

Для цитирования: Ильясов Б. Г., Макарова Е. А., Закиева Е. Ш., Гиздатуллина Э. С. Оценка данных о доходах населения в региональном разрезе методом главных компонент // Экономика региона. — 2019 — Т. 15, вып. 2 — С. 601-617

doi 10.17059/2019-2-22

УДК 681.3.06

**Б. Г. Ильясов, Е. А. Макарова, Е. Ш. Закиева, Э. С. Гиздатуллина**

Уфимский государственный авиационный технический университет  
(Уфа, Российская Федерация; e-mail: zakievae@mail.ru)

## ОЦЕНКА ДАННЫХ О ДОХОДАХ НАСЕЛЕНИЯ В РЕГИОНАЛЬНОМ РАЗРЕЗЕ МЕТОДОМ ГЛАВНЫХ КОМПОНЕНТ<sup>1</sup>

*В статье рассматривается решение задачи анализа статистических данных о доходах домохозяйств и их основных составляющих в абсолютных и относительных единицах с учетом ряда дополнительных показателей, в том числе социальных трансфертов, с использованием метода главных компонент. Цель анализа состоит в выявлении закономерности типа «кластеризация», что предполагает, во-первых, выявление кластеров регионов РФ, различающихся по структуре доходов населения с учетом объемов субсидий и субвенций, во-вторых, формирование обобщенных характеристик выявленных кластеров и представление их в виде правил кластеризации. Показано, что кластерная структура сектора домохозяйств, рассмотренная на региональном уровне, является достаточно поляризованной. Выявлены малочисленные кластеры регионов, характеризующиеся высоким уровнем денежных доходов домохозяйств при достаточно высокой численности населения (например, г. Москва, Ханты-Мансийский автономный округ). Сформированы достаточно населенные кластеры регионов, характеризующиеся значительным объемом неденежных доходов в виде натуральных поступлений продуктов питания при низком и среднем уровне денежных доходов и малой положительной динамике численности населения (например, Брянская, Курская области). Показано, что регионы, характеризующиеся достаточно низкими денежными доходами, имеют в структуре доходов высокую долю натуральных поступлений в виде продуктов питания (например, Республики Дагестан и Ингушетия). Показано также, что регионы, характеризующиеся высокими денежными доходами, имеют в структуре доходов малую долю привлеченных средств и израсходованных сбережений (Ямало-Ненецкий автономный округ и другие). Построенные кластеры регионов, их количество и состав, а также их обобщенные характеристики, представленные в виде правил кластеризации, используются для определения структурных и параметрических характеристик при разработке динамической модели сектора домохозяйств и модуля интеллектуального управления в составе системы имитационного динамического моделирования и интеллектуального управления (СИДМИУ) процессами формирования доходов населения. Применение СИДМИУ предполагает проведение сценарных исследований различных вариантов принятия решений при управлении доходами населения на региональном уровне с учетом дифференциация населения регионов по уровню доходов.*

**Ключевые слова:** доходы населения, метод главных компонент, выборка, кластеризация, весовой коэффициент, коэффициент информативности, интегральный признак, диаграмма рассеяния, кластеры регионов, имитационная динамическая модель

### Введение

Одними из ключевых показателей в рыночной экономике являются показатели, отража-

ющие уровень доходов населения. К ним относятся номинальные и реальные доходы населения, уровень заработной платы наемных работников, размер пенсий и стипендий. Косвенной характеристикой является размер сбережений населения [1]. Важность показа-

<sup>1</sup> © Ильясов Б. Г., Макарова Е. А., Закиева Е. Ш., Гиздатуллина Э. С. Текст. 2019.

теля уровня доходов объясняется его взаимосвязью с различными сферами жизни государства. С одной стороны, размер дохода позволяет оценить возможности человека или семьи, определить степень удовлетворения потребностей человека. С другой стороны, доход — это источник налоговых поступлений в бюджеты различных уровней.

Переход России к рыночной экономике по-разному сказался на российских регионах, в связи с чем одни регионы поднялись в экономическом плане, у других, наоборот, наблюдается кризис в экономике [2, 3]. В такой ситуации неизбежна экономическая дифференциация населения регионов по уровню доходов. Наблюдаются рост регионального неравенства, усиление межрегиональной дифференциации по прожиточному минимуму, среднедушевым денежным доходам, покупательной способности доходов населения и покупательной способности потребительских расходов [4–8]. Это создает ряд проблем для государства: сдерживание экономического развития, необходимость перераспределения части ресурсов на региональное выравнивание, социальная напряженность и сепаратистские настроения, опасность дезинтеграции национальной экономики [9–11]. Поэтому стратегически важным для России является проведение государственной региональной политики, направленной на сглаживание дифференциации в уровне социально-экономического развития регионов, оптимальную организацию государственной и региональной систем социальной защиты, способствующую снижению численности бедного населения, смягчению социальной дифференциации, созданию среднего класса как основы стабильности и процветания страны [2, 12–15].

В этой связи актуальны исследования динамики процессов управления доходами сектора домохозяйств в многосекторной макроэкономической системе (ММЭС), одной из целей которых является разработка программного обеспечения для системы имитационного динамического моделирования и интеллектуального управления (СИДМИУ) процессами формирования доходов сектора домохозяйств в составе ММЭС [5, 6].

Основными компонентами существующей СИДМИУ являются модуль имитационного моделирования динамики доходов и расходов населения в рамках воспроизводственного процесса ММЭС в целом, а также модуль интеллектуального управления макроэкономической системой при реализации сценариев

инвестиционно ориентированного и потребительски ориентированного роста экономики, рассматриваемых на федеральном уровне. Региональный аспект управления доходами и расходами секторов экономики ранее не рассматривался. Поэтому для целей анализа, моделирования и управления доходами населения на региональном уровне с учетом дифференциации населения регионов по уровню доходов, необходимо расширение существующей динамической модели, которое должно быть выполнено путем декомпозиции модели сектора домохозяйств по региональному принципу. Поскольку моделирование динамики доходов населения всех 85 регионов России в отдельности не представляется возможным в силу сложности и необозримости такой модели, то предложено выполнять процесс декомпозиции сектора домохозяйств на кластеры домохозяйств регионов. Кластеры формируются путем объединения регионов, сходных по доходам населения и их основным составляющим, определяющим структуру доходов.

В связи с этим необходимо решение задачи анализа статистических данных о доходах домохозяйств и их основных составляющих в абсолютных и относительных единицах с учетом ряда дополнительных показателей (например, таких как динамика численности населения, уровни заболеваемости и безработицы, социальные трансферты), а также выявление по результатам анализа закономерностей типа «кластеризация». Данная закономерность предполагает формирование кластеров регионов и их характеристик в виде правил кластеризации. Полученные результаты позволят определить структурные и параметрические характеристики расширенной динамической модели сектора домохозяйств. Структурные характеристики определяют искомое количество подсистем в расширенной модели сектора домохозяйств в виде количества кластеров регионов и их состава, а параметрические характеристики определяют диапазоны изменения анализируемых признаков в пределах каждого выделенного кластера в отдельности. Параметрические характеристики могут быть представлены как в количественном выражении (для целей имитационного моделирования), так и в качественном выражении — например, с помощью термножеств «высокий», «средний», «низкий» (для целей формирования правил кластеризации и принятия решений нечеткой базы знаний в модуле интеллектуального управления СИДМИУ).

Интерес представляет также анализ данных о принимаемых решениях по выделению трансфертов населению регионов и их сопоставление с обобщенными характеристиками доходов и расходов населения. Результатом такого анализа является формирование кластеров регионов и их характеристик, которые интегрируют в себе текущий уровень доходов населения регионов и принимаемые решения по предоставлению субсидий, субвенций и дотаций на выравнивание бюджетной обеспеченности. Эти результаты позволят настроить параметры текущего состояния модуля интеллектуального управления СИДМИУ на региональном уровне и сформировать дополнительные корректирующие правила принятия решений, направленные на снижение дифференциация населения регионов по уровню доходов.

### Методы

Статья посвящена вопросам анализа статистических данных о доходах домохозяйств и их основных составляющих в абсолютных и относительных единицах с учетом ряда дополнительных показателей, например, таких, например, как динамика численности населения, уровень безработицы, социальные трансферты. Цель анализа состоит в выявлении закономерности типа «кластеризация», что предполагает, во-первых, выявление кластеров регионов РФ, различающихся по структуре доходов населения, численности населения и ее динамике, объемам субсидий и субвенций, направляемых на поддержку социально незащищенных слоев населения, во-вторых, формирование обобщенных характеристик выявленных кластеров и представление их в виде правил кластеризации. Полученные результаты в виде кластеров, а также их состава и обобщенных характеристик в виде правил позволят, во-первых, определить количество подсистем расширенной динамической модели сектора домохозяйств, предназначенной для исследования и управления доходами населения на региональном уровне с учетом дифференциация населения регионов по уровню доходов, во-вторых, сформировать базу нечетких правил модуля интеллектуального управления СИДМИУ процессами формирования доходов сектора домохозяйств на региональном уровне в составе ММЭС в целом.

Для исследования применяется метод главных компонент, который позволяет визуализировать многомерные данные в пространстве меньшей размерности [16–22], которое представлено в виде обобщенных характеристик, интегрирующих важные исходные признаки.

Источником информации для проведения исследований являются данные Федеральной службы государственной статистики и справочно-правовой системы «Консультант Плюс»<sup>1</sup>.

Процедура проведения исследования основывается на предложенных ранее модели качества жизни и методе интеллектуального анализа данных о доходах населения и состоит из ряда шагов [23–25].

На первом шаге, согласно сформулированной цели анализа, формируется множество признаков, включающих доходы населения, их составляющие, социальные трансферты, а также дополнительные признаки, характеризующие различные аспекты уровня жизни населения. Поскольку это множество включает более тридцати признаков, возникает необходимость деления этого множества на отдельные группы признаков для целей проведения компонентного анализа. Выделены следующие три группы признаков.

К первой группе отнесены признаки, характеризующие основные виды доходов населения, выраженные в абсолютных единицах, а также данные о численности населения. Это такие признаки, как среднедушевой денежный доход, среднемесячная номинальная начисленная заработная плата, средний размер назначенных пенсий, стоимость натуральных поступлений в виде продуктов питания, товаров и услуг, сумма привлеченных средств и израсходованных сбережений, прожиточный минимум, а также признаки: численность постоянного населения, общие коэффициенты рождаемости и смертности, уровень заболеваемости населения.

Во вторую группу входят признаки, характеризующие основные виды доходов населения, выраженные в относительных единицах, а также данные о численности населения и уровне безработицы.

К третьей группе отнесены признаки, характеризующие распределение межбюджетных трансфертов, субсидий и субвенций с уче-

<sup>1</sup> Регионы России. Социально-экономические показатели. 2015 [Электронный ресурс]. URL: [http://www.gks.ru/wps/wcm/connect/rosstat\\_main/rosstat/ru/statistics/publications/catalog/doc\\_1138623506156](http://www.gks.ru/wps/wcm/connect/rosstat_main/rosstat/ru/statistics/publications/catalog/doc_1138623506156) (дата обращения: февраль 2017); Доходы, расходы и потребление домашних домохозяйств в 2015 году (по итогам выборочного обследования бюджетов домохозяйств) [Электронный ресурс]. URL: [http://www.gks.ru/bgd/regl/b15\\_102/Main.htm](http://www.gks.ru/bgd/regl/b15_102/Main.htm) (дата обращения: февраль 2017); О федеральном бюджете на 2015 год и на плановый период 2016 и 2017 годов [Электронный ресурс]. Доступ из справочно-правовой системы «КонсультантПлюс». URL: [http://www.consultant.ru/document/cons\\_doc\\_LAW\\_171692](http://www.consultant.ru/document/cons_doc_LAW_171692) (дата обращения: февраль 2017).

том показателей уровня жизни населения. Это такие признаки, как дотации на выравнивание бюджетной обеспеченности субъектов РФ, единая субвенция бюджетам субъектов РФ, субвенции на выплаты инвалидам компенсаций страховых премий по договору ОСАГО, субвенции на социальные выплаты безработным гражданам, субвенции на выплату пособия при всех формах устройства лишенных родительского попечения детей в семью, межбюджетные трансферты в области лекарственного обеспечения и профилактики заболеваний, субсидии на предоставление жилых помещений детям-сиротам и детям, оставшимся без попечения родителей, а также такие признаки, как среднемесячная номинальная начисленная заработная плата, располагаемые ресурсы.

На втором шаге проводится расчет матрицы весовых коэффициентов главных компонент [16, 17].

На третьем шаге осуществляется визуализация диаграммы рассеяния объектов (регионов РФ) в пространстве главных компонент (ГК), которые являются координатными осями нового пространства и математически представлены в виде линейной комбинации исходных признаков [16, 17]. В работе рассматриваются двумерные или трехмерные диаграммы рассеяния.

На четвертом шаге определяется количество анализируемых главных компонент. Как правило, анализу подвергаются две или три главные компоненты, что является достаточным для проведения визуального анализа особенностей распределения объектов в пространстве интегральных признаков.

Следующие три шага (пятый, шестой, седьмой) посвящены формированию названия для первой главной компоненты (ГК1) как новой интегральной характеристики, присущей данным и являющейся латентной [16, 17]. Шаги включают расчет коэффициента информативности, корректировку состава признаков в зависимости от значения коэффициента информативности, предварительное выделение кластеров регионов на основе анализа различий по значениям первой главной компоненты (ГК1).

Восьмой и девятый шаги выполняются для анализа второй главной компоненты (ГК2) и аналогичны предыдущим. Совмещение результатов деления на кластеры по первой и второй главным компонентам выполняется на десятом шаге.

На последнем (одиннадцатом) шаге формулируются закономерности, то есть характер-

ные свойства, присущие выделенным кластерам регионов, что позволяет составить их профили, которые являются важной информацией для поддержки принятия решений при проведении социальной политики. Кластеры могут быть объединены в группы кластеров благополучных, умеренно благополучных и неблагополучных регионов.

### Результаты

Далее представлены результаты выполнения процедуры компонентного анализа данных о доходах населения.

Целью анализа первой выборки является выявление кластеров регионов РФ, различающихся по объемам основных видов доходов домашних хозяйств, представленных в абсолютных единицах, с учетом динамики численности населения в разрезе субъектов РФ.

Анализ проводится для выборки, включающей 85 субъектов РФ (по данным 2015 г.), по 11 признакам. По результатам анализа построены три главные компоненты. Весовые коэффициенты признаков для этих компонент приведены в таблице 1.

В первой главной компоненте наиболее значимыми являются признаки «среднедушевой денежный доход», «среднемесячная номинальная начисленная заработная плата», «средний размер назначенных пенсий», «прожиточный минимум» (данные признаки выделены в таблице жирным шрифтом). Таким образом, первая главная компонента интегрально характеризует величину денежных доходов населения с учетом прожиточного минимума.

Во второй главной компоненте большой по величине положительный коэффициент имеет признак «общие коэффициенты рождаемости», а признаки «общие коэффициенты смертности», «стоимость натуральных поступлений в виде товаров и услуг» и «сумма привлеченных средств и израсходованных сбережений» имеют большие по модулю отрицательные коэффициенты.

Таким образом, вторая главная компонента характеризует динамику численности населения, а также величину неденежных доходов (в виде товаров и услуг) и объем привлеченных средств и израсходованных сбережений населения. Высокие значения второй главной компоненты характеризуют высокую динамику численности населения (высокую рождаемость и низкую смертность), низкие величины неденежных доходов (в виде товаров и услуг) и суммы привлеченных средств и израсходованных сбережений.

Таблица 1

## Весовые коэффициенты признаков (первая выборка)

№ п/п	Признак	Весовые коэффициенты признака		
		компонента 1	компонента 2	компонента 3
$X_1$	Среднедушевой денежный доход (в месяц)	0,406101	0,00426384	0,232717
$X_2$	Среднемесячная номинальная начисленная заработная плата	0,441066	0,015903	0,101109
$X_3$	Средний размер назначенных пенсий	0,434259	-0,121878	-0,0795199
$X_4$	Прожиточный минимум (в среднем на душу населения)	0,429159	-0,0328116	-0,0445831
$X_5$	Общие коэффициенты рождаемости (число родившихся на 1000 чел. населения)	0,095573	0,566814	-0,374134
$X_6$	Уровень заболеваемости населения (зарегистрировано больных с диагнозом, установленным впервые в жизни) (число заболевших на 1000 чел. населения)	0,219938	-0,104101	-0,264374
$X_7$	Общие коэффициенты смертности (число умерших на 1000 чел. населения)	-0,210309	-0,553815	0,125638
$X_8$	Стоимость натуральных поступлений в виде продуктов питания (в среднем на члена домашнего хозяйства в месяц)	-0,185518	-0,371213	-0,343977
$X_9$	Стоимость натуральных поступлений в виде товаров и услуг (в среднем на члена домашнего хозяйства в месяц)	0,292598	-0,301003	-0,0292048
$X_{10}$	Сумма привлеченных средств и израсходованных сбережений (в среднем на члена домашнего хозяйства в месяц)	0,214067	-0,333148	-0,143736
$X_{11}$	Численность постоянного населения	0,0138617	0,0747165	0,749547

В третьей главной компоненте большой по величине положительный коэффициент имеет признак «численность постоянного населения», а признаки «уровень заболеваемости населения» и «стоимость натуральных поступлений в виде продуктов питания» имеют большие по модулю отрицательные значения коэффициентов. Таким образом, третья главная компонента характеризует численность постоянного населения с учетом уровня заболеваемости и неденежных доходов населения (натуральных поступлений в виде продуктов питания). Высокие значения третьей главной компоненты характеризуют высокую численность населения, низкий уровень заболеваемости, низкую величину неденежных доходов (в виде продуктов питания).

Построена диаграмма рассеяния в пространстве первой и второй главных компонент (рис. 1). Каждая ось разбита на интервалы, а плоскость — на прямоугольники. Так, на оси ГК1 выделено пять интервалов значений компоненты: низкий (Н), довольно низкий (ДН), средний (С), довольно высокий (ДВ), высокий (В), на оси ГК2 — интервалы значений: низкий (Н), средний (С), высокий (В). По результатам анализа данной выборки выделено восемь кластеров регионов.

В первый кластер вошли регионы, характеризующиеся высокими величинами денежных доходов, прожиточного минимума и средними величинами неденежных доходов, а также стабильной динамикой численности населения, но низкой численностью и высоким уровнем заболеваемости населения. Это такие регионы, как Ямало-Ненецкий автономный округ и Ненецкий автономный округ.

Во второй кластер вошел Чукотский автономный округ, характеризующийся высокими величинами денежных и неденежных доходов, прожиточного минимума, а также низкой динамикой численности населения, низкой численностью и высоким уровнем заболеваемости населения.

В третий кластер вошел Камчатский край, отличающийся более низкими величинами денежных доходов (расположен по оси ГК1 в области более низких значений).

В четвертый кластер вошли четыре региона: город-мегаполис Москва, Ханты-Мансийский автономный округ, Магаданская область и Республика Саха (Якутия). Эти регионы характеризуются довольно высокими величинами денежных доходов, прожиточного минимума и средними величинами неденежных до-

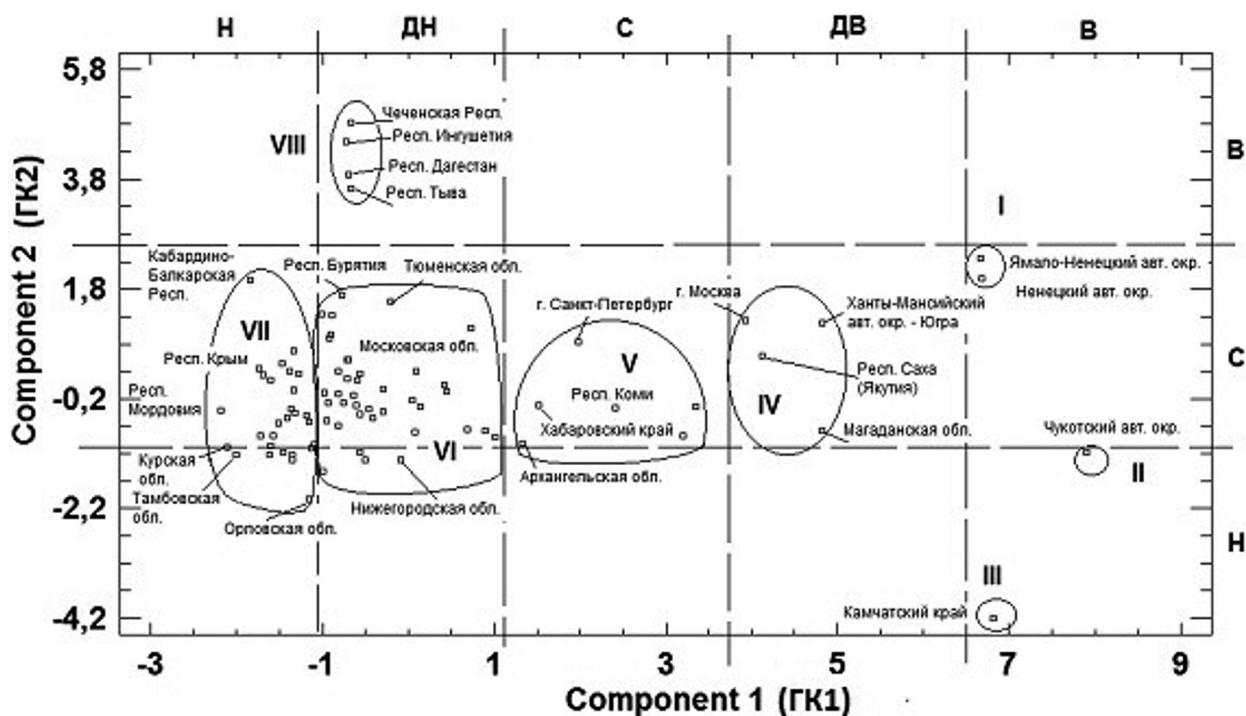


Рис. 1. Проекция множества регионов РФ на плоскость двух главных компонент (первая выборка)

ходов, а также стабильной динамикой численности населения.

В пятый кластер вошли шесть субъектов: Сахалинская, Мурманская, Архангельская области, Хабаровский край, Республика Коми и город-мегаполис Санкт-Петербург. Для них характерны средние величины денежных и неденежных доходов, прожиточного минимума. Динамика численности населения стабильная.

В шестой, самый многочисленный кластер вошли 39 регионов с довольно низким уровнем денежных доходов, прожиточного минимума и высоким или средним уровнем неденежных доходов от натуральных поступлений продуктов питания, товаров и услуг, со средней или низкой динамикой численности населения. Это такие регионы, как Республика Карелия, Самарская область, Республика Татарстан, Республика Башкортостан, Нижегородская область, Республика Бурятия, Тюменская область, Калужская область и другие.

Для регионов, вошедших в седьмой кластер, характерны низкий уровень денежных доходов населения, прожиточного минимума, высокий или средний уровень неденежных доходов, сумм привлеченных средств и израсходованных сбережений. Динамика численности населения средняя или низкая. Это такие регионы, как Брянская, Курская, Воронежская, Рязанская области, Республики Мордовия, Мари Эл, Крым, Кабардино-Балкарская Республика, г. Севастополь и другие.

Для регионов, вошедших в восьмой кластер, характерны довольно низкие уровни денежных доходов и прожиточного минимума, низкий уровень неденежных доходов. Динамика численности населения высокая. Это такие регионы, как Республики Тыва, Дагестан, Ингушетия и Чеченская Республика.

В целом, выделенные восемь кластеров различаются по уровню основных видов доходов населения и динамике численности населения. Малочисленные кластеры 1–5 включают регионы-лидеры с высоким, довольно высоким и средним уровнем доходов и стабильной динамикой численности населения. Это северные регионы и два мегаполиса. Кластеры 6 и 7 являются самыми многочисленными и характеризуются довольно низким и низким уровнем денежных доходов и средним или высоким уровнем неденежных доходов. Для регионов, вошедших в кластер 8 (это в основном Северо-Кавказские республики), характерны довольно низкий уровень денежных доходов и низкий уровень неденежных доходов (в виде товаров и услуг), а также высокая динамика численности населения.

Профили кластеров, составленные с учетом выделенных зон в пространстве главных компонент (рис. 1), являются основой для составления правил кластеризации базы знаний СИДМИУ (табл. 2).

Пример составления правил кластеризации с учетом состава признаков, определяю-

Таблица 2  
Профили кластеров в пространстве главных компонент

№ кластера	Значение ГК1	Значение ГК2
Кластер 1	Высокое	Среднее или высокое
Кластер 2	Высокое	Среднее или низкое
Кластер 3	Высокое	Низкое
Кластер 4	Довольно высокое	Среднее
Кластер 5	Среднее	Среднее
Кластер 6	Довольно низкое	Среднее или низкое
Кластер 7	Низкое	Среднее или низкое
Кластер 8	Довольно низкое	Высокое

щих первые две главные компоненты, для кластера 8:

Если среднедушевой денежный доход средний, среднемесячная номинальная начисленная заработная плата средняя, средний размер назначенных пенсий средний, прожиточный минимум средний, динамика численности населения положительная высокая, стоимость натуральных поступлений в виде товаров и услуг низкая, сумма привлеченных средств и израсходованных сбережений низкая, то кластер 8.

Для оставшихся двух выборок профили кластеров и правила кластеризации составляют аналогично.

Целью анализа второй выборки является выявление кластеров регионов РФ, различающихся по уровню доходов населения и их структурным особенностям. Признаки, характеризующие основные составляющие доходов населения, выражены в относительных единицах (в процентах). Дополнительно введен признак «уровень безработицы», косвенно харак-

теризующий положение домохозяйств регионов по уровню доходов.

При проведении компонентного анализа выборки, включающей 85 регионов РФ, (по данным 2015 г.), построены три главные компоненты. Анализируемых признаков — 10. Весовые коэффициенты признаков для этих компонент приведены в таблице 3.

Первая главная компонента характеризует в целом уровень жизни, который в значительной степени определяется уровнем доходов населения (заработная плата, размер назначенных пенсий) и величиной прожиточного минимума, а также долей неденежных доходов (натуральных поступлений в виде продуктов питания, товаров и услуг) в структуре располагаемых ресурсов населения. Высокие значения первой главной компоненты характеризуют высокий уровень денежных доходов, прожиточного минимума, низкий уровень доли неденежных доходов от натуральных поступлений в виде продуктов питания и высокий уровень доли неденежных доходов от натуральных поступлений в виде товаров и услуг. Выявлена важная структурная характеристика компоненты в виде обратно пропорционального соотношения долей натуральных поступлений продуктов питания и доли натуральными поступлениями товаров и услуг, представляющими неденежные доходы.

Вторая главная компонента имеет еще одну структурную характеристику, представленную в виде соотношения между денежным доходом и суммой привлеченных средств и израсходованных сбережений в структуре располагаемых ресурсов населения. Выявлено, что вы-

Таблица 3

Весовые коэффициенты признаков (вторая выборка)

№ п/п	Признак	Весовые коэффициенты признака		
		компонента 1	компонента 2	компонента 3
$X_1$	Прожиточный минимум (в среднем на душу населения)	0,488824	0,139211	0,0878907
$X_2$	Среднедушевой денежный доход (в месяц), %	-0,120376	0,631281	-0,232349
$X_3$	Среднемесячная номинальная начисленная заработная плата	0,498779	0,152372	-0,0691273
$X_4$	Общие коэффициенты рождаемости (число родившихся на 1000 чел. населения)	0,0374876	0,334753	0,489264
$X_5$	Стоимость натуральных поступлений в виде продуктов питания, %	-0,370348	-0,205221	0,100789
$X_6$	Стоимость натуральных поступлений в виде товаров и услуг, %	0,204835	-0,199283	0,0379958
$X_7$	Сумма привлеченных средств и израсходованных сбережений, %	0,206	-0,564834	0,207174
$X_8$	Уровень безработицы	-0,127277	0,207657	0,615515
$X_9$	Средний размер назначенных пенсий	0,508378	0,0529384	0,0337592
$X_{10}$	Численность постоянного населения	0,0132471	0,00615294	-0,509519

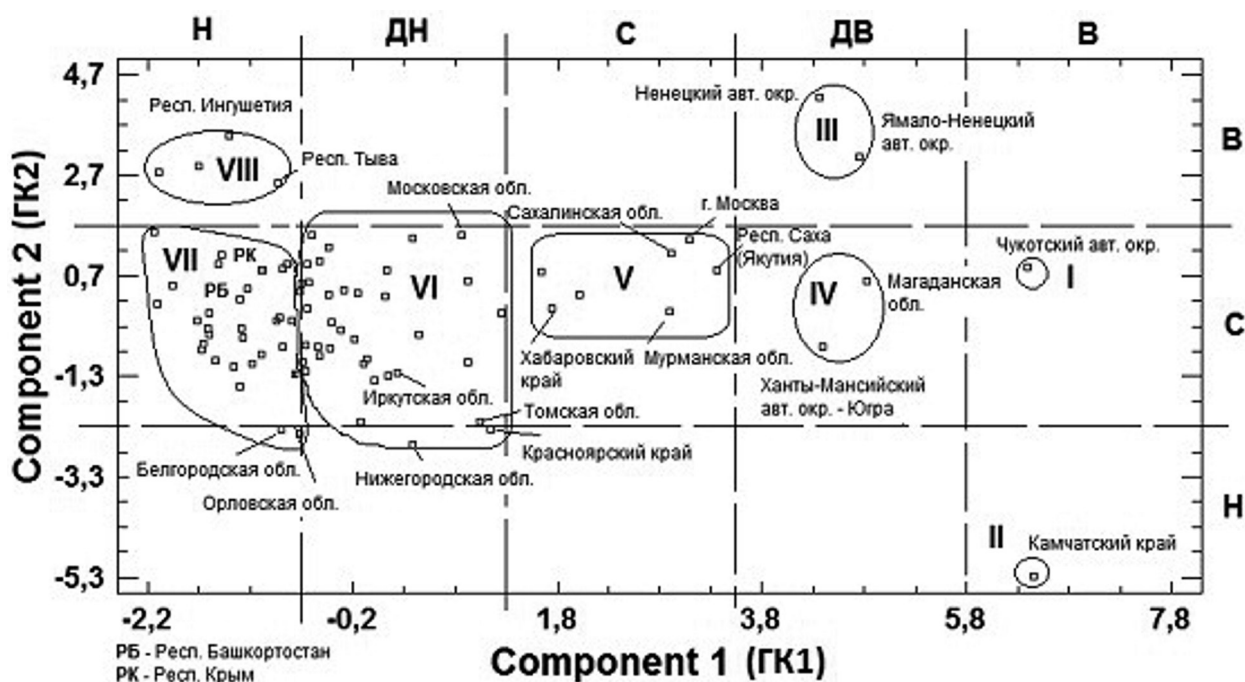


Рис. 2. Проекция множества регионов РФ на плоскость двух главных компонент (вторая выборка)

соким значениям второй компоненты соответствуют высокая доля денежных доходов и низкая доля доходов в виде суммы привлеченных средств и израсходованных сбережений.

Третья главная компонента характеризуется уровнем безработицы и численности населения с учетом влияния признака «коэффициент рождаемости». Уровень безработицы и численность населения связаны обратной пропорциональной зависимостью. Выявлено, что высоким значениям третьей главной компоненты соответствуют высокий уровень безработицы, высокий коэффициент рождаемости, но низкая численность населения. Справедливо и утверждение обратной зависимости.

Построена диаграмма рассеяния для рассматриваемого множества регионов РФ в пространстве выделенных двух главных компонент (рис. 2).

Исследуемая совокупность всех регионов разделилась на достаточно выраженные группы.

Первый кластер включает один субъект РФ — Чукотский автономный округ. Для данного региона характерен высокий уровень жизни (высокие уровни заработной платы и пенсии), высокий уровень прожиточного минимума, низкая доля неденежных доходов в виде натуральных поступлений (продуктов питания), что обусловлено климатическим и географическим расположением региона. Структура располагаемых ресурсов характеризуется средними по величине долями денежных доходов и сумм привлечен-

ных средств и израсходованных сбережений. Для данного региона характерны также средние уровни безработицы, рождаемости и численности населения.

Второй кластер включает в себя один субъект — Камчатский край, который также характеризуется высоким уровнем жизни, однако в структуре располагаемых ресурсов доля привлеченных средств и израсходованных сбережений преобладает над долей денежного дохода, а доля натуральных поступлений в виде продуктов питания также мала. Для данного региона характерны высокие уровни безработицы и рождаемости, низкая численность населения.

Третий кластер включает в себя 2 региона РФ: Ненецкий автономный округ и Ямало-Ненецкий автономный округ. Этот кластер регионов характеризуется довольно высоким уровнем жизни, но с измененной по сравнению со вторым кластером структурой располагаемых ресурсов: доля денежного дохода высокая, а доля привлеченных средств и израсходованных сбережений низкая, доля натуральных поступлений в виде продуктов питания также мала. Для данных регионов характерны также средний уровень безработицы, численности населения, коэффициента рождаемости.

Четвертый кластер включает в себя 2 субъекта РФ: Ханты-Мансийский автономный округ и Магаданскую область. Для данного кластера характерен довольно высокий уровень жизни. Структура располагаемых ресурсов характеризуется средними по величине до-

лями денежных доходов и сумм привлеченных средств и израсходованных сбережений, доля натуральных поступлений в виде продуктов питания немного выше, чем в рассмотренных выше кластерах.

*Пятый кластер* включает в себя 7 субъектов РФ, это регионы со средним уровнем жизни населения: г. Москва, г. Санкт-Петербург, Мурманская область, Республика Саха (Якутия), Сахалинская область, Республика Коми, Хабаровский край. Доля натуральных поступлений в виде продуктов питания возрастает. Структура располагаемых ресурсов по второй компоненте аналогична четвертому кластеру. Для данных регионов характерны также средний или низкий уровень безработицы, средняя численность населения (исключение составляет г. Москва, для которого характерна высокая численность населения).

*Шестой кластер* включает в себя 37 регионов РФ: Московская область, Красноярский и Пермский край, Нижегородская, Костромская, Кемеровская, Калининградская, Тюменская, Свердловская, Ивановская области и Республика Татарстан и др. Для регионов характерен довольно низкий уровень жизни, при этом продолжает возрастать доля натуральных поступлений в виде продуктов питания. По компоненте ГК2 структура располагаемых ресурсов не изменена. Данные регионы характеризуются также средним уровнем безработицы, численности населения.

*Седьмой кластер* включает 31 регион РФ: Республики Башкортостан, Мари Эл, Адыгея, Кабардино-Балкарская, Алтай и Брянскую, Воронежскую, Белгородскую, Орловскую области, и др. Это регионы с низким уровнем жизни. Уровень безработицы средний.

*Восьмой кластер* включает 4 субъекта РФ: республики Тыва, Ингушетия, Дагестан и Чеченскую Республику, которые характеризуются очень низким уровнем жизни, а также высокой долей натуральных поступлений в виде продуктов питания. В структуре располагаемых ресурсов по ГК2 преобладает доля денежных доходов, доля сбережений мала. Для данных регионов характерны низкая численность населения, высокий уровень безработицы, высокая рождаемость.

В целом, выделенные восемь кластеров различаются по уровню доходов, а также по структуре располагаемых ресурсов населения. Показано, что кластерная структура регионов РФ является поляризованной: несколько малочисленных кластеров находятся в области довольно высоких доходов, выделен немногочис-

ленный кластер, характеризующийся средним уровнем доходов, многочисленные кластеры находятся в области низких доходов. Выявлено, что с ростом доходов наблюдается снижение доли натуральных поступлений в виде продуктов питания в структуре доходов. Для малочисленных третьего и восьмого кластеров характерно преобладание доли денежных доходов, для второго кластера — преобладание доли привлеченных средств и израсходованных сбережений. Для остальных (довольно многочисленных) кластеров поддерживается баланс двух источников дохода в виде денежных доходов, а также привлеченных средств и израсходованных сбережений.

Цель анализа третьей выборки состоит в выявлении кластеров регионов РФ, различающихся по уровню располагаемых ресурсов, предоставляемых субсидий и субвенций с учетом дотации на выравнивание бюджетной обеспеченности субъектов РФ, представленных в абсолютных единицах.

Дотация в бюджетных отношениях — это определенная денежная сумма, выделяемая из бюджета более высокого уровня по решению соответствующего представительного органа власти в нижестоящий бюджет, не оговариваемая каким-либо целевым направлением средств. Дотация предоставляется в случаях, если закрепленных и регулирующих доходных источников недостаточно для формирования минимального бюджета нижестоящего территориального уровня. Субвенция в бюджетных отношениях — это денежная сумма, выделяемая из вышестоящего бюджета в нижестоящие бюджеты для конкретных целей на безвозмездной и безвозвратной основах. Субсидия — имеющая целевое назначение оплата предоставляемых гражданам материальных благ или оказываемых услуг. Переход к субвенциям связан с обострением необходимости более экономного и эффективного использования бюджетных средств. В случае использования субвенций не по целевому назначению уполномоченный Правительством Российской Федерации федеральный орган исполнительной власти, осуществляющий функции по контролю и надзору в финансово-бюджетной сфере, вправе осуществить взыскание средств в порядке, установленном законодательством Российской Федерации.

Особенностью выполнения интеллектуального анализа третьей выборки является проведение компонентного анализа для полной (анализируется 85 регионов РФ) и усеченной (анализируется 66 регионов РФ) выборки.

Весовые коэффициенты признаков (третья полная выборка)

№ п/п	Признак	Весовые коэффициенты признака		
		компонента 1	компонента 2	компонента 3
X <sub>1</sub>	Среднемесячная номинальная начисленная заработная плата	0,0988606	0,622079	0,307568
X <sub>2</sub>	Располагаемые ресурсы (в среднем на члена домашнего хозяйства)	0,161372	0,587087	0,292275
X <sub>3</sub>	Дотации на выравнивание бюджетной обеспеченности субъектов РФ	-0,0426704	-0,0449155	0,36364
X <sub>4</sub>	Межбюджетные трансферты на реализацию отдельных полномочий в области лекарственного обеспечения	0,389514	0,0956144	-0,304709
X <sub>5</sub>	Межбюджетные трансферты на реализацию мероприятий по профилактике ВИЧ-инфекции и гепатитов В и С	0,415228	-0,00454754	-0,23027
X <sub>6</sub>	Единая субвенция бюджетам субъектов РФ	0,430328	0,0294885	0,0642543
X <sub>7</sub>	Субвенции на выплаты инвалидам компенсаций страховых премий по договору ОСАГО	0,364081	0,100582	-0,398142
X <sub>8</sub>	Субвенции на социальные выплаты безработным гражданам	0,3688	-0,255649	0,0977104
X <sub>9</sub>	Субсидии на предоставление жилых помещений детям-сиротам и детям, оставшимся без попечения родителей, лицам из их числа по договорам найма специализированных жилых помещений бюджетам субъектов РФ	0,215839	-0,387716	0,508536
X <sub>10</sub>	Субвенции на выплату пособия при всех формах устройства детей, лишенных родительского попечения, в семью	0,371339	-0,174644	0,333271

В первой части компонентного анализа для полной выборки построены три главные компоненты. Анализируемых признаков — 10. Весовые коэффициенты признаков для этих компонент приведены в таблице 4.

Первые три главные компоненты содержат 81,41 % информации, что является достаточным для проведения исследований.

*Первая главная компонента* характеризует в целом объем межбюджетных трансфертов и субвенций. Высоким значениям первой главной компоненты соответствуют регионы с большим объемом межбюджетных трансфертов и субвенций.

*Вторая главная компонента* интегрально характеризует уровень доходов домохозяйств, который определяется в основном признаками «заработная плата» и «располагаемые ресурсы». Для объектов, расположенных в области высоких значений второй главной компоненты, характерен высокий уровень доходов домохозяйств.

*Третья главная компонента* характеризует объем дотаций на выравнивание бюджетной обеспеченности субъектов РФ. Общий объем дотаций определяется исходя из необходимости достижения минимального уровня расчетной бюджетной обеспеченности субъекта РФ, который зависит от бюджетных расходов и налогового потенциала субъекта. В структуре федеральных расходов в отношении регионов

объем дотаций является одной из крупнейших статей. Высоким значениям третьей главной компоненты соответствуют регионы с большим объемом дотаций.

Построена диаграмма рассеяния множества регионов РФ в пространстве выделенных двух первых главных компонент, по результатам анализа которой выделены следующие кластеры регионов (рис. 3).

*Первый кластер* составляет г. Москва. Он характеризуется высокими уровнями располагаемых ресурсов, заработной платы, а также межбюджетных трансфертов и субвенций. Для него характерен также самый низкий уровень дотации на выравнивание бюджетной обеспеченности по сравнению с другими кластерами.

*Второй кластер* включает 16 регионов РФ. Это регионы со средним уровнем межбюджетных трансфертов и субвенций, со средним или низким уровнем располагаемых ресурсов, а также со средним или высоким уровнем дотаций на выравнивание бюджетной обеспеченности: г. Санкт-Петербург, Камчатский край, Московская область, Свердловская область, Краснодарский край, Челябинская область, Алтайский край, Республика Татарстан, Красноярский край, Пермский край, Республика Башкортостан, Иркутская область, Новосибирская область, Нижегородская область и др.

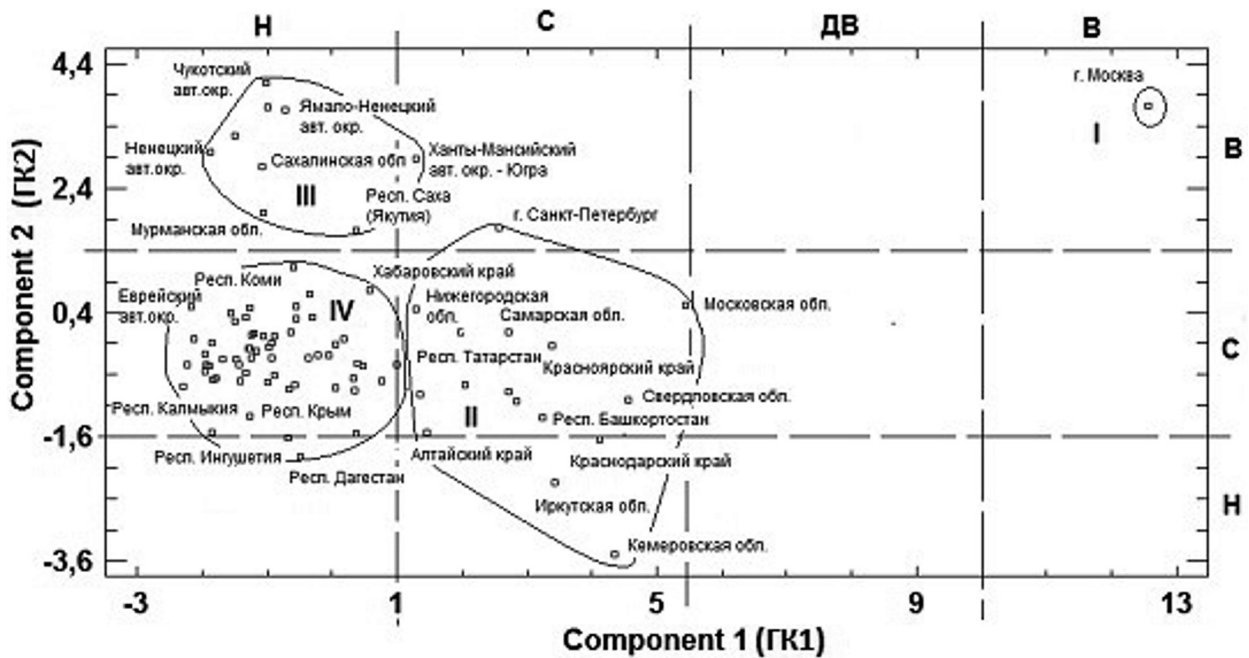


Рис. 3. Проекция множества регионов на пространство двух главных компонент (третья выборка)

Третий кластер включает 9 регионов: Республика Саха (Якутия), Камчатский край, Чукотский, Ямало-Ненецкий, Ненецкий, Ханты-Мансийский автономные округа, Сахалинская, Мурманская и Магаданская области. Для данных регионов характерен высокий уровень располагаемых ресурсов и низкий уровень межбюджетных трансфертов и субвенций.

Четвертый кластер, самый многочисленный, включает 59 субъектов РФ. Это регионы со средним уровнем располагаемых ресурсов и заработной платы и низким уровнем межбюджетных трансфертов и субвенций, однако с высоким объемом дотации на выравнивание бюджетной обеспеченности. Это такие регионы, как Хабаровский край, Архангельская, Оренбургская, Брянская, Ивановская, Липецкая, Костромская области, Республика Калмыкия, Республика Ингушетия, Республика Дагестан, Чеченская Республика, Еврейский автономный округ и др.

По результатам первой части компонентного анализа выделен крупный кластер, включающий наибольшее количество регионов РФ. Поскольку выявить особенности этого кластера при наличии регионов, сильно отличающихся от него, не представляется возможным, возникает необходимость анализа усеченной выборки.

Во второй части компонентного анализа составлена неполная выборка, включающая 84 региона РФ. Удален первый кластер (г. Москва).

Проанализированы первые три главные компоненты. Весовые коэффициенты признаков для этих компонент приведены в таблице 5.

Как показали результаты анализа коэффициентов информативности, изменений в характере компонент по сравнению с полной выборкой не произошло, поэтому названия компонент сохранены. Построена диаграмма рассеяния множества регионов РФ в пространстве выделенных двух первых главных компонент (рис. 4). Исследуемая совокупность всех регионов разделилась на семь достаточно выраженных групп.

Интерес представляют особенности расположения групп (кластеров) регионов на плоскости главных компонент.

Во-первых, в верхней правой части плоскости кластеры регионов отсутствуют, а в нижней левой части плоскости находятся достаточно населенные кластеры. В целом, если представить всю плоскость поделенной на две части с помощью диагонали (от левого верхнего угла и правому нижнему углу), то хорошо формируются две треугольные области, одна из которых заполнена объектами-регионами, а другая — пустая. Эта особенность хорошо интерпретируется с помощью выявленного смыслового содержания первых двух главных компонент. Регионы с высоким уровнем жизни в меньшем объеме охвачены субсидиями и субвенциями. По мере снижения уровня жизни и, соответственно, доходов домохозяйств регионов, объемы субсидий и субвенций, а также дотаций растут, что соответствует целям проводимой социально-экономической политики государства.

Во-вторых, распределение объектов-регионов внутри треугольной области неравномер-

Весовые коэффициенты признаков (третья неполная выборка)

№ п/п	Признак	Весовые коэффициенты признака		
		компонента 1	компонента 2	компонента 3
X <sub>1</sub>	Среднемесячная номинальная начисленная заработная плата	-0,0182104	0,689969	0,0474201
X <sub>2</sub>	Располагаемые ресурсы (в среднем на члена домашнего хозяйства)	0,0622961	0,685025	-0,034541
X <sub>3</sub>	Дотации на выравнивание бюджетной обеспеченности субъектов РФ	-0,0233045	0,00809917	0,792195
X <sub>4</sub>	Межбюджетные трансферты на реализацию отдельных полномочий в области лекарственного обеспечения	0,396408	-0,023611	-0,104347
X <sub>5</sub>	Межбюджетные трансферты на реализацию мероприятий по профилактике ВИЧ-инфекции и гепатитов В и С	0,408169	-0,0423179	-0,106468
X <sub>6</sub>	Единая субвенция бюджетам субъектов РФ	0,421942	0,148934	0,0463415
X <sub>7</sub>	Субвенции на выплаты инвалидам компенсаций страховых премий по договору ОСАГО	0,314723	-0,0332128	-0,427671
X <sub>8</sub>	Субвенции на социальные выплаты безработным гражданам	0,383214	-0,122552	0,143967
X <sub>9</sub>	Субсидии на предоставление жилых помещений детям-сиротам и детям, оставшимся без попечения родителей, лицам из их числа по договорам найма специализированных жилых помещений бюджетам субъектов РФ	0,301613	-0,10357	0,338929
X <sub>10</sub>	Субвенции на выплату пособия при всех формах устройства детей, лишенных родительского попечения, в семью	0,39572	0,0569201	0,161596

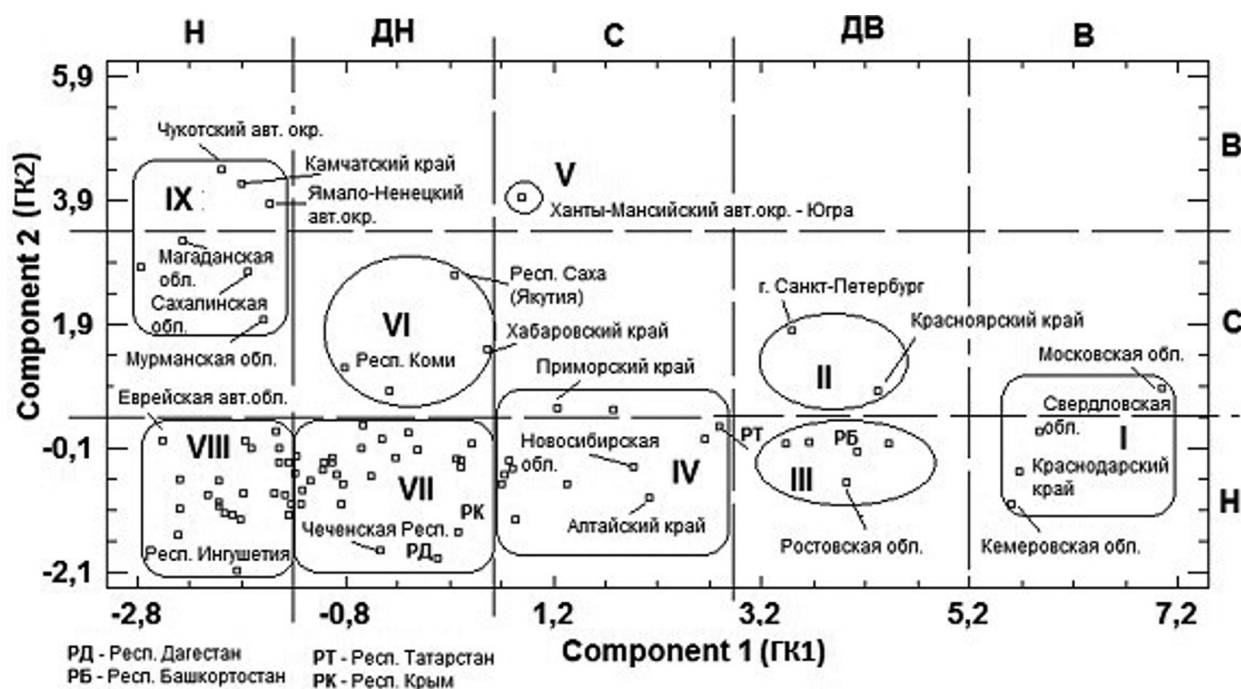


Рис. 4. Проекция множества регионов на пространство двух главных компонент (третья усеченная выборка)

ное. При движении справа налево населенность кластеров возрастает, однако возрастание плотности расположения объектов-регионов происходит вдоль координатной оси, соответствующей ГК1. Это значит, что часть регионов при наличии достаточно низкого уровня жизни получают недостаточные по объему суб-

сидии, субвенции и дотации. В качестве желаемого (идеального) расположения объектов-регионов, которому соответствует сбалансированная поддержка государства в виде межбюджетных трансфертов, можно было бы назвать расположение кластеров регионов вдоль диагонали, что предполагает оказание большей

помощи регионам с более низким уровнем жизни, и наоборот.

Более детальная характеристика кластеров представлена ниже.

Первый кластер включает 4 региона РФ. Это регионы с высоким объемом предоставляемых субсидий, субвенций и трансфертов, а также средним и низким уровнем располагаемых ресурсов, заработной платы, и дотации на выравнивание бюджетной обеспеченности. Это такие регионы, как Кемеровская область, Московская область, Краснодарский край, Свердловская область.

Второй кластер включает 2 субъекта РФ: г. Санкт-Петербург и Красноярский край. Это регионы с достаточно высоким объемом предоставляемых субсидий, субвенций и трансфертов, а также средним уровнем располагаемых ресурсов, заработной платы и дотации на выравнивание бюджетной обеспеченности.

Третий кластер включает 5 регионов РФ: Ростовская область, Республика Башкортостан, Иркутская область, Самарская область, Челябинская область. Это регионы с достаточно высоким объемом предоставляемых субсидий, субвенций и трансфертов, а также низким уровнем располагаемых ресурсов, заработной платы и дотаций на выравнивание бюджетной обеспеченности.

Четвертый кластер включает 12 регионов РФ: Приморский край, Республика Татарстан, Пермский край, Алтайский край, Оренбургская, Нижегородская, Новосибирская области и другие. Это регионы со средним объемом предоставляемых субсидий, субвенций и трансфертов, низким уровнем располагаемых ресурсов, заработной платы, а также средним и низким уровнем дотаций на выравнивание бюджетной обеспеченности.

Пятый кластер включает 1 регион — Ханты-Мансийский автономный округ — Югра. Для этого региона характерен средний объем предоставляемых субсидий, субвенций и трансфертов, а также высокий уровень располагаемых ресурсов, заработной платы и средний уровень дотаций на выравнивание бюджетной обеспеченности.

Шестой кластер включает 4 субъекта РФ: Республика Саха (Якутия), Республики Коми, Хабаровский край, Архангельская область. Это регионы с довольно низким объемом субвенций, субсидий и трансфертов, средним уровнем располагаемых ресурсов. Для регионов данного кластера характерен высокий объем предоставляемых дотаций на выравнивание бюджетной обеспеченности.

Далее два самых многочисленных кластера.

Седьмой кластер включает 26 регионов РФ: республики Крым, Бурятия и Дагестан, Чеченская Республика, Тверская, Владимирская, Курская и Кировская области и др. Это регионы с довольно низким объемом субвенций, субсидий и трансфертов, низким уровнем располагаемых ресурсов. Для регионов данного кластера характерен средний объем предоставляемых дотаций на выравнивание бюджетной обеспеченности.

Восьмой кластер включает 23 регионов РФ. Это регионы с низким уровнем располагаемых ресурсов и низким объемом предоставляемых дотаций на выравнивание бюджетной обеспеченности, низким уровнем предоставляемых субсидий и субвенций. Это такие регионы, как республики Ингушетия, Калмыкия, Адыгея, Алтай, Северная Осетия, Кабардино-Балкарская Республика, Псковская область, Орловская область, Еврейский автономный округ и другие.

Девятый кластер включает 7 субъектов РФ: Чукотский автономный округ, Камчатский край, Ямало-Ненецкий автономный округ, Ненецкий автономный округ, Магаданская область, Мурманская область и Сахалинская область. Это регионы с низким объемом субвенций, субсидий и трансфертов, с высоким и средним уровнем располагаемых ресурсов. Для регионов данного кластера характерен средний объем предоставляемых дотаций на выравнивание бюджетной обеспеченности.

### Заключение

Таким образом, с помощью метода главных компонент проведен анализ статистических данных о доходах домохозяйств регионов РФ. Цель анализа состоит в выявлении закономерности типа «кластеризация», что предполагает, во-первых, выявление кластеров регионов РФ, различающихся по структуре доходов населения, численности населения и ее динамике, а также объемам дотаций, субсидий и субвенций, направляемых на поддержку социально незащищенных слоев населения, и, во-вторых, формирование обобщенных характеристик выявленных кластеров регионов и представление их в виде правил кластеризации. Результаты анализа представлены в виде количества кластеров регионов РФ и их состава, различающегося по структуре доходов населения, а также в виде профилей кластеров и правил кластеризации.

Статистические данные объединены в три выборки.

Определены особенности построенных кластеров регионов РФ, которые учтены при формировании профилей кластеров и правил кластеризации.

По результатам анализа первой выборки выявлены малочисленные кластеры регионов, характеризующиеся высоким уровнем денежных доходов домохозяйств при достаточно высокой численности населения. Сформированы достаточно населенные кластеры регионов, характеризующиеся значительным объемом денежных доходов в виде натуральных поступлений продуктов питания при низком и среднем уровне денежных доходов и малой положительной динамике численности населения.

По результатам анализа второй выборки выявлен состав кластеров регионов РФ, различающихся по типовым пропорциям в структуре доходов с учетом различий в расходах на конечное потребление. Показано, что регионы, характеризующиеся достаточно низкими денежными доходами, имеют в структуре доходов высокую долю натуральных поступлений в виде продуктов питания. Показано также, что регионы, характеризующиеся высокими денежными доходами, имеют в структуре доходов малую долю привлеченных средств и израсходованных сбережений.

Полученные результаты анализа первых двух выборок позволяют заключить, что кластерная структура сектора домохозяйств, рассмотренная на региональном уровне, является достаточно поляризованной. Это предполагает наличие малочисленных кластеров благополучных регионов с достаточно высокими доходами населения и в то же время существование многочисленных кластеров регионов, население которых имеет незначительные доходы. Определены структурные и параметри-

ческие характеристики расширенной динамической модели сектора домохозяйств СИДМИУ. Построенные кластеры регионов, их количество и состав, а также их обобщенные характеристики в виде правил используются для определения количества подсистем в расширенной динамической модели сектора домохозяйств СИДМИУ и их особенностей.

По результатам анализа третьей выборки выявлены кластеры домохозяйств регионов, характеризующиеся либо высоким уровнем доходов при достаточно высоких трансфертах, либо большим объемом дотации на выравнивание бюджетной обеспеченности при низком и среднем уровне доходов.

Полученные результаты анализа третьей выборки в виде кластеров регионов и их характеристик интегрируют в себе текущий уровень доходов населения регионов и принимаемые решения по предоставлению субсидий, субвенций и дотаций на выравнивание бюджетной обеспеченности. Эти результаты позволят настроить параметры текущего состояния модуля интеллектуального управления СИДМИУ на региональном уровне и сформировать дополнительные корректирующие правила принятия решений, направленные на снижение дифференциации населения регионов по уровню доходов.

В целом, расширенная динамическая модель сектора домохозяйств с кластерным делением регионов в составе СИДМИУ совместно с расширенным модулем интеллектуального управления позволяет проводить сценарные исследования различных вариантов принятия решений при управлении доходами населения на региональном уровне с учетом дифференциации населения регионов по уровню доходов.

### Благодарность

*Работа выполнена при финансовой поддержке РФФИ (проект № 17-08-01155).*

### Список источников

1. Попова Г. Л., Филатова О. В. Доходы населения: сравнительный анализ и прогноз // Вопросы современной науки и практики. Университет им. В. И. Вернадского. — 2013. — № 1(45). — С. 185–190.
2. Бакуменко Л. П., Мхитарян В. С. Анализ территориальных различий уровня доходов населения в мезосистеме // Экономика, статистика и информатика. — 2010. — № 2. — С. 86–91.
3. Абдусаламова М. М. Анализ территориальных различий уровня доходов и качества жизни населения // Региональная экономика. Теория и практика. — 2013. — № 47(326). — С. 62–68.
4. Бобков В. Н., Гулюгина А. А. Мониторинг доходов и уровня жизни населения России в IV квартале 2013 года // Уровень жизни населения регионов России. — 2014. — № 1(191). — С. 26–37.
5. Малкина М. Ю. Социальное благополучие регионов Российской Федерации // Экономика региона. — 2017. — Т. 13, вып. 1 — С. 49–62. — DOI: 10.17059/2017-1-5.
6. Малкина М. Ю. Динамика и факторы внутрирегиональной и межрегиональной дифференциации доходов населения РФ // Пространственная экономика. — 2014. — № 3. — С. 44–66. — DOI: 10.14530/se.2014.3.44-66.

7. Нижегородцев Р. М., Пискун Е. И., Кудревич В. В. Прогнозирование показателей социально-экономического развития региона // Экономика региона. — 2017. — Т. 13, вып. 1 — С. 38–48. — DOI: 10.17059/2017–1-4.
8. Стукаленко Е. А. Дифференциация доходов населения. Причины и последствия // Вестник Омского университета. — 2014. — № 1. — С. 183–187. — (Экономика).
9. Lambert R. J. The Distribution and Redistribution of Income; third ed. — UK, Manchester : Manchester University Press, 2002. — 336 p.
10. Howarth R. B., Kennedy K. Economic growth, inequality, and well-being // Ecological Economics. — 2016. — Vol. 121. — Pp. 231–236. — DOI: 10.1016/j.ecolecon.2015.10.005.
11. Carlsson F, Daruvala D., Johansson-Stenman O. Are people inequality-averse, or risk-averse? // Economica. — 2005. — Vol. 72 (287). — Pp. 375–396. — DOI: 10.1111/j.0013–0427.205.00421.
12. Куклин А. А., Леонтьева А. Г. Дифференциация доходов населения и проблема бедности // Экономика региона. — 2011. — № 1. — С. 93–98.
13. Кузнецов С. В., Растова Ю. И., Растов М. А. Рейтинговая оценка качества жизни в российских регионах // Экономика региона. — 2017. — Т. 13, вып. 1. — С. 137–146. — DOI: 10.17059/2017–1-13.
14. Кузнецов С. В., Межевич Н. М. Новые практики территориального управления в России и вопросы управления качеством жизни // Управленческое консультирование. — 2015. — № 7. — С. 25–34.
15. Плотников В. А., Шамахов В. А. Стратегии территориального развития и качество жизни // Управленческое консультирование. — 2015. — № 7. — С. 57–64.
16. Анализ данных и процессов: учеб. пособие / А. А. Барсебян, М. С. Куприянов, И. И. Холод и др. 3-е изд., перераб. и доп. — СПб. : БХВ-Петербург, 2009. — 512 с.
17. Кулаичев А. П. Методы и средства комплексного анализа данных: учебное пособие. 4-е изд., перераб. и доп. — М. : Форум; Инфра-М, 2013. — 312 с.
18. Dean J. Big Data, Data Mining, and Machine Learning. — Hoboken, New Jersey : John Wiley & Sons, Inc., 2014. — 265 p.
19. Han J, Kamber M., Pei J. Data Mining: Concepts and Techniques. — San Mateo: Morgan Kaufmann Publishers is an imprint of Elsevier, 2012. — 703 p.
20. Mitra S., Acharya T. Data Mining, Multimedia, Soft Computing, and Bioinformatics. — Hoboken, New Jersey: John Wiley & Sons Inc., 2003. — 424 p.
21. Zaki M. J., Meira W. Data Mining and Analysis: Fundamental Concepts and Algorithms. — New York : Cambridge University Press, 2014. — 562 p.
22. Ruppert D., Matteson D. S. Statistics and Data Analysis for Financial Engineering with R Examples. — New York : Springer Science+Business Media, 2015. — 719 p.
23. Объективно-субъективный подход к оценке качества жизни / Ильясов Б. Г., Герасимова И. Б., Макарова Е. А., Закиева Е. Ш. // Качество. Инновации. Образование. — 2016. — № 2 (129). — С. 47–57.
24. Гузаиров М. Б., Десярева И. В., Макарова Е. А. Расходы населения регионов Российской Федерации на покупку продуктов питания: компонентный и кластерный анализ // Экономика региона. — 2015. — № 4. — С. 145–157. — DOI: 10.17059/2015–4-12.
25. Ильясов Б. Г., Макарова Е. А., Валитов Р. Р. Регулирование доходов населения и анализ их влияния на динамику потребительского спроса на основе имитационного моделирования // Научно-технические ведомости СПбГПУ. — 2012. — № 5. — С. 67–71.

### Информация об авторах

**Ильясов Барый Галеевич** — доктор технических наук, профессор, профессор кафедры технической кибернетики, Уфимский государственный авиационный технический университет; Scopus Author ID: 6603426467 (Российская Федерация, 450000, Республика Башкортостан, г. Уфа, ул. К.Маркса, 12, 6–216; e-mail: ilyasov@tc.ugatu.ac.ru).

**Макарова Елена Анатольевна** — доктор технических наук, профессор, профессор кафедры технической кибернетики, Уфимский государственный авиационный технический университет; Scopus Author ID: 57193092121 (Российская Федерация, 450000, Республика Башкортостан, г. Уфа, ул. Карла Маркса, 12; e-mail: ea-makarova@mail.ru).

**Закиева Елена Шавкатовна** — кандидат технических наук, доцент, доцент кафедры технической кибернетики, Уфимский государственный авиационный технический университет; ORCID: 0000–0002–6921–7473; Researcher ID: D-4447–2017; Scopus Author ID: 57193094790 (Российская Федерация, 450000, Республика Башкортостан, г. Уфа, ул. К.Маркса, 12, 6–211; e-mail: zakievae@mail.ru).

**Гиздатуллина Эмма Салаватовна** — аспирант, Уфимский государственный авиационный технический университет (Российская Федерация, 450000, Республика Башкортостан, г. Уфа, ул. К.Маркса, 12, 6–109; e-mail: gizdatullina@mail.ru).

For citation: Ilyasov, B. G., Makarova, E. A., Zakieva, E. S. & Gizdatullina, E. S. (2019). Analysing the Data on Incomes in the Regional Context by the Principal Component Method. *Ekonomika regiona [Economy of region]*, 15(2), 601-617

**B. G. Ilyasov, E. A. Makarova, E. S. Zakieva, E. S. Gizdatullina**  
Ufa State Aviation Technical University (Ufa, Russian Federation; e-mail: zakievae@mail.ru)

### Analysing the Data on Incomes in the Regional Context by the Principal Component Method

The article focuses on solving the task of analysing statistical data on households' income and their main components in absolute and relative units. We took into account a number of additional indicators, including social transfers, and applied the principle component method. The analysis' purpose was to identify patterns of «clustering». The first step was to identify clusters of the Russian Federation regions, which vary in terms of population's revenue structure taking into account the volumes of subsidies and subventions. The second step was to determine the generalized characteristics of the revealed clusters and their representation in a form of clustering rules. We have shown that the cluster structure of the households sector at the regional level is sufficiently polarized. We have revealed the small clusters of regions characterized by a high level of households' monetary income and relatively large population (e. g. Moscow, Khanty-Mansi Autonomous Okrug). Alternatively, there are sufficiently inhabited clusters of regions with both a considerable volume of non-monetary income in a form of food combined and the low or average level of monetary income and small positive dynamics of population (Bryansk, Kursk Oblasts). On the other hand, in the regions with a relatively low monetary income, the revenue structure includes a high share of natural supplies in the form of food (for example, Republic of Dagestan and Republic of Ingushetia). Moreover, in the regions with a high monetary income, there is a small share of the raised funds and spent savings in revenue structure (Yamalo-Nenets Autonomous Okrug and others). We have constructed clusters of regions and established their quantity, structure and generalized characteristics presented in the form of clustering rules. We used that data for defining structural and parametrical characteristics when developing a dynamic model of the households sector and the module of intellectual management. These dynamic model and the module became a part of the system of imitating dynamic modelling and intellectual management (SIDMIM) of population income generation. The application of SIDMIM involves scenario studies for decision-making in managing the population income at the regional level considering differentiation in the income level.

**Keywords:** population income, principal component method, sample, clustering, weight coefficient, coefficient of information content, integrated sign, scatterplot, clusters of regions, imitating dynamic model

### Acknowledgments

The article has been supported by Russian Foundation for Basic Research (the project № 17-08-01155).

### References

1. Popova, G. L. & Filatova, O. V. (2013). Dokhody naseleniya: sravnitelnyy analiz i prognoz (na primere Tambovskoy i Volgogradskoy oblastey) [Population income: comparative analysis and forecast (the Tambov and Volgograd regions)]. *Voprosy sovremennoy nauki i praktiki. Universitet im. V. I. Vernadskogo [Problems of contemporary science and practice. Vernadsky University]*, 1(45), 185–190. (In Russ.)
2. Bakumenko, L. P. & Mhitaryan, V. S. (2010). Analiz territorialnykh razlichiy urovnya dokhodov naseleniya v mezosisteme [Analysis of regional differences of income levels in mesosystem]. *Ekonomika, statistika i informatika [Economics, statistics and Informatics]*, 2, 86–91. (In Russ.)
3. Abdusalomova, M. M. (2013). Analiz territorialnykh razlichiy urovnya dokhodov i kachestva zhizni naseleniya [Analysis of territorial distinctions of level of income and quality of life of the population]. *Regionalnaya ekonomika: teoriya i praktika [Regional economy: theory and practice]*, 47(326), 62–68. (In Russ.)
4. Bobkov, V. N. & Gulyugina, A. A. (2014). Monitoring dokhodov i urovnya zhizni naseleniya Rossii v IV kvartale 2013 goda [Monitoring incomes and living standards of the population of Russia in the 4th quarter of 2013]. *Uroven zhizni naseleniya regionov Rossii [Living standards and quality of life]*, 1(191), 26–37. (In Russ.)
5. Malkina, M. Yu. (2017). Sotsialnoe blagopoluchie regionov rossiyskoy federatsii [Social well-beng of the Russian Federation regions]. *Ekonomika regiona [Economy of Region]*, 13(1), 49–62. DOI: 10.17059/2017-1-5. (In Russ.)
6. Malkina, M. Yu. (2014). Dinamika i factory vnutriregionalnoy i mezhregionalnoy differentsiatsii dokhodov naseleniya RF [Dynamics and determinants of intra and inter-regional income differentiation of the population of the Russian Federation regions]. *Prostranstvennaya ekonomika [Spatial economics]*, 3, 44–66. DOI: 10.14530/se.2014.3.44-66. (In Russ.)
7. Nizhegorodtsev, R. M., Piskun, E. I. & Kudrevich, V. V. (2017). Prognozirovanie pokazateley sotsialno-ekonomicheskogo razvitiya regiona [The forecasting of regional social and economic development]. *Ekonomika regiona [Economy of Region]*, 13(1), 38–48. DOI: 10.17059/2017-1-4. (In Russ.)
8. Stukalenko, E. A. (2014). Differentsiatsiya dokhodov naseleniya: prichiny i posledstviya [Differentiation of population's incomes: causes and consequences]. *Vestnik Omskogo universiteta. Seriya «Ekonomika» [Herald of Omsk University. Series "Economics"]*, 1, 183–187. (In Russ.)
9. Lambert, R. J. (2002). *The Distribution and Redistribution of Income* (3rd edition). UK, Manchester: Manchester University Press, 336.
10. Howarth, R. B. & Kennedy, K. (2016). Economic growth, inequality, and well-being. *Ecological Economics*, 121, 231–236. DOI: 10.1016/j.ecolecon.2015.10.005.

11. Carlsson, F., Daruvala, D. & Johansson-Stenman, O. (2005). Are people inequality-averse, or risk-averse? *Economica*, 72(287), 375–396. DOI: 10.1111/j.0013-0427.205.00421.
12. Kuklin, A. A. & Leonteva, A. G. (2011). Differentsiatsiya dokhodov naseleniya I problema bednosti [Differentiation of population income and poverty problem (based on the data of Tyumen region)]. *Ekonomika regiona [Economy of Region]*, 1, 93–98. (In Russ.)
13. Kuznetsov, S. V., Rastova, Yu. I. & Rastov, M. A. (2017). Reytingovaya otsenka kachestva zhizni v rossiyskikh regionakh [Rating evaluation of the quality of life in Russian regions]. *Ekonomika regiona [Economy of Region]*, 13(1), 137–146. (In Russ.)
14. Kuznetsov, S. V. & Mezhevich, N. M. (2015). Novye praktiki territorialnogo upravleniya v Rossii I voprosy upravleniya kachestvom zhizni [New practices of territorial governance in Russia and quality of life management issues]. *Upravlencheskoe konsultirovanie [Administrative consulting]*, 7, 25–34. (In Russ.)
15. Plotnikov, V. A. & Shamakhov, V. A. (2015). Strategii territorialnogo razvitiya i kachestvo zhizni [Strategies of territorial development and quality of life]. *Upravlencheskoe konsultirovanie [Administrative consulting]*, 7, 57–64. (In Russ.)
16. Barsegyan, A. A., Kupriyanov, M. S., Holod, I. I., Tess, M. D. & Elizarov, S. I. (2009). *Analiz dannykh I protsessov [Data and processes analysis]*. Saint-Petersburg: BHV-Petersburg, 512. (In Russ.)
17. Kulaichev, A. P. (2013). *Metody I sredstva kompleksnogo analiza dannykh [Methods and means of complex data analysis]*. Moscow: FORUM: INFRA-M, 312. (In Russ.)
18. Dean, J. (2014). *Big Data, Data Mining, and Machine Learning*. Hoboken, New Jersey, John Wiley & Sons, Inc., 265.
19. Han, J., Kamber, M. & Pei, J. (2012). *Data Mining: Concepts and Techniques*. San Mateo: Morgan Kaufmann Publishers is an imprint of Elsevier, 703.
20. Mitra, S. & Acharya, T. (2003). *Data Mining, Multimedia, Soft Computing, and Bioinformatics*. Hoboken, New Jersey, John Wiley & Sons, Inc., 424.
21. Zaki, M. J. & Meira, W. (2014). *Data Mining and Analysis: Fundamental Concepts and Algorithms*. New York: Cambridge University Press, 562.
22. Ruppert, D. & Matteson, D. S. (2015). *Statistics and Data Analysis for Financial Engineering with R Examples*. New York: Springer Science + Business Media, 719.
23. Ilyasov, B. G., Gerasimova, I. B., Makarova, E. A. & Zakieva, E. Sh. (2016). Obektivno-subektivnyy podkhod k otsenke kachestva zhizni [Objective-subjective approach to the assessment of quality of life]. *Kachestvo. Innovatsii. Obrazovanie [Quality. Innovation. Education]*, 2(129), 47–57. (In Russ.)
24. Guzairov, M. B., Degtyareva, I. V. & Makarova, E. A. (2015). Raskhody naseleniya regionov Rossiysko Federatsii na pokupku produktov pitaniya: komponentnyy i klasternyy analiz [Regional Population Expenditure for Foodstuffs in the Russian Federation: Componential and Cluster Analyses]. *Ekonomika regiona [Economy of region]*, 4, 145–157. (In Russ.)
25. Ilyasov, B. G., Makarova, E. A. & Valitov, R. R. (2012). Regulirovanie dokhodov naseleniya I analiz ikh vliyaniya na dinamiku potrebitelskogo sprosa na osnove imitatsionnogo modelirovaniya [Regulation of incomes and the analysis of their influence on dynamics of consumer demand is based on simulation modelling]. *Nauchno-tekhnicheskie vedomosti SPbGPU [St. Petersburg State Polytechnical University Journal]*, 5, 67–71. (In Russ.)

### Authors

**Baryy Galeevich Ilyasov** — Doctor of Engineering, Professor, Department of Engineering Cybernetics, Ufa State Aviation Technical University; Scopus AuthorID: 6603426467 (12/6, K. Marksa st., Ufa, Bashkortostan, 450000, Russian Federation; e-mail: ilyasov@tc.ugatu.ac.ru).

**Elena Anatolyevna Makarova** — Doctor of Engineering, Professor, Department of Engineering Cybernetics, Ufa State Aviation Technical University; Scopus Author ID: 57193092121 (12/6, K. Marksa st., Ufa, Bashkortostan, 450000, Russian Federation; e-mail: ea-makarova@mail.ru).

**Elena Shavkatovna Zakieva** — PhD in Engineering, Associate Professor, Department of Engineering Cybernetics, Ufa State Aviation Technical University; ORCID: 0000-0002-6921-7473; Researcher ID: D-4447-2017; Scopus AuthorID: 57193094790 (12/6, K. Marksa st., Ufa, Bashkortostan, 450000, Russian Federation; e-mail: zakievae@mail.ru).

**Emma Salavatovna Gizdatullina** — PhD Student, Ufa State Aviation Technical University (12/6, K. Marksa st., Ufa, Bashkortostan, 450000, Russian Federation; e-mail: gizdatullina@mail.ru).