

УДК 658.012.1

А. А. Гаврилова

*Харьковский национальный экономический университет, Харьков*

## ИСПОЛЬЗОВАНИЕ ПРИКЛАДНОГО ПАКЕТА STATISTICA ДЛЯ ВЫЯВЛЕНИЯ ОДНОЗНАЧНЫХ ГРУПП В ГОСУДАРСТВЕННОМ КОРПОРАТИВНОМ СЕКТОРЕ ЭКОНОМИКИ УКРАИНЫ

*Обоснована необходимость использования информационной технологии STATISTICA 6.0 для проведения финансового анализа деятельности корпоративных предприятий с государственным пакетом акций. Проведен кластерный анализ исследуемой совокупности по методу иерархической агломеративной кластеризации и методу K-средних. Проведено сравнение полученных результатов, сделаны выводы о взаимосвязи показателей финансовой деятельности с оптимизацией корпоративного портфеля государства.*

**Ключевые слова:** кластеризация, метод Варда, расстояние Чебышева, внутриклассовая дисперсия, евклидово расстояние, Z-счет Альтмана, показатели финансовой деятельности.

### Введение

**Постановка проблемы и анализ литературы.** Современная экономическая наука характеризуется широким использованием математики. Математические методы являются составной частью методов любой экономической науки, включая корпоративное управление. обстоятельный экономический анализ явлений и процессов, происходящих в период трансформационных изменений в государственном корпоративном секторе, открывает новые возможности для экономической науки и практики.

В ходе анализа явлений и процессов, происходящих в период деятельности предприятий корпоративной формы собственности, перед самими предприятиями и кругом лиц, заинтересованных в его бесперебойной и эффективной работе, стоит множество проблем, связанных, как с определением производственного потенциала предприятия, так и с определением его инвестиционной привлекательности. Поэтому, среди всего многообразия методов исследования объектов необходимо выбрать те, которые в наилучшей форме и как можно более полно могут охарактеризовать те или иные изменения, происходящие в процессе функционирования предприятий.

Существуют определенные методики анализа деятельности предприятий, разработанные как отечественными, так и зарубежными учеными-экономистами, с помощью которых можно оценить потенциал каждого из них [1 – 3]. Но возможность проведения комплексного анализа, выражающегося в разделении всей исследуемой совокупности на подмножества, для которых будут характерны одинаковые тенденции, сделать вручную достаточно сложно. А, следовательно, при разделении всей совокупности государственного корпоративного сектора на однородные классы, обязательно не

только наличие методов и методик проведения подобного рода исследований, но и средств, позволяющих в оптимальные сроки и с допустимыми издержками реализовать данную методику [4 – 6]. Для подобного рода исследований большое значение имеет использование компьютерных средств и на их основе – новых информационных технологий в области математической статистики.

Поэтому **целью данной статьи** является обоснование выбора информационной технологии, реализующей процедуру финансового анализа экономических данных о деятельности государственного корпоративного сектора на примере промышленного сектора экономики и проведение самого анализа с интерпретацией полученных результатов.

### Основной материал

Формирование модели корпоративного контроля в акционерных обществах неразрывно связано с поиском «эффективного собственника», могущего и умеющего оказывать положительное влияние на деятельность этих предприятий. Причем это влияние позитивно не только с позиций финансовых результатов деятельности, но и с позиций выполнения основных требований и положений, указанных в международных и национальных принципах корпоративного управления. Следовательно, решение данной проблемы может рассматриваться в двух аспектах: во-первых, представление рассматриваемой статистической совокупности предприятий в виде однородных групп, признаками для которых будут являться коэффициенты финансово-хозяйственной деятельности; во-вторых, степень соблюдения основных принципов корпоративного управления в каждой группе. Таким образом, можно сузить область поиска, что позволит в максимальной степени адекватно идентифицировать произвольным образом выбранные корпорации, которые и

будут являться устойчивыми носителями их индивидуальных свойств. Предварительная обработка статистических данных представлена значениями основных показателей финансового состояния предприятий. Выбор финансовых коэффициентов зависит от исследователя, однако в данной работе выбраны коэффициенты, входящие в состав Z-счета Альтмана [1].

В основу решения данной задачи должен быть положен принцип классификации объектов. Классификация исследуемых объектов путем их группировки в определенное количество категорий или классов является заданием, решаемым во многих отраслях науки, которое обозначается общим термином «таксономия». Исследователь может осуществлять классификацию субъективно (интуитивно) без каких бы то ни было предварительных статистических или алгебраических расчетов. В этом случае необходимо владеть информацией одного из двух типов [2]: описание объектов с помощью определенного количества характеристик (или переменных); показателями, которые отображают близость объектов, взятых попарно. От одного типа информации к другому можно перейти, используя методы расчета показателей близости (расстояние или степень подобности). В общем случае проблему, которая требует решения методами классификации, можно сформулировать так [3]: а) задается совокупность объектов (наблюдений), описанных определенным количеством характеристик; б) все объекты группируются в такие группы, чтобы они были однородными (гомогенными) в середине каждой группы, и чтобы группы объектов отличались, насколько это возможно, одна от другой (были гетерогенными); в) однородность и неоднородность объектов измеряют на совокупности переменных, которые описывают свойства объектов.

В статистике принято считать, что любую обработку статистических данных (усреднение, установление связей и др.) необходимо осуществлять только в однородных группах наблюдений. Процесс выделения однородных групп в нашем случае основан на двух видах группировок [4]: типологической и структурной. Типологической группировкой считается распределение совокупности на качественно однородные группы, характеризующие некоторые типы (классы) явлений. Выборка, состоящая из крупных и средних промышленных предприятий, в настоящем исследовании представляет собой данный вид группировки. Структурной же группировкой будет расчленение качественно однородной совокупности на группы, характеризующие строевые совокупности, ее структуру [4].

Реализация процедуры кластерного анализа, осуществляемая ручным способом, является громоздким и трудоемким процессом, облегчить кото-

рый может помочь использование современных средств вычислительной техники и новых информационных технологий в области обработки статистических данных.

В своем исследовании мы остановились на программном продукте фирмы Microsoft STATISTICA 6.0 for Windows, в котором реализованы традиционные статистические методы. В данном случае выбор был обусловлен набором математических критериев, по которым можно оценивать качество кластеризации. Этот программный продукт, как и многие другие программы статистического анализа данных, ориентирован преимущественно на выполнение вычислительных функций, тогда как математическому аспекту анализа (обоснованию выбранного алгоритма классификации, определению метрики расстояния между объектами) несмотря на наличие критериев оптимального выбора, внимание не уделяется.

Из множества алгоритмов кластеризации реализованы три: алгоритм «иерархическое дерево», метод K-средних («K-means clustering») и метод двуходового объединения («Two way joining»). Для кластеризации алгоритмов в одной точке, было необходимо оценить расстояние между кластерами (объектами), следовательно, стояла задача выбора масштаба. Поскольку каждый из коэффициентов имеет различные меры весов, то данные необходимо нормализовать (стандартизировать). Процесс выбора количества кластеров наиболее эффективно отражает процедура иерархической классификации объектов. В данном случае в качестве расстояния между кластерами был выбран метод Варда, а расстояния между объектами в кластерах – расстояние Чебышева [5] (рис. 1).

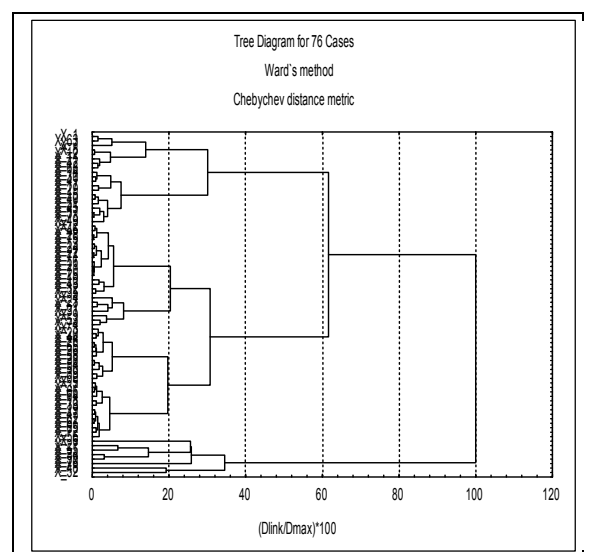


Рис. 1. Дендрограмма иерархической агломеративной кластеризации предприятий промышленного сектора Украины

При данных установках, количество кластеров при 60% разрезе равно трем. Процедура иерархической кластеризации анализа признаков предназначена для малого числа объектов. Ее преимущество состоит в том, что можно детально рассмотреть каждый объект. Но эта процедура не подходит для громадных социологических данных из-за трудоемкости ее агломеративного алгоритма и больших размеров дендрограмм [6].

Поэтому в данном исследовании использовался метод К-средних, который относится к так называемым алгоритмам прямой кластеризации. Он относится к процедурам эталонного типа, поэтому для него, как и для других алгоритмов, существуют определенные функционалы качества классификации, среди которых выделяют: критерий минимизации внутриклассовой дисперсии; минимизацию функции потерь, которые возникают при классификации; максимизацию суммы межклассовых расстояний. В качестве меры близости объектов в пространстве признаков для процедур эталонного типа используется одна из метрик близости – евклидово расстояние, где центроид представляет собой центр тяжести класса: его координаты определяются как средние значения признаков, которые характеризуют группу объектов, принадлежащих этому классу [7]. При этом объект присоединяется к тому кластеру, к центру которого он расположен ближе всего.

В терминах математики результатом любого процесса классификации является формирование определенного набора отношений между группами объектов. Кроме того, качественный характер сформированных отношений является признаком разделения самих алгоритмов классификации.

Так, в процедуре классификации («K-means clustering») результатом классификации является формирование отношений эквивалентности, т.е. группировки с непересекающимися классами. Отношения такого типа предусматривают, что все объекты внутри найденного класса считаются тождественными, а объекты внутри разных классов – нет. Таким образом, в данном случае, осуществляя группировку государственного корпоративного сектора по кластерам (классам), предполагается, что принадлежность к определенной группе предприятия должна быть однозначной.

Тем не менее, в случае возникновения необходимости изменения критериев принадлежности, появляется необходимость в формировании других отношений между объектами, к которым традиционно относят [8]: группировку с пересекающимися классами, которая задается введением степени принадлежности объекта к классу (из теории «размытых множеств»), определением вероятности принадлежности к классу; группировку типа иерархического дерева (процедуры классификации находят

правила подчинения одних классов объектов другим и представляют их в виде иерархии); отношения произвольной структуры, поиск которых предусматривает аппроксимацию исходной выборки.

Таким образом, осуществляется идентификация предприятия, т.е. отнесение его к определенному классу финансового потенциала. В этом случае применение метода «K-means clustering», формирующего отношения эквивалентности между классами, является достаточно обоснованным.

Краткие результаты, которые были получены при применении кластерного анализа данных о финансовом состоянии 77 анализируемых корпоративных предприятий Украины с государственным пакетом, представлены ниже.

Кластеризация предприятий осуществлялась по признаку возможности быстрого погашения краткосрочных обязательств за счет оборотных средств (значению общего коэффициента ликвидности или коэффициента покрытия). При этом учитывались как показатели финансовой устойчивости (коэффициент финансовой автономии, коэффициент маневренности собственного капитала), так и показатели ликвидности оборотных средств (коэффициент общей ликвидности, коэффициент срочной ликвидности) и деловой активности (коэффициент оборачиваемости активов).

В результате анализа было выявлено, что все предприятия исследуемой выборочной совокупности на основании значений 5 финансовых коэффициентов достаточно четко разделяются на 3 группы. Разбиение объектов на три кластера подтверждается графически (рис. 2) и на основании таблиц объектов, входящих в кластеры (рис. 3).

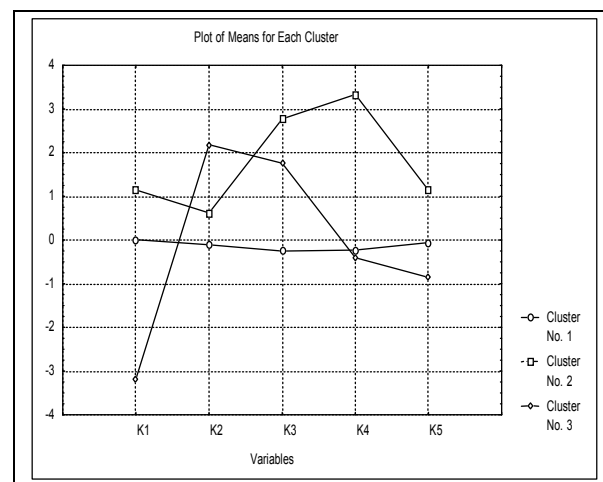


Рис. 2. График средних для каждого кластера

Первый кластер представлен самыми неустойчивыми предприятиями, которые характеризуются значительной выручкой от реализации продукции, высоким уровнем прибыльности, но текущая хозяйственная деятельность при этом финансируется

преимущественно из внешних источников финансирования.

and Distances from Respective Cluster Center Cluster contains 69 cases						
Case No.	Case No.	Case No.	Case No.	Case No.	Case No.	Case No.
x_1	x_2	x_3	x_5	x_7	x_8	
Distance	.889311	.369421	.340683	.380576	1.391607	.390516

and Distances from Respective Cluster Center Cluster contains 5 cases					
Case No.	Case No.	Case No.	Case No.	Case No.	
x_4	x_27	x_34	x_55	x_59	
Distance	.978552	1.228325	.362680	.919073	.706580

and Distances from Respective Cluster Center Cluster contains 2 cases	
Case No.	Case No.
x_6	x_52
Distance	1.593175

Рис. 3. Таблицы распределения объектов по кластерам методом К-средних

Ко второму кластеру относятся относительно финансово устойчивые предприятия, которые характеризуются средним значением привлеченного капитала, нормальным уровнем рентабельности, средним или высоким значением выручки от реализации, текущая деятельность которых достаточно обеспечена собственными оборотными средствами.

И третий кластер объединяет предприятия, которые характеризуются низкой (пассивной) деловой активностью, низким уровнем оборотных средств, которые почти не используют внешние источники финансирования, но направляют свою прибыль на развитие производства.

### Выводы

Данная классификация позволяет выделить те предприятия, выкуп контрольного пакета акций которых позволит сосредоточить власть в руках эффективного собственника с последующей менее

тяжелой модернизацией производства и выходом на отечественные и мировые рынки продукции, а также произвести оптимизацию корпоративного пакета государства.

### Список литературы

1. Дубров А.М. Многомерные статистические методы / А.М. Дубров, В.С. Мхитарян, Л. И. Трошин. – М.: Финансы и статистика, 2000. – 352 с.
2. Айвазян С.А. Классификация многомерных наблюдений / С.А. Айвазян, З.И. Бежаева, О.В. Староверов. – М.: Статистика, 1974. – 365 с.
3. Косенков С.И. Маркетингові дослідження / С.И. Косенков. – К.: Скарби, 2004. – 464 с.
4. Райзен Дж. Классификация и кластер / Дж. Райзен. – М.: Мир, 1980. – 361 с.
5. Мандель И.Д. Кластерный анализ / И.Д. Мандель. – М.: Финансы и статистика, 1988. – 176 с.
6. Анализ социологических данных с применением пакета SPSS [Электронный ресурс]. – Режим доступа к ресурсу: [http://www.ieie.nsc.ru/%7Emetansk/docs/Rostovtsev/book\\_datan](http://www.ieie.nsc.ru/%7Emetansk/docs/Rostovtsev/book_datan).
7. Плюта В. Сравнительный многомерный анализ в экономическом моделировании / В. Плюта; пер. с польского В.В. Иванова. – М.: Финансы и статистика, 1989. – 174 с.
8. Плюта В. Сравнительный многомерный анализ в экономических исследованиях: методы таксономии и факторного анализа / В. Плюта; пер. с польск. В.В. Иванова; науч. ред. В.М. Жуковской. – М.: Статистика, 1980. – 151 с.

Поступила в редколлегию 26.04.2011

Рецензент: д-р техн. наук, проф. В.Б. Мокин, Винницкий национальный технический университет, Винница.

### ВИКОРИСТАННЯ ПРИКЛАДНОГО ПАКЕТА STATISTICA ДЛЯ ВИЯВЛЕННЯ ОДНОЗНАЧНИХ ГРУП В ДЕРЖАВНОМУ КОРПОРАТИВНОМУ СЕКТОРІ ЕКОНОМІКИ УКРАЇНИ

А.А. Гаврилова

Обґрунтовано необхідність використання інформаційної технології STATISTICA 6 для проведення фінансового аналізу діяльності корпоративних підприємств з державним пакетом акцій. Проведено кластерний аналіз сукупності, яка досліджується, з використанням методу ієрархічної агломеративної кластеризації та методу К-середні. Проведено порівняння отриманих результатів, зроблені висновки щодо взаємозв'язку показників фінансової діяльності з оптимізацією корпоративного портфелю держави.

**Ключові слова:** кластеризація, метод Варда, відстань Чебишева, внутрішньокласова дисперсія, евклідова відстань, Z-рахунок Алтмана, показники фінансової діяльності.

### USE THE APPLICATION PACKAGE STATISTICA FOR DETECTION OF UNIQUE GROUPS IN THE PUBLIC CORPORATE SECTOR OF UKRAINE

A.A. Gavrilova

The necessity of using information technology STATISTICA 6 for the financial analysis of corporate enterprises with state shares. A cluster analysis of the study population according to the method of hierarchical agglomerative clustering and method K-means. A comparison of the results and draw conclusions about the relationship of indicators of financial activity with the optimization of the corporate portfolio of the state.

**Keywords:** clustering, Ward's method, Chebyshev distance, intraclass variance, Euclidean distance, Z-account Altman, financial performance.