

Алексеева М.О.

студент

4 курс, Институт экономики управления и финансов

Марийский государственный университет

Россия, г. Йошкар-Ола

Пайдыганова М.Ю.

студент

4 курс, Институт Экономики, Управления и Финансов

Марийский Государственный Университет

Россия, г. Йошкар-Ола

ДИСКРИМИНАНТНЫЙ АНАЛИЗ ПОКАЗАТЕЛЕЙ ПРОМЫШЛЕННЫХ ПРЕДПРИЯТИЙ

***Аннотация:** В представленной статье проводится дискриминантный анализ 10 промышленных предприятий по показателям производительности труда и рентабельности. Предприятия разделяются на группы по разделяющим переменным.*

***Ключевые слова:** статистика, дискриминантный анализ, классификация, группирующие переменные, выборка.*

***Annotation:** This article provides a discriminant analysis of 10 industrial enterprises in terms of labor productivity and profitability. Enterprises are divided into groups by dividing variables.*

***Key words:** statistics, discriminant analysis, classification, grouping variables, sampling.*

Дискриминантный анализ — раздел вычислительной математики, представляющий набор методов статистического анализа для принятия решения

о том, какие переменные разделяют возникающие наборы данных (так называемые «группы») [2, с. 52].

Из годовых отчетов двух групп промышленных организаций получены следующие данные, определяющие их функционирование (табл. 1).

Таблица 1 – Данные предприятий

Группы предприятий	Производительность труда работающего, млн.руб. (x_1)	Рентабельность продукции, % (x_2)
Первая группа	4,0	6
	4,9	6,3
	6,1	7,0
	5,3	7,1
	5,8	6,8
Вторая группа	8,7	9
	10,3	10,5
	11,6	10,9
	10,8	11,0

Нужно показать, можно ли записать во вторую группу организацию, у которой наблюдаемые переменные характеризуются значениями $x_1 = 8,4$; $x_2 = 7,5$.

Решение изучаемой задачи рассмотрено методом дискриминантного анализа в программе STATISTICA. Дискриминантный анализ проведен стандартным способом.

В таблице 2 представлены средние значения для каждой группы и для комбинации всех совокупностей.

Таблица 2 – Средние значения

	Производительность труда работающего, млн.руб. (x_1)	Рентабельность продукции, % (x_2)	N
G_1:1	5,22000	6,64000	5
G_2:2	10,35000	10,35000	4
Все гр.	7,50000	8,28889	9

Средние значения, которые рассчитаны по первой обучающей выборке, меньше средних значений, рассчитанных по второй обучающей выборке. Следовательно, по изучаемым показателям экономическая ситуация промышленных организаций второй группы лучше, чем первой.

Для обзора независимых влияний каждой переменной в общую дискриминацию между показателями организации была построена таблица 3 с выводами дискриминантного анализа.

Таблица 3 – Итоги анализа дискриминантных функций для двух переменных

	Уилкса Лямбда	Частная Лямбда	F- исключ (1,6)	p-уров.	Толер.	1-толер. (R-кв.)
Производительность труда работающего, млн.руб. (x ₁)	0,101677	0,987132	0,078215	0,789122	0,174329	0,825671
Рентабельность продукции, % (x ₂)	0,109592	0,915842	0,551348	0,485810	0,174329	0,825671

Итоги анализа показывают, что:

- число переменных в модели 2;
- значение лямбды Уилкса: 0,10037;
- приближенное значение F –статистики, связанной с лямбда Уилкса при степенях свободы (2, 6) = 26,89; (Критическое значение (0,05;2;6)=5,14;
- P- уровень значимости < 0,00.

По данным значения лямбды Уилкса, который равен 0,10037 и по F-критерию равного 26,89, следует, что полученная классификация почти корректная. Потому как значение лямбды Уилкса по мере возможности должно быть ближе к нулю, а значение F критерия как можно больше.

В общем, критерий Уилкса лямбда является стандартной статистикой, используемой для обозначения статистической значимости мощности дискриминации в текущей модели [1, с. 63].

Частная лямбда Уилкса используется для расчета одиночного влияния переменной в общую дискриминацию между группами. Так как лямбда со

значением 0,0 обозначает полную дискриминацию, то чем ниже ее значение в этом столбце, тем больше одиночный вклад соответствующей переменной в степень дискриминации. Частичная статистика Уилкса лямбда сообщает, что переменная X1 - производительность труда работающего (млн.руб.) имеет больше влияние, чем переменная X2 – рентабельность продукции (%). Поэтому можно сделать вывод, что производительность труда работающего(млн. руб.) считается главным показателем, позволяющим производить дискриминацию между разными промышленными организациями.

При проведении дискриминации объектов по всем переменным может оказаться, что некоторые объекты неправильно отнесены к той или иной группе. Для этого проводится классификация наблюдений [3, с. 74]. В данной работе не было отмечено неправильных классификаций, поэтому все показатели распределены правильно.

В таблице 4 показаны квадраты расстояний Махаланобиса до групповых центров.

Таблица 4 – Квадраты расстояний Махаланобиса до центров для 10 наблюдений

	Наблюд. Класс.	G_1:1 p=,55556	G_2:2 p=,44444
1	G_1:1	1,63991	40,62109
2	G_1:1	0,32157	33,28009
3	G_1:1	1,19212	22,79954
4	G_1:1	1,95200	24,85401
5	G_1:1	0,81696	25,53890
6	G_2:2	12,10071	3,76375
7	G_2:2	30,12912	0,38455
8	G_2:2	40,24992	2,17495
9	G_2:2	38,53457	1,75420
10	---	25,01246	34,18058

Здесь некорректные классификации отмечены звездочками.

По полученным данным видно, что новое предприятие можно отнести к первому классу, так как его показатель в первом классе меньше ($25,01 < 34,18$).

Здесь еще видно, что новое предприятие можно отнести к первому классу ($0,99 > 0,008$).

Таблица 5 – Матрица классификации для двух групп

	Процент правиль.	G_1:1 p=,55556	G_2:2 p=,44444
G_1:1	100,0000	6	0
G_2:2	100,0000	0	4
Всего	100,0000	6	4

Из матрицы классификаций можно сказать, что предприятия были корректно отнесены к полученным группам.

С целью получения минимальной вероятности ошибки при группировке каждого предприятия, в дискриминантном анализе рассматриваются функции классификации (табл. 6), их количество равно количеству групп.

Таблица 6 – Функции классификации для двух переменных

	G_1:1 p=,55556	G_2:2 p=,44444
Производительность труда работающего, млн.руб. (x_1)	-19,4126	-17,912
Рентабельность продукции, % (x_2)	38,8606	44,395
Конст-та	-78,9382	-137,865

В нашей задаче вид функций классификации записывается моделями, описанными ниже:

$$S1 = -19,41 * X1 + 38,86 * X2 - 78,9382;$$

$$S2 = -17,912 * X1 + 44,395 * X2 - 137,865.$$

С помощью полученных функций вычисляются классификационные показатели для вновь рассматриваемых объектов. Добавленный объект будет записан к той группе, для которой классификационный показатель максимален.

Использував описанные формулы, рассмотрена принадлежность всех предприятий к своему классу (табл. 7):

Таблица 7 – Распределение предприятий по группам

№ предприятия	1 гр	2 гр	макс	№ гр
1	76,575	56,857	76,575	1
2	70,7618	54,0547	70,7618	1
3	74,6691	63,6368	74,6691	1
4	94,0853	82,4059	94,0853	1
5	72,7208	60,1314	72,7208	1
6	49,4505	44,6367	49,4505	1
7	101,918	105,856	105,856	2
8	129,148	143,789	143,789	2
9	119,456	138,261	138,261	2
10	138,872	157,03	157,03	2

Полученные данные совпали с данными, полученными в системе Statistica, что говорит о правильности выделенных данных.

Литература

1. Айвазян, С.А. Эконометрика / С.А. Айвазян, С.С. Иванова. - М.: Маркет ДС, 2017. - 104 с.
2. Тихомиров, Н. Методы эконометрики и многомерного статистического анализа / Н. Тихомиров. - М.: Экономика, 2017. - 989с.
3. Шилов, В.В. Библиотечная Эконометрика. Сборник Научных Трудов. Вып.2 / В.В. Шилов. - Москва: Огни, 2016. - 120 с.