнению (3) и представленной на рис. 6.

Как видно из графика, с увеличением денежного дохода домохозяйств Архангельской области в указанном диапазоне расходы на

продукты питания также увеличиваются, однако скорость роста таких расходов, рассчитываемая по формуле

$$v = \frac{dy}{dx} = 42,31x^{-0,5},$$

убывает с увеличением дохода.

Далее, в полученном уравнении регрессии в виде степенной функции (3) степень независимой переменной имеет четкое экономическое истолкование – она является коэффициентом эластичности E расходов на продукты питания по денежному доходу. Действительно, по определению [5]

$$E = \frac{dy}{dx} \frac{x}{y} = 0,5.$$

В нашем исследовании это означает, что при увеличении денежных доходов населения Архангельской области на 1% расходы на продукты питания возрастают в среднем на 0,5%. Полученный результат ($0 < E < 1$) также подтверждает то, что продукты питания для населения Архангельской области являются товарами первой необходимости [5].

Наконец, введем долю δ расходов на продукты питания по доходу в виде

$$\delta = \frac{y}{x}$$

и рассчитаем эластичность E_δ данной величины по денежному доходу. Согласно определению, получим

$$E_\delta = \frac{d\delta}{dx} \frac{x}{\delta} = \frac{dy}{dx} \frac{x}{y} - 1 = E - 1.$$

Тогда в нашем случае $E_\delta = -0,5 < 0$, что означает, что при увеличении денежных доходов

населения Архангельской области на 1% доля расходов на продукты питания убывает в среднем на 0,5%.

Такое поведение расходов на продукты питания находится в полном согласии с известным в теории потребления законом Энгеля [5], утверждающем о том, что с увеличением доходов расходы на продукты питания хоть и растут, но доля таких расходов уменьшается. Ранее [6] нами была проведена более подробная проверка выполнения закона Энгеля с помощью статистического анализа потребления продуктов питания населением всей Российской Федерации. Теперь, как сказано выше, подтвердили справедливость закона Энгеля для потребительского поведения населения Архангельской области.

В заключение подчеркнем, что выбранная дважды логарифмическая регрессионная модель между денежным доходом и расходами на продукты питания (или, после преобразования, степенная зависимость между этими величинами) оказалась статистически значимой и адекватной для прогнозирования результатов и расчета эконометрических показателей в первую очередь благодаря тому, что переменные – в нашем случае логарифмы исследуемых величин – имели нормальное распределение, что является необходимым условием для проведения регрессионного анализа [2]. Дополнительный анализ других возможных регрессионных моделей для исследуемых данных показал, что для них условие о нормальном распределении переменных не выполняется.

Литература

1. Официальный сайт Федеральной службы государственной статистики [Электронный ресурс]. – Режим доступа: www.gks.ru (дата обращения: 23.06.2016).
2. Гржибовский А.М. Однофакторный регрессионный анализ // Экология человека. – 2010. – №10. – С. 55–64.
3. Суринов А.Е. Доходы населения. Опыт количественных измерений. – М., 2000. – 432 с.
4. Вильчинская О.В., Тарханова Л.А. Корреляционно-регрессионный анализ в оценке взаимосвязи показателей социально-экономического развития му-

References

1. Official site of Federal state statistics service [Electronic resource]. – Access mode: www.gks.ru (reference date: 23.06.2016).
2. Grzhibovsky A.M. One-factorial regression analysis // Human ecology. – 2010. – №10. – С. 55–64.
3. Surinov A.E. Income of the population. Experience of quantitative measurements. – M., 2000. – 432 с.
4. Vilchinskaya O. V., Tarkhanova L.A. The correlation and regression analysis in an assessment of interrelation of indicators of social and economic development of

ниципальных образований // Проблемы региональной экономики. – 2010. – Т. 8, № 3, часть 2. – С. 148–159.

5. Попов А.И. Экономическая теория. – СПб, 2006. – 900 с.

6. Кочкин С.А. Проверка закона Энгеля с помощью статистического анализа потребления продуктов питания населением России // Вестник современной науки. – 2016. – № 5, Ч. 1. – С. 66–70.

municipalities // Problems of regional economy. – 2010. – Т. 8, № 3, часть 2. – С. 148–159.

5. Popov A. I. Economic theory. – SPb, 2006. – 900 с.

6. Kochkin S.A. Verification of Engel's law by means of the statistical analysis of food consumption by the population of Russia // Vestnik sovremennoy nauki. – 2016. – № 5, P.1. – С. 66–70.

Сведения об авторе

Сергей Алексеевич Кочкин

к.ф.-м.н., доцент,

Северный (Арктический) федеральный университет

им. М.В. Ломоносова

Тел.: (8182) 21-61-00

Эл. почта: s.kochkin@narfu.ru

Архангельск, Россия

Information about the author

Sergey A. Kochkin

Candidate of Physico-mathematical Sciences, Associate Professor,

Northern (Arctic) Federal University

named after M.V. Lomonosov

Tel.: (8182) 21-61-00

E-mail: s.kochkin@narfu.ru

Arkhangelsk, Russia