

МНОГОМЕРНЫЙ АНАЛИЗ В ИССЛЕДОВАНИИ РЕГИОНАЛЬНОЙ ДИФФЕРЕНЦИАЦИИ

П.В. Герасенко

Многомерный анализ региональной дифференциации основан на дисперсионном, регрессионном и кластерном анализе, которые позволяют получить синергический эффект достоверности прогнозной оценки уровня жизни и оптимизации варианта стратегии устойчивого развития региона, подтверждают или опровергают полученные при применении только одного метода выводы, обосновывают достоверность результатов исследования.

Для проведения многомерного анализа в исследовании региональной дифференциации нами были составлены матрицы: показателей уровня жизни населения и экономического развития регионов, включающие статистические и расчетные данные. Матрица показателей уровня жизни населения по регионам Гомельской области отражает различные аспекты и включает следующие показатели: номинальная среднемесячная зарплата, тыс. руб. / месяц; обеспеченность благоустроенным жильем жителя области, кв.м на человека; обеспеченность благоустроенным жильем жителя города, кв.м на человека; обеспеченность благоустроенным жильем жителя села, кв.м на человека; численность студентов на 10 тыс. населения, человек; интегральный уровень образования; показатель чистоты воздуха; средняя плотность загрязнения, $ku / км^2$; первичная заболеваемость всего населения (случаев на 100 тыс. населения); первичная заболеваемость детей до 14 лет (случаев на 100 тыс. населения); обеспеченность медицинскими кадрами (число врачей и среднего медицинского персонала), на 10 тыс. населения; посещение театров на тыс. чел., раз; посещение музеев на тыс. чел., раз; количество преступлений в районе на 10 тыс. человек.

Матрица показателей экономического развития регионов Гомельской области предназначена для отображения экономического развития и включает в себя следующие показатели: экспорт товаров на работника, тыс. долл.; импорт товаров на работника, тыс. долл.; сальдо на работника, тыс. долл.; x_5 – прибыль (убыток) предприятий и организаций на работника,

тыс. руб.; x_6 – платные услуги населению тыс. руб. / чел. и др.

Так как уровень жизни населения характеризуется финансовой устойчивостью регионов, нами для ее оценки предлагается использовать показатель рентабельности. Известно, что на уровень рентабельности существенное влияние оказывают как экономические, так и социальные факторы. Поэтому нами при проведении *регрессионного анализа* финансовой устойчивости регионов наряду с экономическими использовались и социальные факторы.

Постановка задачи. В основу данного исследования положена следующая научная задача: существует некая статистически значимая связь между показателями системы устойчивого развития региона. Установление характера связи и степени ее влияния на рост уровня жизни населения позволит выявить особенности функционирования системы региональной экономики и на их основе предсказать будущее поведение объекта при изменении каких-либо параметров и предложить корректировку курса региональной экономической политики.

В работе используются логический анализ изучаемых показателей, сравнение статистических характеристик (коэффициенты множественной корреляции и детерминации) на основе статистических моделей множественной регрессии. В подобных моделях значения экономических переменных определяются обычно влиянием не одного, а нескольких факторов. Для искомых региональных уравнений регрессии прямолинейного вида будем использовать уравнение (1):

$$y = a_0 + a_1x_1 + a_2x_2 + \dots + a_nx_n, \quad (1)$$

где: y – функция (анализируемый показатель); a_0 – свободный член уравнения (в отдельных случаях при экономическом анализе имеет логический смысл только регрессионное уравнение, в котором $a_0 = 0$); x_1, x_2, \dots, x_n – факторные признаки, определяющие результативный показатель; a_1, a_2, \dots, a_n – коэффициенты регрессии при факторных признаках, характеризующие уровень влияния каждого фактора на результативный показатель в его единицах измерения.

Тесноту связи между анализируемым показателем и всеми факторами, включенными в модель, определяет *множественный коэффициент корреляции* ($R_{y/x}$). По практическим соображениям выделим следующие диапазоны тесноты связи между переменными в зависимости от величины коэффициента корреляции: при $R_{y/x} < 0,3$ – слабая степень тесноты связи, при $R_{y/x} = 0,31 - 0,5$ – умеренная, при $R_{y/x} = 0,51 - 0,7$ – заметная и при $R_{y/x} > 0,7$ – высокая степень тесноты связи.

Квадрат коэффициента корреляции называется *коэффициентом детерминации* [R – квадрат (R^2)]. Он показывает, на сколько процентов вариация результативного показателя зависит от избранных факторов.

При регрессионном анализе данных, отражающих уровень устойчивости регионального развития, нами выделено восемь базовых статистических задач. Наиболее важные статистические характеристики по выделенным задачам приведены в таблице 1.

Таблица 1
Статистические характеристики уравнений регрессии по регионам Гомельской области

№ задачи	Изучаемый фактор	Факторные признаки	Коэффициенты		Характеристики тесноты связи	Влияние факторов на выбранный показатель	
			множественный R	детерминации (R^2)		на увеличение	на уменьшение
1	Уровень рентабельности (p_1)	x_1-x_4	0,617	0,381	заметная	x_4, x_2, x_1	
2	Уровень рентабельности (p_2)	x_1, x_2, x_6	0,690	0,476	заметная	x_6, x_5	
3	Преступность количество преступлений на 10 тыс. населения (x_{14})	$x_{11}, x_5, x_6, x_{12}, x_{13}$	0,598	0,358	заметная		x_6
4	Уровень рентабельности (p_3)	$x_1, x_2, x_6, x_8, x_{10}$	0,697	0,486	заметная		x_8
5	Уровень рентабельности (p_4)	x_1, x_2, x_5, x_8, x_9	0,695	0,484	заметная	x_6	x_9
6	Экспорт товаров на работника, тыс. долл. (x_2)	x_6-x_8	0,602;	0,362	заметная	x_6	
7	Уровень рентабельности (p_5)	x_4-x_8	0,733	0,537	высокая	x_6	
8	Номинальная среднемесячная зарплата работников (x_1)	x_2, x_5, x_7-x_{11}	0,488	0,238	умеренная	x_2	x_8

По уравнению регрессии, отражающему результаты решения задачи 7 получен множественный коэффициент корреляции ($R \geq 0,70$), который показывает высокую тесноту связи между изучаемым фактором и факторными признаками. При этом коэффициент детерминации (R^2) составил 0,537. Следовательно, изучаемый фактор объясняет зависимость от факторных признаков на 53,7%. По задачам 1–6 получены более низкие коэффициенты корреляции (0,598–0,697) и более низкие коэффициенты детерминации (0,362–0,486), что характеризует тесноту связи в полученном корреляционном уравнении как заметную. В задаче 8 получен низкий коэффициент корреляции и соответственно низкий коэффициент детерминации. Следовательно, в данной задаче тесноту связей надобно оценить как слабую. Однако в задаче 8 полученное корреляционное уравнение может быть использовано для выявления ориентированных прогнозных ошибок.

Из всех факторных признаков в семи задачах, приведенных в таблице 1, на увеличение выбранного показателя чаще всего оказывает влияние факторный признак x_6 – платные услуги населению, тыс. руб. / чел. (в четырех задачах из шести – 66,7%) и только один раз (16,7%) – на уменьшение.

В двух задачах (66,7% от количества решенных) на уменьшение выбранного показателя оказывает влияние факторный признак x_8 – средняя плотность загрязнения, ku / km^2 .

Дисперсионный анализ дает возможность установить, существенное ли влияние оказывает тот или иной из рассматриваемых факторов или их взаимодействие на изменчивость изучаемого параметра, а также вычислить удельный вес каждого источника изменчивости в общей совокупности.

Постановка задачи. В практической деятельности, связанной с проведением различных региональных исследований, мы часто сталкиваемся с необходимостью выявить и оценить влияние отдельных факторов

на изменчивость какого-либо изучаемого параметра.

Результаты регионального дисперсионного анализа. По традиционной схеме проведения однофакторного дисперсионного анализа выполнены конкретные исследования,

позволяющие ответить на вопрос: существенно ли влияние региональных особенностей на изменение изучаемого экономического или социального параметра по регионам Гомельской области. Полученные результаты показаны в таблице 2.

Таблица 2

Сводные результаты дисперсионного анализа изучаемых региональных экономических и социальных параметров по регионам Гомельской области за 2005 г.

№ задач	Изучаемые региональные экономические параметры	Численное значение F-критерия Фишера при 5%-ном уровне значимости		Влияние регионального фактора (специфики регионов)
		Наблюдаемые значения (F набл.)	Табличные значения (F табл.)	
1	Уровень рентабельности (p_1)	2,621	1,000	Имеется, сильное
2	Уровень рентабельности (p_2)	3,018	1,000	Имеется, сильное
3	Преступность количество преступлений на 10 тыс. населения (x_{14})	2,385	1,000	Имеется, сильное
4	Уровень рентабельности (p_3)	2,385	1,000	Имеется, сильное
5	Уровень рентабельности (p_4)	2,385	1,000	Имеется, сильное
6	Экспорт товаров на работника, тыс. долл. (x_2)	3,018	1,000	Имеется, сильное
7	Уровень рентабельности (p_5)	2,385	1,000	Имеется, сильное
8	Номинальная среднемесячная зарплата работников (x_1)	2,108	1,000	Имеется, сильное
9	Города и районы Гомельской области в целом по двенадцати показателям уровня жизни (двадцать четыре задачи)	1,952	1,913	Имеется, сильное
10	Города и районы Гомельской области по каждому показателю уровня жизни (двенадцать задач)	1,620	1,500	Имеется, сильное

При этом номера задач дисперсионного и выполненного выше регрессионного анализа совпадают, так как эти задачи решались синхронно. Кроме того, дисперсионный анализ выполнен по всей Гомельской области по двенадцати показателям уровня жизни населения за шесть лет и отдельно по каждому показателю уровня жизни населения в разрезе городов и районов.

Полученные результаты регионального дисперсионного анализа позволяют сформулировать следующие выводы, представляющие теоретический и практический интерес:

1. В результате изучения социально-экономической природы влияния регионального фактора на исследуемые региональные социально-экономические процессы создается возможность объяснить, почему в одних локальных задачах влияние регионального фактора на изучаемый параметр проявляется более сильно, а в других – ослабевает.

2. Выявлено сильное влияние регионального фактора при изучении параметров всех решенных задач.

Постановка задачи. Расширение масштаба проводимого многомерного анализа, имеющего целью более глубокое изучение комбинаций факторных признаков и в итоге выделение однородных региональных зон, может быть реализовано с применением кластерного анализа [1]. Наиболее существенные методологические черты кластерного анализа: образование единой меры, охватывающей ряд характеризующих признаков, и группировка объектов наблюдения.

В кластерном региональном анализе группировочные признаки подвергаются объединению с помощью функции евклидова расстояния.

Последовательный процесс кластеризации начинается с рассмотрения n районов, затем два наименее сходных района объединяются в один кластер, и число кластеров становится

равным $n - 1$. Процесс повторяется до тех пор, пока все n районов не попадут в один кластер.

Для представления результатов решения задачи регионального кластерного анализа будем использовать «дендограмму» как графическое изображение результатов процесса последовательной кластеризации, который осуществляется в терминах матрицы расстояний или сходства.

Результаты регионального кластерного анализа. Социально-экономическое зонирование в Гомельской области по выделенным показателям уровня жизни населения выполнено с использованием SPSS Version 10 (анализ статистических данных и восстановление скрытых закономерностей) [2, с.384–417] регионального кластерного анализа. При этом использованы следующие показатели, характеризующие уровень жизни населения: – номинальная среднемесячная зарплата, тыс. руб. / месяц; обеспеченность благоустроенным жильем одного жителя, кв.м на человека; численность студентов на 10 тыс.

населения, человек; интегральный уровень образования; показатель чистоты воздуха (единицу делим на выбросы и умножаем на 1 млн); средняя плотность загрязнения, $ku / км^2$; первичная заболеваемость всего населения (случаев на 100 тыс. населения); первичная заболеваемость детей до 14 лет (случаев на 100 тыс. детей); обеспеченность медицинскими кадрами (суммарное число врачей и среднего медицинского персонала на 10 тыс. чел.); посещение театров на тыс. чел. населения, раз; посещение музеев на тыс. чел. населения, раз; количество преступлений в районе на 10 тыс. населения.

Полученные в ходе исследования результаты по ранжированию районов (городов) в Гомельской области приведены в таблице 3. Определяем группировки регионов Гомельской области. Например, Рогачевский и Чечерский районы имеют одинаковый первый номер приоритета, Октябрьский – второй, Кормянский район – третий, Брагинский и Мозырский районы имеют одинаковый – шестой номер приоритета.

Таблица 3

Распределение районов и городов Гомельской области по приоритетам кластерного анализа за 2005 г.

Наименование района (региона)	Номера показателей и присвоенные им ранги												Сумма рангов	Приоритеты группировки регионов
	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉	X ₁₀	X ₁₁	X ₁₂		
Рогачевский	2	17	4	4	9	7	16	18	2	4	19	6	108	1
Чечерский	7	4	2	9	21	11	5	16	12	2	12	7	108	1
Октябрьский	12	10	9	7	16	4	10	7	3	9	17	15	119	2
Кормянский	16	9	12	11	19	12	1	1	17	12	1	10	121	3
Ельский	8	6	16	10	11	8	12	6	1	16	8	20	122	4
Петриковский	14	7	6	15	8	5	17	19	6	6	7	13	123	5
Брагинский	13	11	3	21	22	10	11	3	7	3	6	18	128	6
Мозырский	18	5	11	13	1	3	19	17	23	11	2	5	128	6
Лоевский	15	3	10	14	18	13	2	4	18	10	9	14	130	7
Наровлянский	3	16	8	20	23	15	6	2	13	8	16	2	132	8
Хойникский	10	23	1	8	13	9	18	10	20	1	10	16	139	9
Ветковский	6	1	21	17	20	14	9	9	8	21	13	1	140	10
Калинковичский	5	15	15	2	6	20	13	21	10	15	15	3	140	10
Добрушский	1	13	19	16	14	6	4	13	15	19	5	17	142	11
Буда-Кошелевский	9	8	20	19	10	21	3	8	14	20	3	8	143	12
Лельчицкий	4	12	13	18	17	18	23	5	5	13	11	12	151	13
Житковичский	11	2	17	12	5	16	21	14	4	17	14	19	152	14
Речицкий	19	22	7	5	3	22	7	15	11	7	18	21	157	15
Гомельский	17	19	18	1	12	19	15	20	19	18	4	4	166	16
Светлогорский	21	14	5	6	7	23	14	22	9	5	21	22	169	17
Жлобинский	23	18	14	3	15	17	20	11	16	14	20	11	182	18
г. Гомель	20	21	23	23	2	2	8	12	22	23	23	9	188	19
г. Мозырь	22	20	22	22	4	1	22	23	21	22	22	23	224	20
Сумма рангов	276	276	276	276	276	276	276	276	276	276	276	276		

Распределение показателей уровня жизни населения Гомельской области по приоритетам кластерного анализа за 2005 г. приведено в таблице 4.

**Распределение показателей уровня жизни населения Гомельской области
по приоритетам кластерного анализа за 2005 г.**

Наименование района (региона)	Номера показателей и присвоенные им ранги												Сумма рангов
	x ₁	x ₂	x ₃	x ₄	x ₅	x ₆	x ₇	x ₈	x ₉	x ₁₀	x ₁₁	x ₁₂	
Брагинский	9	4	1	6	10	5	11	12	7	2	3	8	78
Буда-Кошелевский	9	5	3	6	10	4	11	12	7	1	2	8	78
Ветковский	9	3	1	5	10	4	11	12	7	2	6	8	78
Гомельский	8	5	3	6	10	4	11	12	7	1	2	9	78
Добрушский	9	4	1	6	10	5	11	12	7	2	3	8	78
Ельский	9	4	1	6	10	5	11	12	7	2	3	8	78
Житковичский	9	6	1	4	10	3	11	12	7	2	5	8	78
Жлобинский	9	4	1	5	10	3	11	12	7	2	6	8	78
Калинковичский	9	4	1	5	10	3	11	12	7	2	6	8	78
Кормянский	7	4	3	6	10	5	11	12	9	1	2	8	78
Лельчицкий	9	5	1	6	10	3	12	11	7	2	4	8	78
Лоевский	9	3	1	6	10	4	11	12	7	2	5	8	78
Мозырский	8	2	3	6	7	4	11	12	10	1	2	9	75
Наровлянский	9	3	1	4	10	5	11	12	7	2	6	8	78
Октябрьский	9	4	3	5	10	1	11	12	7	2	6	8	78
Петриковский	9	5	3	6	10	1	11	12	7	2	4	8	78
Речицкий	10	4	1	5	7	3	11	12	8	2	6	9	78
Рогачевский	9	3	1	4	10	5	11	12	7	2	6	8	78
Светлогорский	9	4	1	5	10	3	11	12	7	2	6	8	78
Хойникский	9	4	1	5	10	6	11	12	7	2	3	8	78
Чечерский	9	3	1	5	10	4	11	12	7	2	6	8	78
г. Гомель	8	1	6	2	7	3	11	12	10	9	4	5	78
г. Мозырь	11	5	7	6	12	3	4	1	10	8	9	2	78
Сумма рангов	205	89	46	120	223	86	247	264	173	55	105	178	
Приоритеты группировки регионов	9	4	1	6	10	3	11	12	7	2	5	8	

По аналогии с предыдущей задачей социально-экономического зонирования определяем группировку показателей. Например, показатель численность студентов на 10 тыс. населения, имеет первый номер приоритета; показатель посещение театров на тыс. чел. населения, (раз) – второй номер; показатель средняя плотность загрязнения ($K_u / \text{км}^2$) – третий номер и т. д. Результаты решения приведенных задач с использованием кластерного анализа представляют интерес для принятия управленческих решений на уровне Гомельского облисполкома.

Социально-экономическое зонирование в Гомельской области. В задаче по показателю обеспеченность благоустроенным жильем одного жителя, кв.м на человека рассматриваются возможные варианты объединения регионов Гомельской области в более крупные территориальные образования с целью более глубокого социально-экономического анализа для повышения обоснованности принимаемых региональных управленческих решений. Задача социально-экономического зонирования также решена на компьютере с

использованием пакета SPSS Version 10 (анализ статистических данных и восстановление скрытых закономерностей).

Полученный количественный результат решения задачи по показателю обеспеченность благоустроенным жильем одного жителя приведен в таблице 5. (Cluster Membership). В данной таблице отражены три наиболее предпочтительных варианта объединения кластеров по условию сходства исследуемого показателя. Например, при выделении в Гомельской области четырех регионов (4 Clusters) согласно направленности решаемой задачи, целесообразно объединить административно-территориальные образования в следующие регионы: первый регион – Брагинский, Буда-Кошелевский, Ветковский, Ельский, Житковичский, Кормянский, Лельчицкий, Лоевский, Мозырский, Октябрьский, Петриковский и Чечерский районы; второй – Гомельский, Жлобинский, Наровлянский и Рогачевский; третий – Добрушский, Калинковичский, Светлогорский; четвертый регион – Речицкий и Хойникский районы, г. Мозырь и Гомель.

Варианты социально-экономического зонирования в Гомельской области по показателю обеспеченности благоустроенным жильем одного жителя, кв.м на человека (Cluster Membership)

Регионы области	Интегральные экономические зоны		
	4 зоны	3 зоны	2 зоны
Брагинский	1	1	1
Буда-Кошелевский	1	1	1
Ветковский	1	1	1
Гомельский	2	2	1
Добрушский	3	2	1
Ельский	1	1	1
Житковичский	1	1	1
Жлобинский	2	2	1
Калинковичский	3	2	1
Кормянский	1	1	1
Лельчицкий	1	1	1
Лоевский	1	1	1
Мозырский	1	1	1
Наровлянский	2	2	1
Октябрьский	1	1	1
Петриковский	1	1	1
Речицкий	4	3	2
Рогачевский	2	2	1
Светлогорский	3	2	1
Хойникский	4	3	2
Чечерский	1	1	1
г. Мозырь	4	3	2
Гомель	4	3	2

Полученные результаты показывают возможность выделения следующих приоритетных вариантов выделения социально-экономических зон в Гомельской области:

- при выделении двух социально-экономических зон к первой зоне следует отнести все административные образования Гомельской области, за исключением Речицкого и Хойникского районов, г. Мозыря и Гомеля;
- при выделении трех социально-экономических зон к первой зоне следует отнести Брагинский, Буда-Кошелевский, Ветковский, Ельский, Житковичский, Кормянский, Лельчицкий, Лоевский, Мозырский, Октябрьский, Петриковский и Чечерский районы; ко второй – Гомельский, Добрушский, Жлобинский, Калинковичский, Наровлянский, Рогачевский и Светлогорский районы; к третьей социально-экономической зоне – Речицкий и Хойникский районы, г. Мозырь и Гомель;
- при выделении четырех социально-экономических зон к первой зоне следует

отнести Брагинский, Буда-Кошелевский, Ветковский, Ельский, Житковичский, Кормянский, Лельчицкий, Лоевский, Мозырский, Октябрьский, Петриковский и Чечерский районы; ко второй – Гомельский, Жлобинский, Наровлянский и Рогачевский районы; к третьей – Добрушский, Калинковичский и Светлогорский районы; к четвертой социально-экономической зоне – Речицкий и Хойникский районы, г. Мозырь и Гомель.

Итак, в статье показана возможность использования кластерного анализа для решения задач по социально-экономическому зонированию. В качестве изучаемых регионов могут быть приняты административно-территориальные районы Гомельской области. Выделены приоритеты объединения регионов в более крупные социально-экономические зоны по показателям уровня жизни населения. Получены конкретные рациональные варианты выделения в Гомельской области двух, трех и четырех социально-экономических зон.

ЛИТЕРАТУРА

1. Герасенко, В.П., Герасенко, П.В. Кластерное зонирование по факторам развития человеческого потенциала // Вопросы статистики. Научно-информационный журнал. М., 2006. №9. С. 61–66.
2. Брююль, А., Цефель, П. SPSS: искусство обработки информации. Анализ статистических данных и восстановление скрытых закономерностей: Пер. с нем. СПб.: ООО «ДиасофтЮП», 2001. С. 268–284.

РЕЗЮМЕ

Показана возможность использования многомерного анализа в исследовании уровня жизни населения по регионам Гомельской области. Многомерный анализ в исследовании региональной дифференциации позволяет получить достоверную прогнозную оценку повышения уровня жизни и разработать оптимальный вариант стратегии устойчивого развития региона.

SUMMARY

The possibility of using the multivariate analysis in studying living standards in Gomel region is shown. The multivariate analysis in regional differentiation research allows to obtain reliable prognosis estimation of the living standard and work out optimum strategy of steady development of the region.

* Статья поступила в редакцию 27 ноября 2005 г.