

УДК 338.431:519.237.8

ИСПОЛЬЗОВАНИЕ КЛАСТЕРНОГО АНАЛИЗА В АНАЛИЗЕ ЭКОНОМИЧЕСКИХ ПРОЦЕССОВ СЕЛЬСКОХОЗЯЙСТВЕННОГО ПРОИЗВОДСТВА

Смагин Борис Игнатьевич

доктор экономических наук, профессор

bismagin@mail.ru

Мичуринский государственный аграрный университет

г. Мичуринск, Россия

Аннотация. Необходимым условием корректного анализа статистической совокупности является ее однородность, для формирования которой используются методы кластерного анализа. В статье приведено обоснование факторных признаков, лежащих в основе проведения кластерного анализа для построения производственных функций. Приведен пример проведения данной процедуры и вычисление зависимостей между величиной затраченных ресурсов и объемом производимой продукции в сельскохозяйственных организациях региона.

Ключевые слова: сельскохозяйственное производство, однородная совокупность, кластерный анализ, стандартизация, производственная функция.

В эконометрических исследованиях сельскохозяйственного производства чаще всего мы имеем только одно наблюдение за год, что не позволяет построить математическую модель по конкретному предприятию. Следовательно, необходимо иметь значительный массив данных по сельскохозяйственным предприятиям региона, а корректную обработку статистических данных можно осуществлять только для однородной группы наблюдений. Выбор однородных переменных необходим потому, что средние величины только тогда имеют смысл, когда они основаны на качественно однородном материале. В противном случае средние величины не отражают характерные черты определенной совокупности, поскольку данные, по которым они рассчитаны, принадлежат к разным совокупностям [1].

Данные же по сельскохозяйственным организациям региона отличаются своей неоднородностью, поэтому для адекватного анализа необходимо формировать соответствующие группы с помощью методов кластерного анализа.

Операцией, предшествующей проведению кластерного анализа, является стандартизация всех переменных. Эта процедура необходима, так как все признаки должны быть приведены к сопоставимому виду путем исключения единиц измерения. При этом в кластерном анализе необходима такая стандартизация, при которой среднее значение каждого факторного признака равно нулю, а дисперсия равна единице [2, 3].

Добиться этого можно, осуществляя процесс стандартизации по формулам:

$$z_{ik} = \frac{x_{ik} - \bar{x}_k}{s_k},$$

где x_{ik} – значение признака k для i -го объекта; \bar{x}_k – среднее арифметическое значение признака k ; s_k – стандартное отклонение признака k .

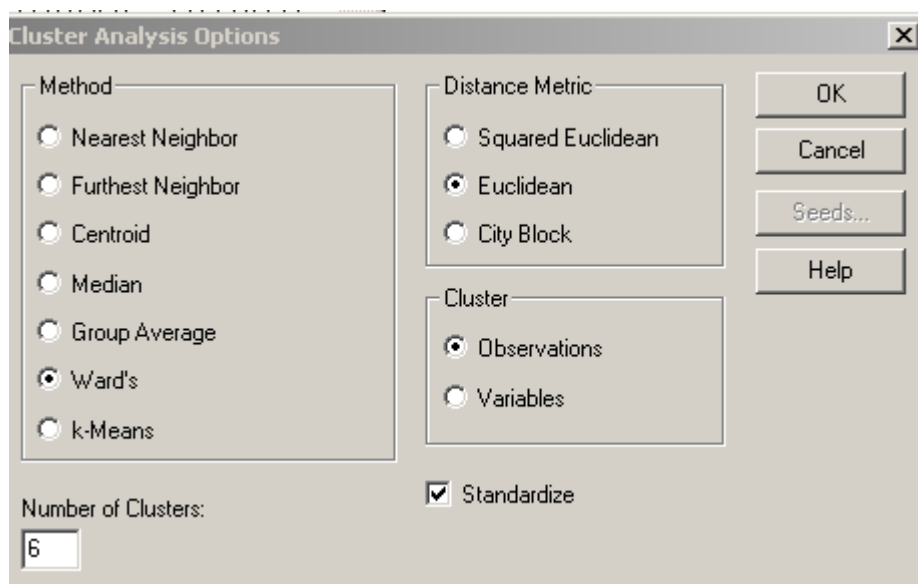


Рисунок 1 - Пример заполнения окна диалога для выбора параметров в программе Statgraphics

Нами на основе использования программы Statgraphics был проведен кластерный анализ при построении производственных функций для сельскохозяйственных организаций Тамбовской области за 2018 год. Так как однородность должна быть обеспечена по тем факторам, которые определяют результативный показатель (а в данном случае это объем валовой продукции), то в качестве кластерных признаков следует использовать объемы и интенсивность использования ресурсов, сбалансированных с учетом специализации предприятия. Тем самым в основу кластеризации нами были положены следующие факторы: x_1 – среднегодовая численность работников, чел.; x_2 – площадь сельскохозяйственных угодий, га; x_3 – среднегодовая стоимость основных производственных фондов, тыс. руб.; x_4 – среднегодовая стоимость производственных оборотных средств, тыс. руб., а также затраты ресурсов в расчете на 100 га сельскохозяйственных угодий: q_1 – количество работников в расчете на 100 га сельскохозяйственных угодий, чел./100га; q_2 – стоимость основных производственных фондов в расчете на 100 га сельскохозяйственных угодий, тыс. руб./100га; q_3 – стоимость производственных оборотных средств в расчете на 100 га сельскохозяйственных угодий, тыс. руб./100га.

Для оценки специализации сельскохозяйственных организаций в основу кластеризации следует положить структуру товарной продукции, %:

z_1 – удельный вес зерна; z_2 – удельный вес масличных культур; z_3 – удельный вес сахарной свеклы; z_4 – удельный вес плодов и ягод; z_5 – удельный вес овощей; z_6 – удельный вес картофеля; z_7 – удельный вес мяса крупного рогатого скота; z_8 – удельный вес мяса свиней; z_9 – удельный вес молока; z_{10} – удельный вес продукции овцеводства.

Для кластерного анализа мы использовали метод Уорда, расстояние между факторными признаками и кластерами измерялось по евклидовой метрике. В результате из 245 сельскохозяйственных организаций 239 были разбиты на 4 репрезентативных кластера; оставшиеся 6 предприятий вошли в непрезентативные кластеры и дальнейший анализ по ним не проводился.

Таблица 1

Результаты кластерного анализа сельскохозяйственных организаций Тамбовской области

Показатели (в среднем на 1 хозяйство)	Кластеры			
	1	2	3	4
Число предприятий	64	77	76	22
Площадь сельскохозяйственных угодий, га	3288	4190	9702	2305
Среднегодовое количество работников, человек	22,7	32,0	99,7	85
Среднегодовая стоимость основных производственных фондов, тыс. руб.	46050	77910	226849	341572
Среднегодовая стоимость оборотных средств, тыс. руб.	52271	93082	284863	159535
Валовая продукция, тыс. руб.	47809	63943	208267	204169
Приходится на 100 га сельскохозяйственных угодий				
Работников, человек	0,7	0,8	1,0	3,7
Основных производственных фондов, тыс. руб.	1400	1859	2338	14819
Оборотных средств, тыс. руб.	1590	2222	2936	6921
Удельный вес в структуре товарной продукции, %				
Зерна	49,0	76,0	49,0	7,4

Масличных культур	45,5	15,9	22,1	3,3
Сахарной свеклы	0	0	18,1	4,7
Картофеля	0	0	2,0	4,8
Овощей	0,1	0	0,1	0,3
Фруктов и ягод	0	0,007	0,05	7,5
Мяса КРС	0,2	0,02	0,9	2,5
Мяса свиней	0	0,02	0,1	52,5
Молока	0,1	0,03	0,9	7,5
Продукции овцеводства	0,004	0,02	0,006	0,2

Среди всех выделенных кластеров сельскохозяйственные предприятия, составляющие первый кластер, имеют самую низкую обеспеченность производственными ресурсами и интенсивность их использования. Следствием этого является и наиболее низкий уровень валового производства – всего 47809 тыс. руб. в среднем на одно хозяйство. Основной удельный вес в структуре товарной продукции занимают зерно и масличные культуры (около 95%).

Предприятия второго кластера также характеризуются низкой обеспеченностью производственными ресурсами (за исключением площади сельскохозяйственных угодий) и наиболее низкой (вслед за предприятиями первого кластера) интенсивностью их использования и уровнем валового производства. Основной удельный вес в структуре товарной продукции занимает зерно (76%).

Сельскохозяйственные организации, сформировавшие третий кластер, характеризуются достаточно высоким уровнем обеспеченности ресурсами, в этой группе в расчете на одно хозяйство наибольшая площадь сельскохозяйственных угодий, количество работников и стоимость оборотных средств, следствием чего является наиболее высокий уровень валового производства. Основной удельный вес в структуре товарной продукции занимает зерно (49%) и другая продукция растениеводства (масличные культуры, сахарная свекла и картофель в сумме занимают 42,2%).

Несколько специфическую группу сформировали сельскохозяйственные организации четвертого кластера. Основной удельный вес в структуре товарной продукции занимает производство мяса свиней (52,5%), другая продукция животноводства (мясо крупного рогатого скота, молоко и продукция овцеводства) занимают 10,2%. Кроме того, производится практически вся продукция растениеводства (зерно, масличные культуры, сахарная свекла, картофель, овощи, плоды и ягоды занимают 23% в структуре товарной продукции). В силу отмеченной специфики в хозяйствах этой группы наблюдается самая низкая площадь сельскохозяйственных угодий при достаточно высокой обеспеченности трудовыми ресурсами, основными производственными и оборотными средствами. Здесь также наблюдается самая высокая интенсивность использования ресурсов. По объему валовой продукции занимая второе место, после предприятий третьего кластера, в данной группе самая высокая эффективность использования земельных, трудовых ресурсов и оборотных средств.

Для расчетов мы использовали кинетическую производственную функцию, обладающую большой гибкостью и показавшую свою применимость для анализа сельскохозяйственного производства [4, 5].

В результате реализации алгоритма пошагового регрессионного анализа для сельскохозяйственных организаций первого кластера была получена модель вида:

$$Y = 0,0894 \cdot x_1^{-0,582} \cdot x_2^{1,221} \cdot x_4^{0,476} \cdot e^{0,0271x_1 - 0,0002x_2}$$

Для сельскохозяйственных организаций, формирующих второй кластер, получена следующая зависимость:

$$Y = 0,0205 \cdot x_1^{-0,4004} \cdot x_2^{1,020} \cdot x_4^{0,726} \cdot e^{-0,0001x_2}$$

Для сельскохозяйственных организаций третьего кластера получена модель вида:

$$Y = 0,029 \cdot x_2^{0,890} \cdot x_4^{0,653} \cdot e^{0,0271x_1 - 0,00004x_2}$$

Наконец, для четвертого кластера построенная кинетическая производственная функция имеет вид:

$$Y = 103596 \cdot x_1^{3,5999} \cdot x_3^{-1,3124} \cdot e^{-0,0206x_1 + 0,000004x_3}$$

Все модели построены на значимых факторах, адекватны с уровнем значимости не менее 99,9%. Коэффициент множественной корреляции изменялся в пределах 0,85 – 0,97. Построенные производственные функции для выделенных кластеров имеют существенно различный характер, то есть невозможно сделать общий вывод о взаимосвязях и взаимозависимостях при производстве сельскохозяйственной продукции в регионе: следует анализировать эти закономерности в кластерах, схожих по объемам используемых ресурсов, интенсивности их использования и уровне специализации предприятий.

Список литературы:

1. Дайитбегов, Д.М. Компьютерные технологии анализа данных в эконометрике / Д.М. Дайитбегов. – М.: ИНФРА-М: Вузовский учебник, 2010. – 578с.
2. Джеймс, Г. Введение в статистическое обучение с примерами на языке R / Г. Джеймс, Д. Уиттон, Т. Хастис, Р. Тибширани. – М.: ДМК Пресс, 2016. – 450с.
3. Ланц, Б. Машинное обучение на R: экспертные техники для прогностического анализа / Б. Ланц. – СПб.: Питер, 2020. – 464 с.
4. Смагин, Б.И. Производственные функции в аграрном секторе экономики: монография / Б.И. Смагин. – Мичуринск: Мичуринский ГАУ, 2018. – 99 с.
5. Смагин, Б.И. Логика формирования производственных функций / Б.И. Смагин, А.Б. Смагина // Развитие агропромышленного комплекса: экономика, моделирование и информационное обеспечение. Сборник научных трудов. – Воронеж: Воронежский ГАУ, 2016. – С. 97 – 105.

UDC 338.431:519.237.8

**USE OF CLUSTER ANALYSIS IN THE ANALYSIS OF ECONOMIC
PROCESSES OF AGRICULTURAL PRODUCTION**

Smagin Boris Ignatievich

Doctor of Economic Sciences, Professor

bismagin@mail.ru

Michurinsk State Agrarian University

Michurinsk, Russia

Annotation. A necessary condition for the correct analysis of a statistical population is its homogeneity, for the formation of which the methods of cluster analysis are used. The article provides a justification for the factor characteristics underlying the cluster analysis for the construction of production functions. An example of this procedure and the calculation of the dependencies between the amount of resources spent and the volume of products produced in agricultural organizations in the region is given.

Key words: agricultural production, homogeneous population, cluster analysis, standardization, production function.