

УДК 574.3.34

**ИССЛЕДОВАНИЕ ВЗАИМОЗАВИСИМОСТИ ЭКОНОМИЧЕСКИХ,
СОЦИАЛЬНЫХ, ДЕМОГРАФИЧЕСКИХ И ЭКОЛОГИЧЕСКИХ
ХАРАКТЕРИСТИК РЕГИОНОВ РОССИЙСКОЙ ФЕДЕРАЦИИ
МЕТОДАМИ ФАКТОРНОГО АНАЛИЗА**

© 2008 г.

С.В. Барышников, А.Я. Моничев, В.А. Басуров

Нижегородский госуниверситет им. Н.И. Лобачевского

vestnik@unn.ru

Поступила в редакцию 12.05.2008

С помощью факторного анализа обнаружены взаимосвязи между различными группами региональных характеристик. Выделены наиболее значимые характеристики, определяющие существенные особенности регионов Российской Федерации.

Ключевые слова: факторный анализ, регион, характеристики.

Каждый регион Российской Федерации характеризуется набором показателей, которые можно подразделить на ряд групп в зависимости от того, к какой сфере человеческой деятельности они относятся. Так, можно выделить в отдельные группы экологические, экономические, медико-демографические и социальные показатели. Наряду с тем представляется очевидным, что многие подобные региональные характеристики, относящиеся к различным группам, тесно связаны между собой. Так, например, высокий уровень экономического развития региона обычно сопровождается высокой интенсивностью воздействия на окружающую среду, а благоприятная социальная обстановка, сопровождающая обычно высокий уровень экономического развития региона, способствует увеличению продолжительности жизни. Отсюда видно, что некоторые характеристики являются ключевыми. Задача выявления ключевых характеристик, безусловно, сопряжена с изучением связей и зависимостей между различного рода региональными показателями. В подобных задачах может быть использован метод многомерного факторного анализа с последующей кластеризацией данных [1], который позволяет определить ключевые факторы как комбинацию взаимосвязанных характеристик, а также дает возможность выявить группы регионов, сходных по ключевым факторам, изучая их положение в многомерном факторном пространстве.

Целью данной работы как раз и является применение метода факторного анализа для выявления взаимосвязей между характеристиками регионов с выделением ключевых характеристик, определяющих их развитие. Данный ме-

тод был применен нами по отношению к статистическим данным, собранным по регионам Российской Федерации специалистами Института народнохозяйственного прогнозирования РАН, г. Москва. Процедуры факторизации и кластеризации данных осуществлялись с помощью стандартной компьютерной программы «Statistica» [2].

Полученное в процессе математической обработки взаимное расположение изучаемых характеристик в трехмерном факторном пространстве представлено на рис. 1. Чтобы не усложнять восприятие рисунка большим количеством подписей, на нем приведены только некоторые из изучаемых характеристик. При этом, как кажется, более ясную картину взаимосвязанности рассматриваемых характеристик дает проекция полученного распределения на одну из факторных плоскостей (рис. 2).

Зная положение каждой характеристики в трехмерном пространстве, мы можем применить метод кластерного анализа к имеющемуся набору показателей для разделения изучаемых характеристик на группы по степени их близости в факторном пространстве, то есть «по средству» (рис. 3).

В результате был получен приведенный ниже состав выделенных кластеров, отображенных на рис. 2 и 3 римскими цифрами (характеристики, входящие в кластеры, пронумерованы в соответствии с дендрограммой на рис. 3).

В первый кластер (помечен I) попали следующие характеристики: уровень рождаемости (1), уровень младенческой смертности (2), доля населения с величиной дохода ниже прожиточного минимума (3), уровень зарегистрированной

безработицы (4), приблизительный уровень фактической безработицы (5), ожидаемая продолжительность жизни мужчин (7), ожидаемая продолжительность жизни женщин (6), средняя ожидаемая продолжительность жизни (8), плотность

населения (9), отношение среднего размера пенсий к величине прожиточного минимума (11).

Второй кластер (II): уровень смертности (17), число умерших от заболеваний органов пищеварения (12), инфекционных и паразитар-

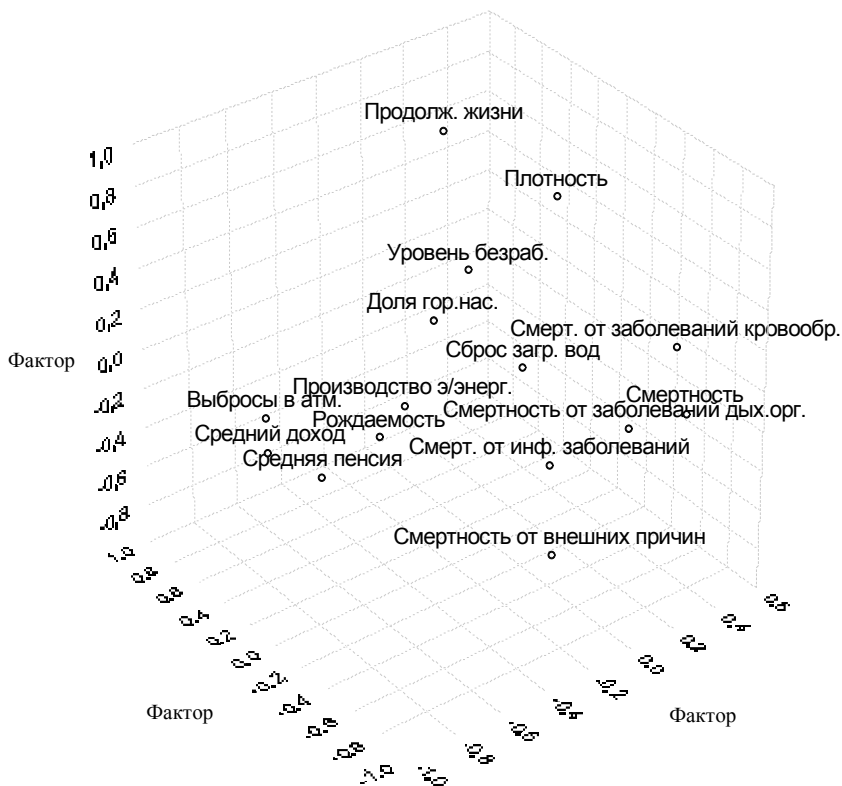


Рис. 1. Распределение региональных характеристик в трехмерном пространстве факторов

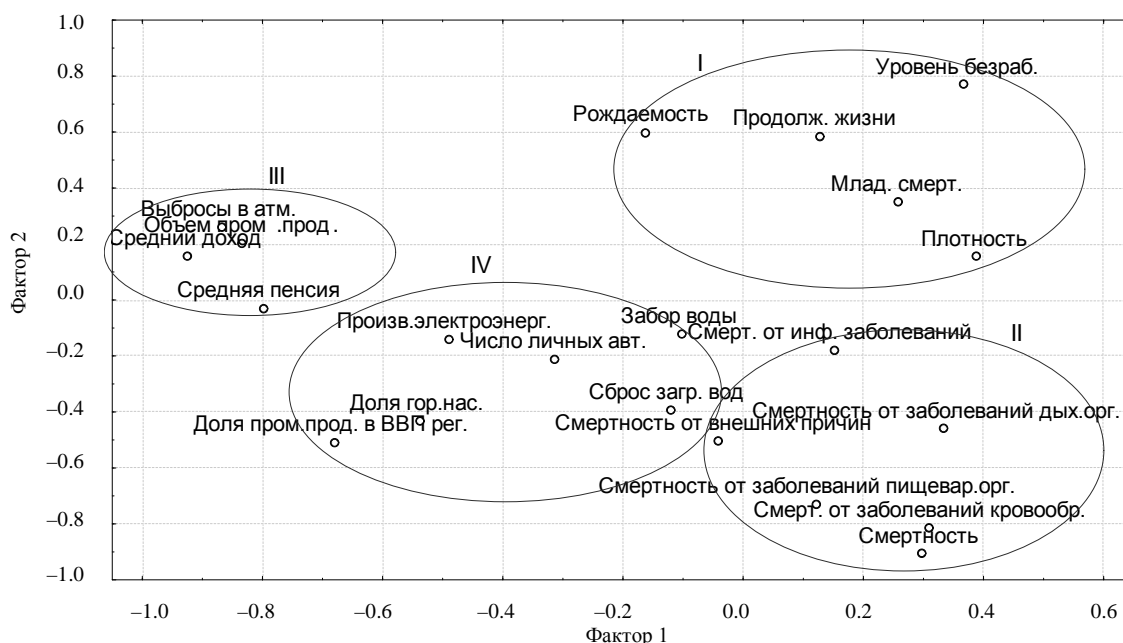


Рис. 2. Проекция пространственного распределения региональных характеристик на одну из факторных плоскостей

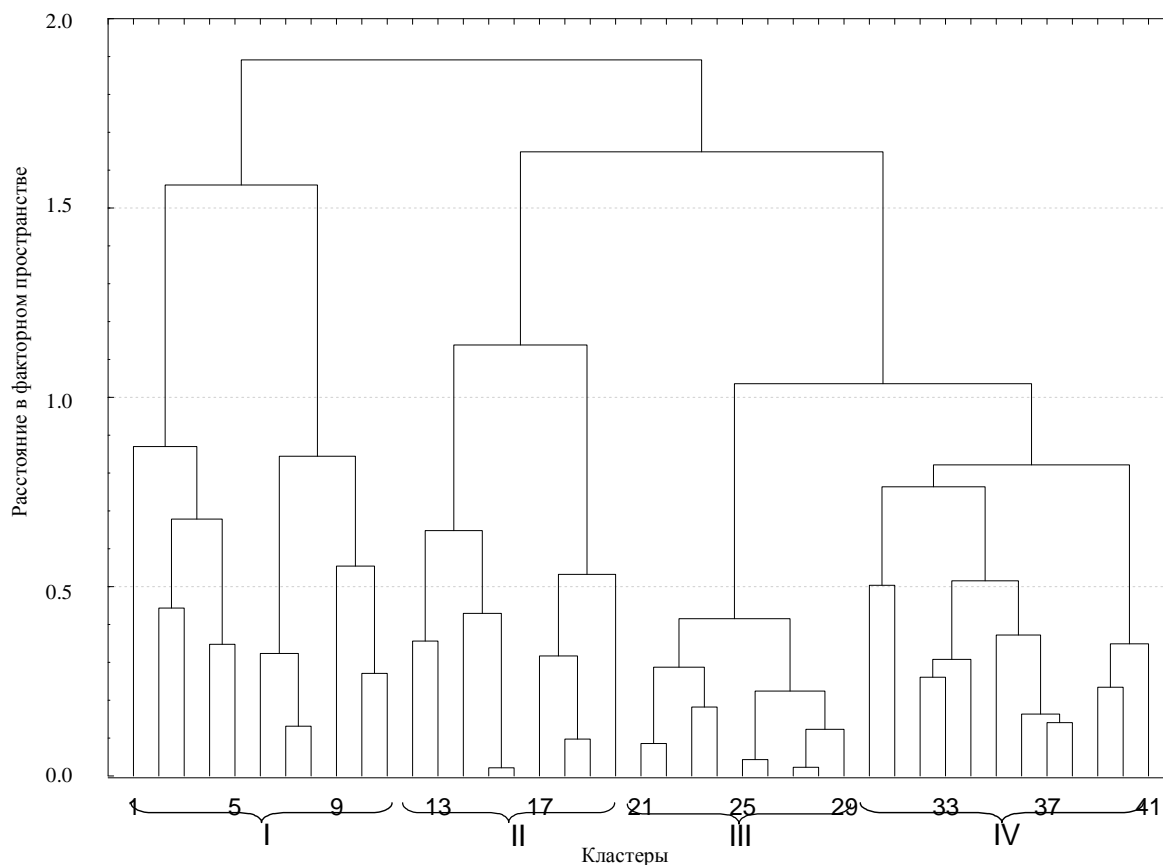


Рис. 3. Дендрограмма кластеризации региональных характеристик

ных заболеваний (16), внешних причин смертности (14), удельная обеспеченность жилплощадью (20).

Третий кластер (III): размер среднедушевого дохода (21), размер заработной платы (22) и пенсии (23), отношение размера заработной платы к величине прожиточного минимума (24), коэффициент Джинни (25), объем выбросов в атмосферу (27), объем валового внутреннего продукта (28), объем произведенной промышленной продукции (29).

Четвертый кластер (IV): число преступлений (31), площадь региона (30), число личных автомобилей (35), объем улавливаемых выбросов в атмосферу (32), объем забора воды (33), объем сброса сточных вод (34), объем оборотной и последовательно используемой воды (37), отношение среднедушевого дохода к прожиточному минимуму (36), количество произведенной электроэнергии (38), доля городского населения в регионе (39), доля промышленной продукции в ВВП региона (40).

Анализируя полученное распределение, можно сделать ряд выводов. Так, например, видно, что в первый кластер входят характеристики, определяющие уровень жизни и социальную обеспеченность в регионе. Эта же груп-

пировка включает в себя важные медико-демографические характеристики, такие как продолжительность жизни и уровень младенческой смертности, что говорит о связи продолжительности жизни с развитием системы социальной защиты в регионе. Присутствие плотности населения в этой же группе свидетельствует о регулирующем воздействии плотности населения на демографические процессы. Второй кластер содержит, в основном, характеристики, определяющие показатели смертности в регионе. Помимо них в этот кластер входят такие характеристики, как уровень преступности и удельная обеспеченность жилплощадью. В третью группу входят характеристики, определяющие уровень социальной обеспеченности населения в регионе, а также показатели, высокие значения которых свидетельствуют об индустриальной развитости. Отсюда следует заключить, что повышение уровня жизни людей связано с улучшением экономической ситуации в регионе. Это свидетельствует о том, что высокая преступность характерна, по-видимому, для регионов с низким уровнем жизни. Последний, четвертый кластер включает в себя характеристики, отражающие уровень материальной обеспеченности населения, промышленную развитость региона и сте-

пень антропогенного воздействия на окружающую среду, что в очередной раз подтверждает предположение о положительной взаимосвязи этих групп показателей.

В процессе факторного анализа число показателей, характеризующих каждый изучаемый объект, значительно уменьшается. В нашем случае расположение характеристик в факторном пространстве определяется тремя факторами. При этом каждый из факторов является обобщающим показателем, составленным на основе всех имеющихся в нашем распоряжении характеристик. Представляет интерес определить смысловое содержание этого фактора. Программа «Statistica» позволяет определить величину вклада каждого из исходных объектов (регионов) в выделенный обобщающий фактор. Это так называемый весовой коэффициент или коэффициент вклада. Анализ первого фактора показал, что наибольший отрицательный вклад в него вносят малонаселенные регионы, расположенные на севере и северо-востоке Российской Федерации, такие как Ямало-Ненецкий автономный округ (коэффициент вклада -3.81), Магаданская область (-1.3), республика Саха (Якутия) (-1.2), Тюменская область (-2.46). Положительный вклад, напротив, вносят регионы, расположенные на юге, такие как республика Ингушетия (2.07), Калмыкия (1.04), Дагестан (1.20). В свою очередь, согласно рис. 2, высокие значения первого фактора свойственны для характеристик, определяющих промышленную развитость региона (ВВП, доля промышленной продукции в ВВП, объем промышленной продукции региона), а также для показателей экономического благополучия населения (средние доход, зарплата, пенсия). Низкие значения первого фактора отмечены для таких характеристик, как плотность населения, смертность, уровень безработицы, доля населения с доходом ниже прожиточного минимума. Таким образом, можно утверждать, что для первой группы регионов (малонаселенные регионы северной части России) наблюдается высокий уровень промышленной развитости, сопровождаемый финансовым благополучием населения. Для южных регионов характерна высокая плотность населения, более низкий уровень доходов и неблагоприятная медико-демографическая ситуация. Из результатов аналогичного анализа для второго фактора следует, что регионы европейской части России (Ленинградская, Тульская, Ивановская, Костромская области) характеризуются высокой смертностью от новообразований, заболеваний системы кровообращения и смертностью от внешних причин. Для многих

периферийных регионов (Дагестан, Калмыкия, Карачаево-Черкесия, Ямало-Ненецкий автономный округ) отмечены высокие рождаемость и продолжительность жизни, но вместе с тем и высокий уровень младенческой смертности. Анализ третьего фактора позволил определить, что ряд регионов (Республика Тыва, Корякский автономный округ, Республика Алтай, Амурская область) отличаются неблагоприятной социальной обстановкой, которая выражается в высоком уровне преступности, смертности от внешних причин, значительном фоне инфекционных заболеваний. Другие же регионы (Республика Татарстан, Московская, Липецкая, Самарская области) характеризуются достаточно высокой продолжительностью жизни и значительной плотностью населения.

Проведя процедуру факторного анализа в варианте, когда объектами являются не характеристики регионов, а непосредственно регионы Российской Федерации, можно получить распределение этих регионов в факторном пространстве уже иных факторов, изображенное на рис. 4. Результат кластеризации полученного распределения представлен на рис. 5. Так же, как в предыдущем случае, на рисунках изображена только часть исследуемых регионов. Регионы пронумерованы в соответствии с дендрограммой рис. 5.

Как и ранее, были рассчитаны весовые коэффициенты, позволяющие определить, какие характеристики дают наибольший вклад в тот или иной фактор. Расчет показал, что наибольший отрицательный вклад в первый фактор вносит площадь региона (-6.18) и показатели материальной обеспеченности населения, такие как средняя заработная плата (-0.46) и средний доход (-0.34). Примечательно, что наибольший вклад в значение второго фактора вносят те же характеристики, с той лишь разницей, что показатели материальной обеспеченности характеризуются отрицательными значениями коэффициента вклада (-4.60 , -3.25), а площадь – положительным значением (0.87).

Отсюда следует, что регионы, находящиеся в пространстве факторов правее или выше других, характеризуются более низким уровнем материальной обеспеченности населения (Ингушетия, Адыгея, Северная Осетия – Алания), а регионы, расположенные в средней части распределения, – более высоким.

Как и раньше, представляет интерес определить, чем обуславливается общность регионов, входящих в тот или иной кластер или, другими словами, по какому параметру отличаются кластеры I–IV.

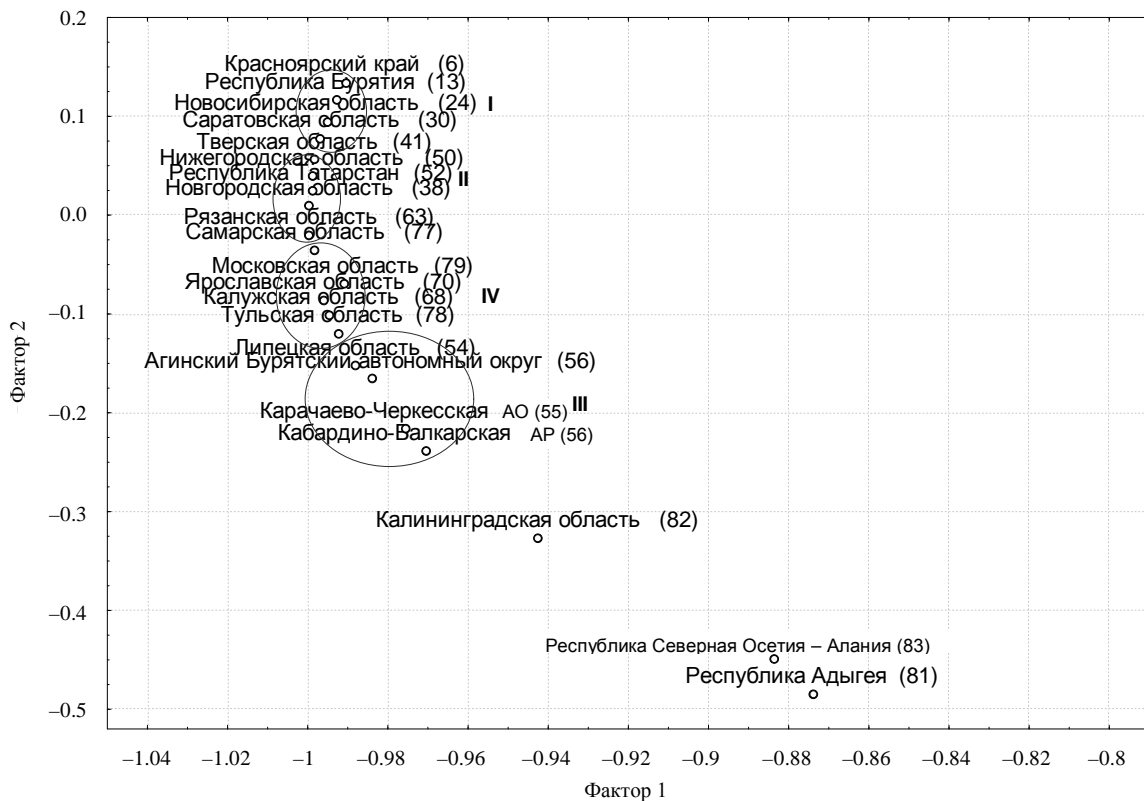


Рис. 4. Распределение регионов Российской Федерации в двумерном факторном пространстве

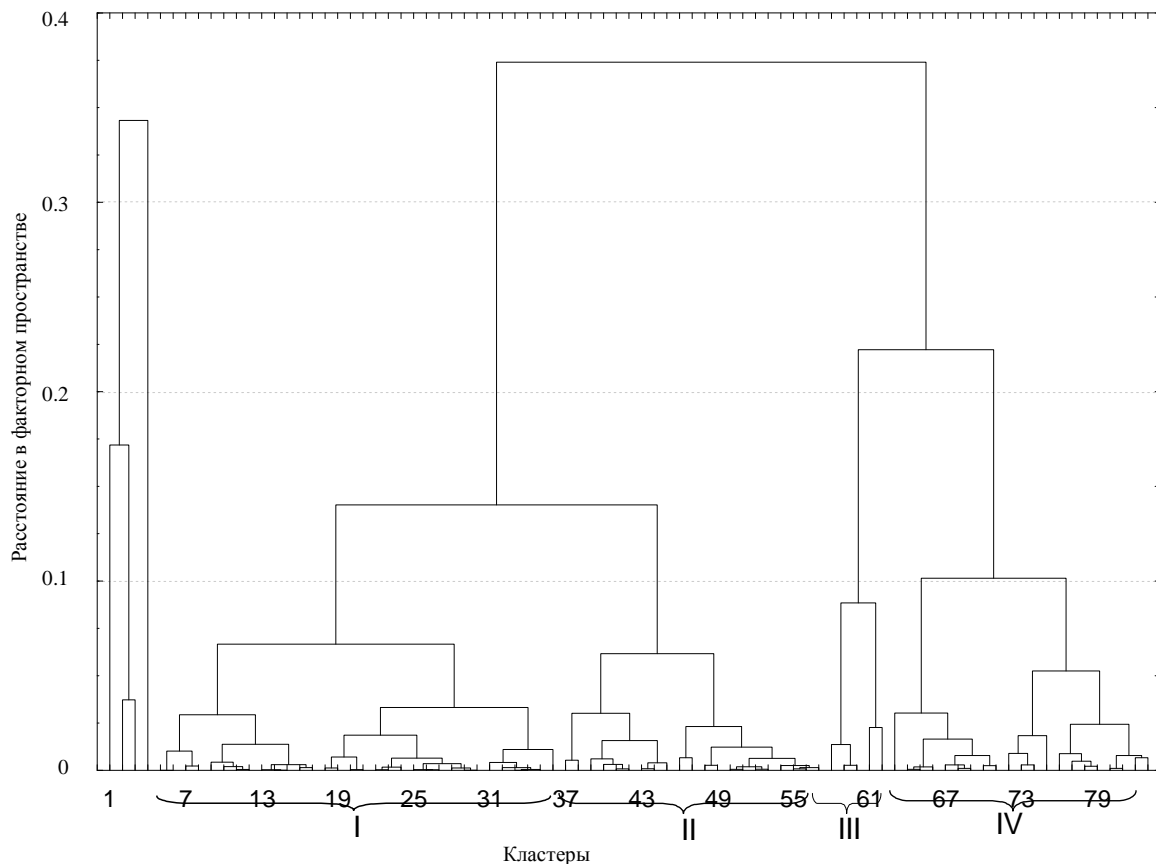


Рис. 5. Дендрограмма кластеризации регионов Российской Федерации в пространстве факторов

Из предыдущего анализа факторов 1 и 2 следует, что главенствующую роль в том и другом играет площадь, а следовательно, и зависящая от площади плотность населения. Ранее [3, 4] мы уже отмечали особую роль плотности населения в демографических процессах. Представляет интерес проанализировать подобную взаимосвязь и в данном случае. Для этого проведем сравнение состава кластеров, в пространстве факторов (рис. 5) с результатами кластеризации по величине плотности населения.

Сравнение проводилось на основе индекса сходства Сьёренсена:

$$I_s = \frac{2 \cdot A}{B + C},$$

где A – число регионов, входящих в обе сравниваемые группы, B и C – общее число регионов в каждой из сравниваемых групп.

Расчет индексов сходства, приведенных в таблице, показал, что, по меньшей мере, для трех кластеров, полученных в результате применения сравниваемых методик, обнаружено значительное сходство по входящим в них объемам.

Таблица

Значения индексов сходства Сьёренсена для двух вариантов кластеризации

		Кластеры по величине плотности			
		1	2	3	4
Кластеры в факторном пространстве	1	0	0.105263	0.36	0.666667
	2	0.121212	0.304348	0.410256	0.163265
	3	0.235294	0.133333	0	0.060606
	4	0.242424	0.565217	0.051282	0.122449

Отмеченное сходство двух вариантов кластеризации, из которых один основан на анализе обобщающих факторов, по нашему мнению,

подтверждает особую роль плотности населения, как одной из основных региональных характеристик, имеющей интегральный характер и определяющей многие другие показатели.

В целом проведенное исследование продемонстрировало возможность и достаточную эффективность факторного анализа для определения взаимосвязей между региональными характеристиками. В частности было показано, что важнейшие медико-демографические показатели, определяющие динамику численности населения в регионе, зависят от уровня экономического благополучия населения и развития системы социальной защиты. В свою очередь, уровень жизни людей повышается при улучшении экономической ситуации, увеличении объемов промышленного производства и индустриальном развитии региона. Следует отметить, что среди наиболее значимых характеристик, определяющих особенности региона, выделяются именно показатели материального благополучия населения. Проведенный анализ позволяет также сделать вывод, что плотность населения является важным интегральным фактором, определяющим характерные особенности изучаемого региона.

Список литературы

1. Дубров А.М., Мхитарян В.С., Трошин Л.И. Многомерные статистические методы. М.: Финансы и статистика, 1998. 352 с.
2. Боровиков В.П. STATISTICA: искусство анализа данных на компьютере. СПб.: Питер, 2001. 656 с.
3. Моничев А.Я., Афонин В.М. Особенности глобального математического прогнозирования с учётом зависимости рождаемости и смертности от численности населения планеты // Вестник Нижегородского университета им. Лобачевского, серия Биология. 2004. Вып. 3(5). С. 92–96.
4. Моничев А.Я., Афонин В.М., Розенберг Г.С., Гелашвили Д.Б. Прогнозирование численности населения на основе модельного представления динамики рождаемости и смертности // Поволжский экологический журн. 2003. № 3. С. 232–238.

INTERRELATIONS OF ECONOMIC, SOCIAL, DEMOGRAPHIC AND ECOLOGICAL CHARACTERISTICS OF RUSSIAN FEDERATION REGIONS STUDIED BY FACTOR ANALYSIS METHODS

S.V. Baryshnikov, A.Ya. Monichev, V.A. Basurov

Interrelations between different groups of regional characteristics have been found using factor analysis. Most significant characteristics determining essential features of Russian Federation regions have been singled out.