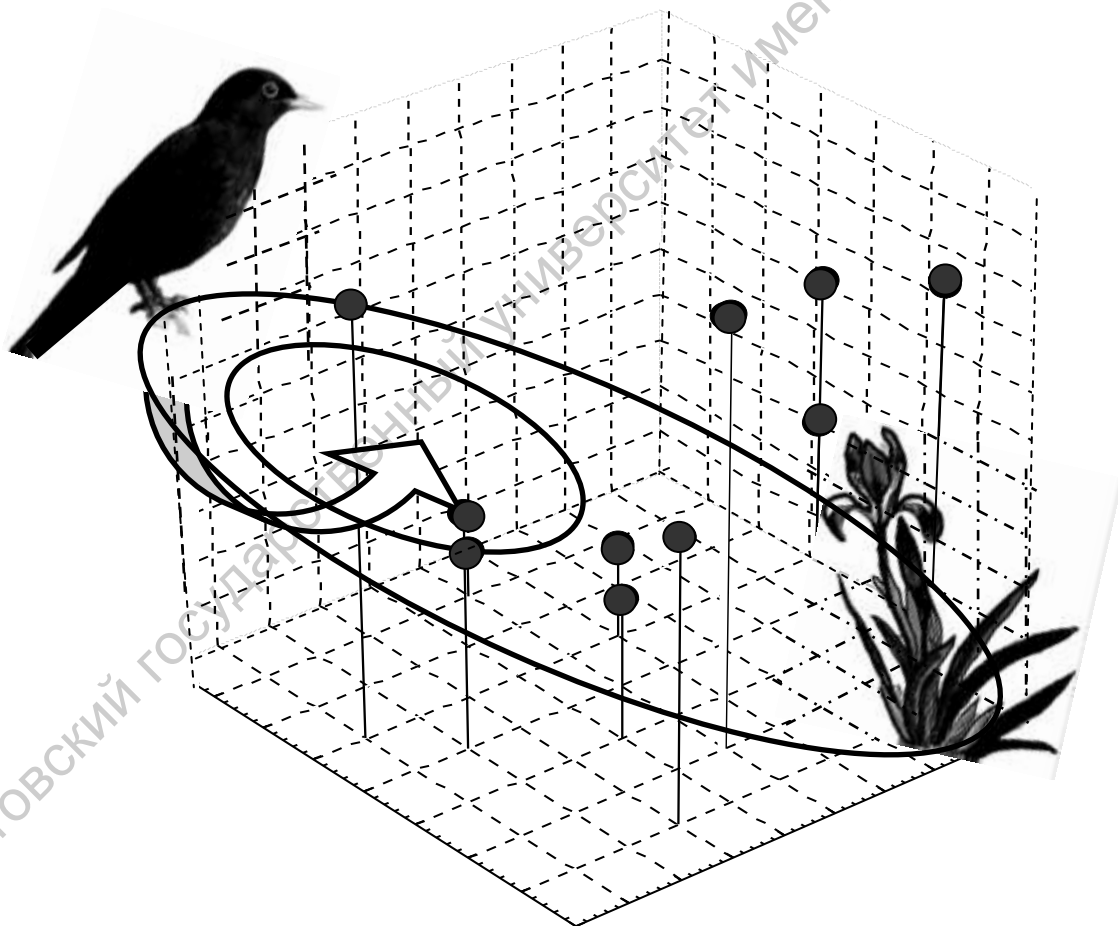


# МНОГОМЕРНЫЕ МЕТОДЫ СТАТИСТИЧЕСКОГО АНАЛИЗА ДАННЫХ В ЭКОЛОГИИ



Саратовский государственный университет им. Н.Г. Чернышевского

# МНОГОМЕРНЫЕ МЕТОДЫ СТАТИСТИЧЕСКОГО АНАЛИЗА ДАННЫХ В ЭКОЛОГИИ

Утверждено Редакционно-издательским советом  
Саратовского государственного университета  
в качестве учебного пособия для студентов биологического факультета,  
обучающихся по специальностям 013100 «Экология», 011600 «Биология»

ИЗДАТЕЛЬСТВО САРАТОВСКОГО УНИВЕРСИТЕТА  
2006

УДК 519.6  
ББК 73  
М73

Авторы: *Т.Н. Давиденко, О.Н. Давиденко,  
В.В. Пискунов, В.А. Болдырев*

**Многомерные методы статистического анализа данных в экологии**  
М73 / Т.Н. Давиденко, О.Н. Давиденко, В.В. Пискунов, В.А. Болдырев: Учеб.  
пособие для студ. биол. фак., обучающихся по спец. 013100 «Экология», 011600  
«Биология». – Саратов: Изд-во Сарат. ун-та, 2006. – 56 с.: ил.  
ISBN 5–292–03604–8

В пособии изложены основы методов статистической обработки данных и методики их применения в экологических исследованиях. Рассмотрены примеры анализа данных в программе *Statistica*, представлены возможные варианты визуализации и объяснения результатов многомерных исследовательских методов, таких как регрессионный, кластерный, факторный и дискриминантный анализы. Приведены варианты заданий для самостоятельного решения.

Для студентов биологического факультета.

Рекомендуют к печати:

Кафедра ботаники и экологии биологического факультета  
Саратовского государственного университета  
Доктор биологических наук, профессор *М.А. Березуцкий*  
Доктор биологических наук, профессор *Н.В. Попов*

*Печатается по решению  
методической комиссии биологического факультета  
Саратовского государственного университета*

УДК 519.6  
ББК 73

ISBN 5–292–03604–8

© Давиденко Т.Н., Давиденко О.Н.,  
Пискунов В.В., Болдырев В.А., 2006

## Оглавление

Предисловие.....	4
<b>1. Общие сведения о программе STATISTICA.....</b>	<b>5</b>
<b>2. Первичный анализ данных.....</b>	<b>6</b>
2.1. Основные понятия описательной статистики.....	6
2.2. Обработка данных и представление результатов.....	7
2.3. Варианты заданий для самостоятельного анализа.....	8
<b>3. Корреляционный анализ.....</b>	<b>11</b>
3.1. Цель и задачи.....	11
3.2. Обработка данных и представление результатов.....	12
3.3. Варианты заданий для самостоятельного анализа.....	14
<b>4. Регрессионный анализ.....</b>	<b>19</b>
4.1. Цель и задачи.....	19
4.2. Обработка данных и представление результатов.....	20
4.3. Варианты заданий для самостоятельного анализа.....	23
<b>5. Кластерный анализ.....</b>	<b>24</b>
5.1. Цель и задачи.....	24
5.2. Обработка данных и представление результатов.....	24
5.3. Методика объяснения результатов.....	27
5.4. Пример обработки данных методом кластерного анализа.....	27
5.5. Варианты заданий для самостоятельного анализа.....	29
<b>6. Факторный анализ.....</b>	<b>33</b>
6.1. Цель и задачи.....	33
6.2. Обработка данных и представление результатов.....	33
6.3. Методика объяснения результатов.....	34
6.4. Пример обработки данных методом факторного анализа.....	35
6.5. Варианты заданий для самостоятельного анализа.....	37
<b>7. Дискриминантный анализ.....</b>	<b>45</b>
7.1. Цель и задачи.....	45
7.2. Обработка данных и представление результатов.....	45
7.3. Пример обработки данных методом дискриминантного анализа.....	46
7.4. Варианты заданий для самостоятельного анализа.....	48
<i>Список рекомендуемой литературы.....</i>	<i>53</i>

## ПРЕДИСЛОВИЕ

Учебное пособие разработано для занятий студентов биологического факультета по спецкурсу "Многомерные статистические методы анализа данных в экологии", который является продолжением лекционного курса "Математические методы в биологии" и позволяет расширить и дополнить полученные ранее знания.

Пособие поможет студентам освоить различные методы статистического анализа данных с использованием специализированной компьютерной программы *STATISTICA* и овладеть многочисленными средствами управления данными, графическими инструментами и возможными вариантами наглядного представления полученных результатов.

Приведенный в пособии материал сгруппирован в два блока – предварительный анализ данных и анализ с применением многомерных статистик. Один из разделов пособия посвящен методу корреляционного анализа, поскольку при работе с многомерными методами рекомендуется исключать из анализа строго скоррелированные переменные. Основное внимание уделено рассмотрению применения в экологических исследованиях многомерных методов, таких как регрессионный, кластерный, факторный и дискриминантный анализы. На примерах из различных областей экологических исследований показаны возможные пути статистического анализа данных с использованием этих методов. По каждому из разбираемых в данном пособии многомерных методов анализа данных приводится сначала информация о целях и задачах конкретного анализа и его возможностях применительно к экологии. Затем рассматривается пример объяснения результатов применения анализа к определенной группе данных. В каждом разделе имеются варианты заданий для самостоятельного решения, которые позволяют углубить понимание разбираемых примеров. Большинство примеров и заданий для самостоятельного анализа являются оригинальными и основаны на собственных данных авторов.

В качестве иллюстративного материала в данном пособии приводятся таблицы и рисунки, выполненные в программе *STATISTICA*, что облегчает в дальнейшем студентам самостоятельную трактовку получаемой графической информации.

## 1. ОБЩИЕ СВЕДЕНИЯ О ПРОГРАММЕ *STATISTICA*

*STATISTICA* является наиболее динамично развивающимся статистическим пакетом и признана мировым лидером на рынке статистического программного обеспечения. Пакет *STATISTICA* дает возможность начинающим пользователям уяснить основные понятия статистики и более полно представить диапазон применения статистических методов. Программа содержит широкий набор методов анализа данных, от самых простых до наиболее сложных, включающих современные достижения в области статистики.

Программа состоит из следующих основных компонентов:

- многофункциональной системы для работы с данными, которая включает в себя электронные таблицы для ввода и задания исходных данных и специальные таблицы для вывода численных результатов анализа;
- графической системы для визуализации исходных данных и результатов статистического анализа;
- набора статистических модулей, в которых собраны группы логически связанных между собой статистических процедур. Каждый модуль является полноценным *Windows*–приложением, поэтому пользователь имеет возможность одновременно работать с несколькими модулями;
- специальных инструментариев для подготовки отчетов.

Статистический анализ данных в программе *STATISTICA* может быть разбит на несколько основных этапов.

1. Ввод данных в электронную таблицу, их предварительное преобразование (структурирование, построение выборок, ранжирование).

2. Визуализация данных при помощи того или иного типа графиков (включает процедуры проверки нормальности распределения данных в выборке, линейности корреляции признаков и др.).

3. Применение конкретной процедуры статистической обработки.

4. Вывод результатов в виде графиков и таблиц.

5. Подготовка и печать отчета.

Для проведения полного статистического анализа не требуется дополнительное программное обеспечение – все этапы статистической обработки, начиная с ввода исходных данных и их преобразований и заканчивая подготовкой отчета, можно выполнить, используя только систему *STATISTICA*. В программе реализован графически ориентированный подход к анализу данных, цель которого состоит в том, чтобы получать всестороннее визуальное представление данных на всех этапах статистической обработки и на основе этого представления выбирать следующий шаг анализа. Встроенный язык программирования *STATISTICA BASIC* позволяет расширять возможности системы, программировать собственные оригинальные методы.

## 2. ПЕРВИЧНЫЙ АНАЛИЗ ДАННЫХ

### 2.1. Основные понятия описательной статистики

Описательная статистика служит для характеристики самых общих свойств изучаемых признаков каких-либо объектов.

*Объем выборки (Valid N)* – количество измерений величин какого-либо признака.

*Сумма (Sum)* – сумма всех значений изучаемой переменной.

*Средняя арифметическая (Mean)* – наиболее типичное (наиболее вероятное) значение в выборке, вокруг которого разбросаны все остальные значения признака, встречающиеся с меньшей вероятностью.

*Ошибка средней арифметической (Std. err. of mean)* – это величина, на которую отличается среднее значение выборки от среднего значения генеральной совокупности при условии, что распределение признака близко к нормальному.

*Стандартное отклонение (среднее квадратическое отклонение) (Standart deviation)* – мера изменчивости признака – показывает, на какую величину в среднем отклоняются варианты от среднего значения признака.

*Дисперсия (Variance)* – средний квадрат отклонений вариант от среднего значения признаков.

*Медиана (Median)* – это такое значение случайной величины, которое разделяет все случаи выборки на две равные по численности части.

*95-% доверительный интервал (Conf. limits for mean)* – интервал, в который с вероятностью 95% попадает среднее значение признака генеральной совокупности.

*Размах (Range)* – расстояние между наибольшим и наименьшим значениями признака.

*Асимметрия (Skewness)* – характеризует степень смещения вариационного ряда относительно среднего значения по величине и направлению. В симметричной кривой коэффициент асимметрии равен нулю. Если правая ветвь кривой больше левой, то асимметрия больше нуля, и наоборот. Асимметрия менее 0,5 считается малой.

*Экссесс (Kurtosis)* – характеризует степень концентрации вариант вокруг среднего значения и является своеобразной мерой крутизны угла наклона кривой распределения признака. В кривой нормального распределения эксцесс равен нулю. Если эксцесс больше нуля, то кривая распределения характеризуется островершинностью, если меньше нуля – кривая плосковершинна. Отрицательным пределом величины эксцесса является число  $-2$ , положительного предела нет.

*Коэффициент вариации (CV)* – отношение стандартного отклонения к среднему значению, умноженное на 100%. При величине коэффициента вариации менее 10% изменчивость оценивается как слабая, 11–25% – средняя, более 25% – сильная.

*Верхний квартиль (Upper quartiles)* – это такое значение случайной величины, больше которого 25% случаев выборки.

*Нижний квартиль (Lower quartiles)* – это такое значение случайной величины, меньше которого 25% случаев выборки.

## 2.2. Обработка данных и представление результатов

Анализ данных начинается с ввода исходных количественных значений в электронную таблицу *Statistica*.

При использовании модуля "Описательная статистика" ввод данных осуществляется следующим образом: по горизонтали (верхняя строка) – анализируемые признаки, по вертикали (по столбцам) – их количественные значения (табл. 1).

Таблица 1

**Расположения данных в таблице *Statistica* при работе с модулем "Описательная статистика"**

Признак 1	Признак 2	Признак 3	Признак 4	Признак 5	Признак 6	Признак 7
790	15,1	62,4	210	2,1	6000	1,2
880	17,5	70	691	3,5	1500	1,9
750	16	62,5	780	3	900	1,6
500	10	30	160	1,5	90	1,1
920	12,1	50,2	960	2,1	120	1,1
450	11,1	29,8	500	1,3	6000	1,1

Программа позволяет выбирать показатели описательной статистики, которые необходимо рассчитать. По итогам анализа на экран выводится таблица результатов, в которой отражены количественные значения всех заданных показателей для всех выбранных признаков (табл. 2).

Таблица 2

**Пример таблицы результатов описательной статистики**

Признаки	Valid N	Mean	Median	Mode	Sum	Variance	Std.Dev.	Skewness	Kurtosis
Признак 1	40	2,84	2,38	3,21	113,98	1,79	1,33	1,08	0,71
Признак 2	40	1,88	1,70	1,00	75,55	0,55	0,74	0,49	-0,97
Признак 3	40	20,92	20,60	6,10	836,80	144,97	12,04	0,52	-0,70
Признак 4	40	47,41	46,90	2,11	1896,70	469,20	21,66	0,20	-0,61
Признак 5	40	5,84	5,35	1,32	233,70	8,74	2,95	1,16	1,71



## 2.3. Варианты заданий для самостоятельного анализа

### ВАРИАНТ 1

На территории лесопарка на учетной площади проводилось измерение диаметров стволов подроста вяза гладкого.

Диаметр стволов вяза, см: 1,4; 1,1; 0,8; 1,7; 0,9; 2,0; 1,8; 0,9; 0,5; 1,4; 1,1; 1,8; 0,8; 1,1; 2,7; 2,5; 1,2; 1,0; 1,2; 2,0; 1,6; 2,6; 2,4; 0,9; 1,6; 1,9; 2,0; 2,1; 2,6; 2,8; 1,9; 2,0; 1,6; 1,6; 1,2; 4,2; 1,9; 2,1; 1,8; 3,2; 2,1; 2,1; 2,2; 0,8; 1,1; 2,7; 2,5; 1,2; 1,0; 1,2; 1,9; 2,1; 1,8; 3,2; 1,5; 4,2; 2,3; 1,5; 2,4; 3,2; 1,1; 2,7; 2,5; 1,2; 1,0; 1,2; 1,1; 2,7; 2,5; 1,2; 1,0; 1,2; 4,2; 1,9; 2,1; 1,8; 3,2; 2,1; 2,1; 2,2; 0,8; 1,1; 2,7; 2,5; 1,2; 4,2; 1,9; 2,1; 1,8; 3,2; 2,1; 2,1; 2,2; 0,8; 1,1; 2,7; 2,5; 1,2; 1,1; 2,7; 2,5; 1,2; 1,0; 1,2; 2,0; 1,6; 2,6; 2,4; 0,9; 1,6; 1,9; 2,0; 2,1; 2,6; 1,1; 2,7; 2,5; 1,2; 1,0; 1,2; 2,0; 1,6; 2,6; 2,4; 0,9; 1,6; 1,9; 2,0; 2,1; 2,6; 1,1; 2,7; 2,5; 1,2; 1,0; 1,2; 2,0; 1,6; 2,6; 2,4; 0,9; 1,6; 1,9; 2,0; 2,1; 2,6; 1,1; 2,7; 2,5; 1,2; 1,0; 1,2; 2,0; 1,6; 2,6; 2,4; 0,9; 1,6; 1,9; 2,0; 2,1; 2,6; 1,1; 2,7; 2,5; 1,2; 1,0; 1,2; 2,0; 1,6; 2,6; 2,4; 0,9; 1,6; 1,9; 2,0; 2,1; 2,6; 1,1; 2,7; 2,5; 1,2; 1,0; 1,2; 3,2; 1,5; 4,2; 2,3; 1,5; 2,4; 3,2; 1,1; 2,7; 2,5; 1,2; 1,0; 1,2; 2,6.

С помощью процедуры "Описательная статистика" вычислите следующие показатели: объем выборки, среднюю арифметическую, стандартное отклонение, верхний и нижний квартили, медиану.

Чему равен коэффициент вариации изучаемого признака?

### ВАРИАНТ 2

На территории парка, расположенного в черте города, проведено измерение высоты подроста вяза шершавого с целью выяснения характера возобновления основной древесной породы.

Высота вяза шершавого, м: 3,3; 3,7; 4,7; 6,4; 6,4; 4,4; 4,3; 2,3; 2,2; 4,4; 5,1; 1,4; 1,8; 1,5; 3,4; 2,1; 3,2; 3,2; 2,7; 2,9; 2,5; 1,8; 1,8; 5,0; 2,2; 2,7; 2,2; 1,4; 3,5; 2,5; 3,5; 2,7; 5,0; 4,5; 2,1; 2,8; 3,2; 3,6; 1,4; 1,9; 2,0; 2,1; 4,0; 2,5; 3,4; 3,8; 2,8; 1,3; 1,5; 1,5; 1,6; 1,8; 1,5; 3,4; 2,1; 3,2; 5,0; 4,5; 2,1; 2,8; 2,2; 2,7; 2,2; 1,8; 1,5; 3,4; 2,1; 3,2; 3,2; 2,7; 5,0; 2,2; 2,7; 2,2; 1,4; 3,5; 1,8; 1,5; 3,4; 2,1; 3,2; 5,0; 4,5; 2,1; 2,8; 2,2; 2,7; 2,2; 1,8; 1,5; 3,4; 1,8; 1,5; 3,4; 2,1; 3,2; 5,0; 4,5; 2,1; 2,8; 2,2; 2,7; 2,2; 1,8; 1,5; 3,4; 1,8; 1,5; 3,4; 2,1; 3,2; 5,0; 4,5; 2,1; 2,8; 2,2; 2,7; 2,2; 1,8; 1,5; 1,4; 1,9; 2,0; 2,1; 4,0; 2,5; 3,4; 3,8; 2,8; 1,3; 1,4; 1,9; 2,0; 2,1; 4,0; 2,5; 3,4; 3,8; 2,8; 1,3; 3,4.

С помощью процедуры "Описательная статистика" вычислите следующие показатели: среднюю арифметическую, ошибку средней, медиану, асимметрию, размах.

Чему равен коэффициент вариации изучаемого признака?



## ВАРИАНТ 5

На участке липняка, расположенного на территории лесопарка, проводилось измерение сомкнутости крон древостоя.

Сомкнутость крон, %: 40, 40, 50, 20, 20, 10, 30, 30, 30, 10, 50, 40, 50, 60, 60, 50, 40, 50, 40, 40, 30, 30, 60, 70, 70, 60, 70, 50, 40, 40, 60, 60, 60, 50, 30, 10, 10, 20, 30, 40, 40, 30, 30, 30, 30, 20, 30, 30, 40, 50, 70, 40, 60, 40, 30, 20, 40, 70, 50, 40, 40, 60, 60, 60, 50, 30, 30, 40, 50, 60, 70, 70, 60, 70, 50, 40, 40, 60, 60, 60, 50, 30, 10, 10, 20, 30, 40, 40, 30, 30, 30, 20, 30, 30, 40, 50, 70, 40, 60, 40, 30, 20, 40, 70, 50, 40, 40, 60, 60, 70, 70, 60, 70, 50, 40, 40, 60, 60, 60, 50, 30, 10, 10, 20, 30, 40, 40, 30, 30, 30, 20, 30, 30, 40, 50, 70, 40, 60, 40, 30, 20, 40, 70, 50, 40, 40, 60, 60, 70, 70, 60, 70, 50, 40, 40, 60, 60, 60, 50, 30, 10, 70, 40, 60, 40, 30, 20, 40, 70, 50, 40, 40, 60, 60, 60, 50, 70, 40, 60, 40, 30, 20, 40, 70, 50, 40, 40, 60, 60, 60, 50, 70, 40, 60, 40, 30, 20, 40, 70, 50, 40, 60, 60, 60, 50, 70, 40, 60, 40, 30, 20, 40, 70, 50, 40, 40, 60, 60, 60, 50, 10.

При помощи "Описательной статистики" вычислите следующие показатели: объем выборки, среднюю арифметическую, ошибку средней, стандартное отклонение, 95-% доверительный интервал, размах, эксцесс. Чему равен коэффициент вариации признака?

## ВАРИАНТ 6

На учетных площадках в пределах кленовых растительных сообществ Хвалынского Национального парка проведено измерение высоты древостоя.

Высота клена платановидного, м: 13,3; 13,7; 14,7; 16,4; 16,4; 14,4; 14,3; 12,3; 12,2; 14,4; 15,1; 11,4; 11,8; 11,5; 13,4; 12,1; 13,2; 13,2; 12,7; 12,9; 12,5; 11,8; 11,8; 15,0; 12,2; 11,7; 12,2; 11,4; 13,5; 12,5; 13,5; 12,7; 15,0; 14,5; 12,1; 12,8; 13,2; 13,6; 11,4; 11,9; 12,0; 12,1; 14,0; 12,5; 13,4; 13,8; 12,8; 11,3; 11,5; 11,5; 11,6; 11,4; 11,8; 11,5; 13,4; 12,1; 13,2; 13,2; 12,7; 11,2; 12,3; 14,1; 12,1; 12,6; 15,1; 11,4; 13,5; 12,5; 13,5; 12,7; 15,0; 14,5; 12,1; 12,8; 11,7; 12,2; 11,4; 13,5; 12,5; 13,5; 12,7; 15,0; 14,5; 12,1; 11,7; 12,2; 11,4; 13,5; 12,5; 13,5; 12,7; 15,0; 14,5; 12,1; 11,9; 12,0; 12,1; 14,0; 12,5; 13,4; 13,8; 12,8; 11,3; 11,5; 11,5; 11,9; 12,0; 12,1; 14,0; 12,5; 13,4; 13,8; 12,8; 11,3; 11,5; 11,5; 12,1; 12,8; 11,7; 12,2; 11,4; 13,5; 12,5; 13,5; 12,7; 15,0; 14,5; 12,1; 11,7; 12,2; 11,4; 13,5; 12,5; 13,5; 12,7; 15,0; 14,5; 12,1; 11,9; 12,0; 12,1; 14,0; 12,5; 13,4; 13,8; 12,8; 12,1; 12,8; 11,7; 12,2; 11,4; 13,5; 12,5; 13,5; 12,7; 15,0; 14,5; 12,1; 11,7; 12,2; 11,4; 13,5; 12,5; 13,5; 12,7; 15,0; 14,5; 12,1; 11,9; 12,0; 12,1; 14,0; 12,5; 13,4; 13,8; 13,5; 12,7; 15,0; 14,5; 12,1; 11,9; 12,0; 12,1; 14,0; 13,5; 12,7; 15,0; 14,5; 12,1; 11,9; 12,0; 12,1; 14,0; 12,8.

С помощью процедуры "Описательная статистика" вычислите следующие показатели: среднюю арифметическую, ошибку средней, медиану, асимметрию, размах. Чему равен коэффициент вариации изучаемого признака?

### 3. КОРРЕЛЯЦИОННЫЙ АНАЛИЗ

#### 3.1. Цель и задачи

Корреляция – это мера зависимости между переменными. Различают линейную и нелинейную зависимости. Линейная корреляция описывается коэффициентом корреляции Пирсона и измеряет степень линейных связей между переменными, т.е. определяет, насколько значения двух переменных пропорциональны друг другу. Коэффициент корреляции является безразмерной величиной. Значение коэффициента корреляции не зависит от масштаба измерения и изменяется в пределах от  $-1,00$  до  $+1,00$ . Значение  $-1,00$  показывает, что переменные скоррелированы строго отрицательно, значение  $+1,00$  означает, что переменные скоррелированы строго положительно. Корреляция является высокой, если на графике зависимость между двумя переменными можно представить в виде прямой линии (с положительным или отрицательным углом наклона) (рис. 1).

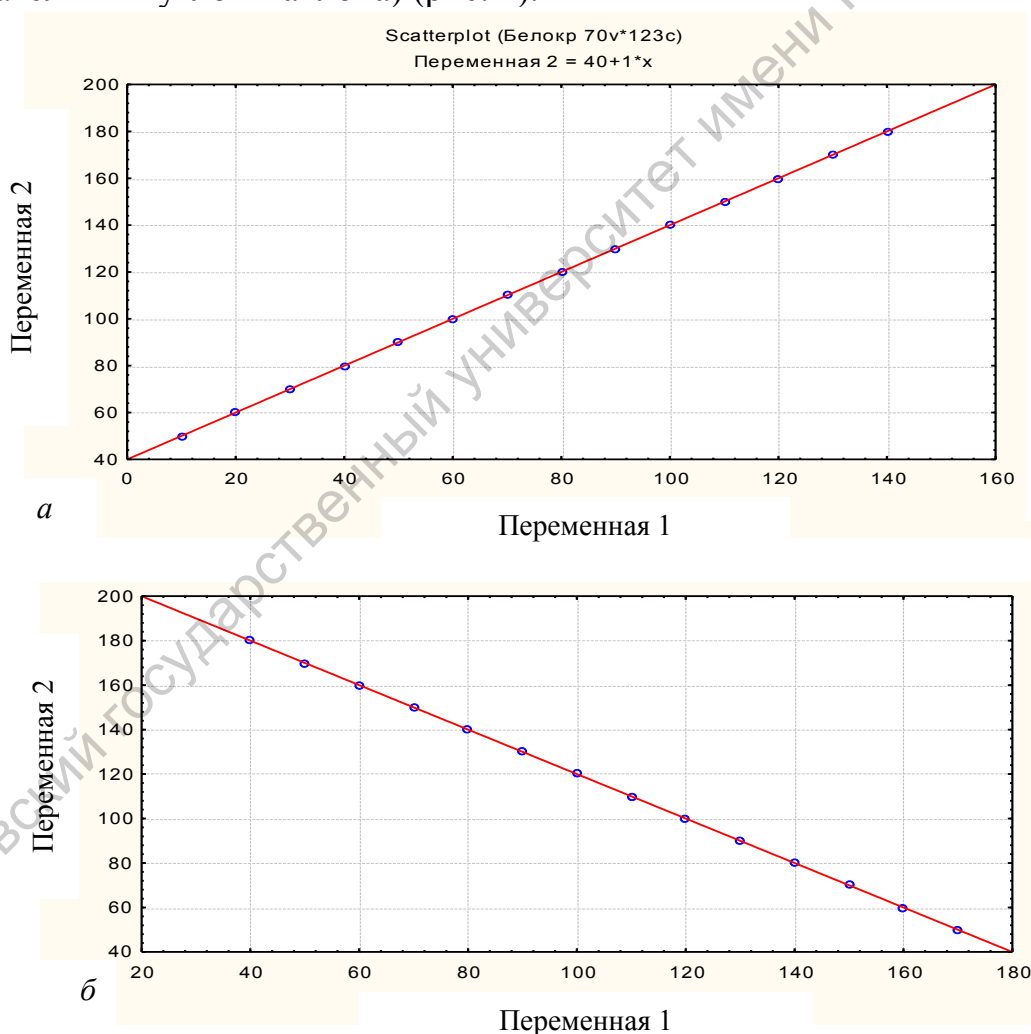


Рис. 1. Графическое изображение линейных положительной (а) и отрицательной (б) корреляционных зависимостей между двумя переменными

Если зависимость между переменными нельзя описать прямой (или близкой к ней) линией, то корреляция между переменными носит нелинейный характер. Для ее нахождения используется непараметрический коэффициент корреляции Спирмена. Его значения также изменяются в пределах от  $-1,00$  до  $+1,00$ .

Главным источником информации о надежности корреляции служит уровень значимости ( $p$ ). Уровень значимости более  $0,05$  считается очень низким, и при таком уровне корреляции считаются недостоверными. В экологических исследованиях для получения достоверных данных используются выборки от 100 и более измерений.

### 3.2. Обработка данных и представление результатов

При работе с корреляционными зависимостями ввод данных в электронную таблицу осуществляется следующим образом: по горизонтали – признаки, между которыми необходимо выявить корреляционные зависимости, по вертикали – средние значения этих признаков на одноименных площадках или у одних и тех же групп растений или животных (табл. 3).

Таблица 3

Пример расположения данных в электронной таблице *Statistica* при работе методом корреляционного анализа

№ площадки	Высота растений	Содержание хлоридов
1	10	0,6
2	10,1	0,56
3	10,6	0,59
4	10	0,6
5	11,1	0,9
6	11,3	0,91
7	10,1	0,7
8	13,5	0,9
9	12,3	0,89
10	14,5	0,9
11	11	0,7
12	12	0,5
13	11,8	0,6
14	13,4	0,75
15	11,4	0,7
16	12	0,85
17	15,6	0,8

Перед тем как проводить корреляционный анализ, необходимо выяснить характер зависимости между переменными: линейный или нелинейный. Для этого при помощи вложения "Графики" строится график зависимости между изучаемыми переменными. Результат выводится на экран в следующей форме (рис. 2).

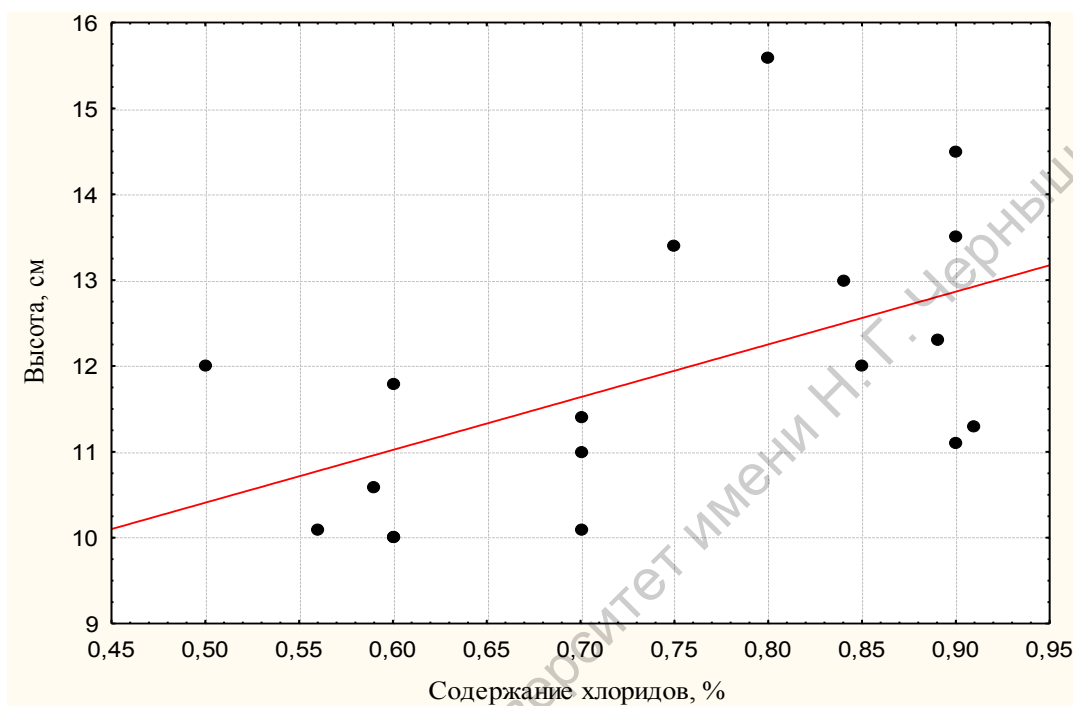


Рис. 2. Графическое изображение зависимости между высотой растений солероса европейского и содержанием хлоридов в почве

В данном случае зависимость носит нелинейный характер (точки не укладываются на линию регрессии), поэтому для дальнейшего анализа мы используем непараметрический коэффициент корреляции Спирмена. Полученный результат выводится на экран в следующей форме (табл. 4).

Таблица 4

**Представление результатов корреляционного анализа в программе Statistica**

	Valid	Spearman	t(N-2)	p-level
Высота растений/содержание хлоридов	20	<b>0,458828</b>	2,190866	0,041863

Примечание: Valid – число измерений признаков, Spearman – значение коэффициента корреляции Спирмена, t(N-2) – число степеней свободы, p-level – уровень значимости. Выделяются значения коэффициента корреляции при уровне значимости  $\leq 0,05$ .

Коэффициент корреляции Спирмена равен 0,46 при уровне значимости 0,04. Это свидетельствует о наличии значимой нелинейной положительной корреляции между содержанием хлоридов в верхней части почвенного профиля и высотой растений солероса европейского.

### 3.3. Варианты заданий для самостоятельного анализа

На 10 участках леса были проведены измерения диаметра стволов деревьев, высоты кустарников, проективного покрытия травостоя и плотности населения птиц-дуплогнездников (табл. 5–8).

Таблица 5

Значения диаметров стволов деревьев, см, на 10 участках леса

№ участка									
1	2	3	4	5	6	7	8	9	10
156,9	92,3	115,2	64,2	84,6	64,2	69,1	40,6	63,4	34,6
104,3	85,4	73,7	68,4	57,2	71,3	45,5	52,1	51,2	57,9
66,9	63,4	77,8	48,3	61,9	63,9	83,4	45,3	64,9	72,8
69,5	69,1	56,3	52,9	63,4	64,8	34,6	67,2	72,1	64,6
43,5	87,3	67,2	69,4	51,2	77,1	78,8	32,1	77,7	80,3
84,2	64,9	84,4	84,9	64,9	79	80,3	25,9	34,6	68,7
79,8	72,8	47,3	77,8	72,1	64,2	81,4	69,1	67,9	85,5
89,9	34,6	39,6	76,3	77,7	34,6	79,3	36,2	25,6	23,1
156,2	112,3	66,3	89,4	34,6	60,1	70,1	21,3	45,2	45,2
56,6	72,8	59,4	94,2	67,9	70,8	61,4	36,9	55,2	66,3
189,3	64,6	78,6	69,8	51,4	77	69,3	68,4	41,3	52,1
94,2	80,3	98,9	67,2	72,3	70,1	70,1	48,3	89,9	13,2
125,2	68,7	114,9	72,9	78,2	69,3	80,3	52,9	18,6	12,9
101,3	85,5	40,9	84,6	69,2	47,8	40,8	69,4	17,3	19,3
147,3	126,3	59,3	82,3	79,8	44,6	20,3	84,9	55,2	15,6
55,9	61,5	109,5	74,8	59,8	98,6	39,6	77,8	19,3	13,2
114,3	81,6	64,2	49,2	87,3	69,1	79,3	76,3	20,1	18,2
109,3	152,8	58,4	89,2	71,8	96,3	70,1	39,9	23,5	19,5
68,3	51,9	62,1	39,4	72,3	79,3	69,1	23,1	25,1	12,1
84,2	64,9	84,4	84,9	64,9	79	80,3	25,9	34,6	68,7
79,8	72,8	47,3	77,8	72,1	64,2	81,4	69,1	67,9	85,5
89,9	34,6	39,6	76,3	77,7	34,6	79,3	36,2	25,6	23,1
156,2	112,3	66,3	89,4	34,6	60,1	70,1	21,3	45,2	45,2
56,6	72,8	59,4	94,2	67,9	70,8	61,4	36,9	55,2	66,3
189,3	64,6	78,6	69,8	51,4	77	69,3	68,4	41,3	52,1
94,2	80,3	98,9	67,2	72,3	70,1	70,1	48,3	89,9	13,2

Таблица 6

Значение высоты кустарникового яруса, м, на 10 участках леса

№ участка									
1	2	3	4	5	6	7	8	9	10
2,6	1,3	2,3	1,3	1,8	0,9	3,2	2,4	3,2	2,4
1,9	1,8	2,6	1,8	1,7	0,6	1,4	1,6	1,4	1,6
1,7	4,2	3,4	3,2	0,9	1,2	1,8	2,4	1,8	2,4
1,2	3,1	3,6	1,7	1,2	1,4	0,9	1,6	0,9	1,6
0,9	6,8	1,1	0,9	1,4	1,8	0,6	3,9	0,6	3,9
1,6	0,9	0,9	2,3	2,3	1,3	0,4	0,9	0,4	0,9
3,2	2,4	1,7	1,9	2,6	1,8	0,9	1,2	0,9	1,2
1,4	1,6	2,1	1,3	3,4	3,2	1,1	1,6	1,1	1,6
1,8	2,4	1,9	1,1	3,6	1,7	1,3	3,2	1,3	3,2
0,9	1,6	1,6	1,9	1,1	0,9	0,9	1,8	0,9	1,8
0,6	3,9	3,4	2,6	0,9	2,3	1,6	1,7	1,6	1,7
0,4	0,9	1,8	3,1	1,7	1,9	1,3	0,9	1,3	0,9
0,9	1,2	2,5	4,8	2,1	1,3	0,9	1,2	0,9	1,2
1,1	1,6	3,1	1,4	1,9	1,1	0,6	1,4	0,6	1,4
1,3	3,2	3,8	0,9	1,6	1,9	1,3	2,3	1,3	2,3
0,9	1,8	0,9	1,6	3,4	2,6	1,8	2,6	1,8	2,6
1,6	1,7	0,6	1,3	1,8	3,1	4,2	3,4	4,2	3,4
1,3	0,9	1,2	1,2	2,5	4,8	3,1	3,6	3,1	3,6
0,9	1,2	1,4	1,6	3,1	1,4	6,8	1,1	6,8	1,1
0,9	1,6	1,6	1,9	1,1	0,9	0,9	1,8	0,9	1,8
0,6	3,9	3,4	2,6	0,9	2,3	1,6	1,7	1,6	1,7
0,4	0,9	1,8	3,1	1,7	1,9	1,3	0,9	1,3	0,9
0,9	1,2	2,5	4,8	2,1	1,3	0,9	1,2	0,9	1,2
1,1	1,6	3,1	1,4	1,9	1,1	0,6	1,4	0,6	1,4
1,3	3,2	3,8	0,9	1,6	1,9	1,3	2,3	1,3	2,3
0,9	1,8	0,9	1,6	3,4	2,6	1,8	2,6	1,8	2,6
1,1	1,6	3,1	1,4	1,9	1,1	0,6	1,4	0,6	1,4
1,3	3,2	3,8	0,9	1,6	1,9	1,3	2,3	1,3	2,3
0,9	1,8	0,9	1,6	3,4	2,6	1,8	2,6	1,8	2,6
1,6	1,7	0,6	1,3	1,8	3,1	4,2	3,4	4,2	3,4
1,1	1,6	3,1	1,4	1,9	1,1	0,6	1,4	0,6	1,4



Таблица 7

Значения проективного покрытия травостоя, %, на 10 участках леса

№ участка									
1	2	3	4	5	6	7	8	9	10
50	0	10	30	20	60	30	20	20	90
70	10	20	20	20	50	20	30	20	50
60	10	20	10	40	90	10	30	20	60
60	10	20	10	40	10	10	20	30	50
60	20	30	20	30	90	20	50	30	40
50	20	30	20	30	90	30	40	20	90
50	30	20	20	20	80	30	10	20	80
40	30	10	20	40	50	20	90	30	80
60	40	10	20	30	60	40	80	30	30
20	30	20	30	50	50	60	50	20	50
40	10	30	30	40	40	50	60	30	60
30	10	30	20	40	20	80	50	40	50
10	20	20	20	30	90	80	40	30	80
10	20	40	30	20	80	80	20	20	90
40	20	30	30	30	70	40	50	10	70
30	50	20	20	30	20	10	70	10	70
50	30	10	30	20	10	90	60	10	50
60	20	30	20	30	90	20	50	30	40
50	20	30	20	30	90	30	40	20	90
50	30	20	20	20	80	30	10	20	80
40	30	10	20	40	50	20	90	30	80
60	40	10	20	30	60	40	80	30	30
20	30	20	30	50	50	60	50	20	50
40	10	30	30	40	40	50	60	30	60
30	10	30	20	40	20	80	50	40	50
10	20	20	20	30	90	80	40	30	80
10	20	40	30	20	80	80	20	20	90
40	20	30	30	30	70	40	50	10	70
50	30	20	20	20	80	30	10	20	80
40	30	10	20	40	50	20	90	30	80
60	40	10	20	30	60	40	80	30	30

## Плотность населения птиц-дуплогнезdnиков, пар/100 га, на 10 участках леса

№ участка									
1	2	3	4	5	6	7	8	9	10
96,6	102,3	96,3	200,3	250,3	200,6	150	325,2	90	190,3
200,3	409,3	100,9	150,6	100,3	400,2	500	102	100	100
213,2	103,5	500,1	100	100	500	250,1	90,5	150,5	100
200,6	114,9	89,9	100	100	360,2	123,5	56,5	150	100
400,2	96,3	45,2	102	96,2	100,5	140,6	90	200	200
500	45,2	400,9	100	409,3	89,9	120	100	100	98,9
360,2	89,9	70,9	350,2	103,5	45,2	100	150,5	150	56,5
500,1	100,3	205,6	100	100,3	200	100	150	96	80,2
89,9	89,9	201,4	200,3	200,3	100,9	203,5	200	105,2	70
45,2	200,1	400,3	500,9	150,6	325,2	200	100	100	65,5
400,9	206,2	350,9	100	100	102	100	150	100	90
406,3	201	100,3	100	100	90,5	150	160,3	89,9	100
325,2	106,3	100	102,3	102	56,5	101,2	120,5	90,6	100
102	205,6	100	105,9	100	90	100	100	100	45,5
100	201,4	96,2	190,3	350,2	100	100	100	100	100
350,2	400,3	409,3	100	100	150,5	150	140	100	205,3
100	350,9	103,5	100	100,9	150	120,9	160	205,2	100
500,1	100,3	205,6	100	100,3	200	100	150	96	80,2
89,9	89,9	201,4	200,3	200,3	100,9	203,5	200	105,2	70
45,2	200,1	400,3	500,9	150,6	325,2	200	100	100	65,5
400,9	206,2	350,9	100	100	102	100	150	100	90
406,3	201	100,3	100	100	90,5	150	160,3	89,9	100
325,2	106,3	100	102,3	102	56,5	101,2	120,5	90,6	100
102	205,6	100	105,9	100	90	100	100	100	45,5
100	201,4	96,2	190,3	350,2	100	100	100	100	100
213,2	103,5	500,1	100	100	500	250,1	90,5	150,5	100
200,6	114,9	89,9	100	100	360,2	123,5	56,5	150	100
400,2	96,3	45,2	102	96,2	100,5	140,6	90	200	200
500	45,2	400,9	100	409,3	89,9	120	100	100	98,9
360,2	89,9	70,9	350,2	103,5	45,2	100	150,5	150	56,5
500,1	100,3	205,6	100	100,3	200	100	150	96	80,2

Необходимо выявить, имеются ли корреляции между:

- а) высотой кустарников и проективным покрытием травостоя;
- б) диаметром стволов деревьев и плотностью гнездования птиц.

По итогам проведения корреляционного анализа сделать вывод о характере корреляции (линейная или нет), ее знаке и значимости.

Последовательность работы

1. Внести в таблицу имеющиеся данные по количественным значениям всех переменных, обозначив наименование параметра и номер площадки.
2. Получить среднее значение для каждого параметра на каждой из площадок, используя описательную статистику.
3. Составить новую таблицу, в которой приведены только средние значения параметров (табл. 9).

Таблица 9

**Расположение в электронной таблице *Statistica* полученных данных по средним значениям параметров на 10 участках леса**

№ участка	Плотность гнездования, пар/100 га	Диаметр древостоя, см	Высота кустарников, м	Покрытие травостоя, %
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				

4. Оценить по полученным средним значениям, какой тип зависимости (линейный или нет) характерен для каждой из пар признаков.
5. Провести корреляционный анализ между заданными параметрами с учетом характера корреляционной зависимости (используя коэффициент Спирмена или Пирсона).
6. Сделать вывод.

## 4. РЕГРЕССИОННЫЙ АНАЛИЗ

### 4.1. Цель и задачи

Регрессионный анализ является одним из методов статистического анализа, который может быть использован в экологии для исследования связей между характеристиками видов растений и животных и факторами окружающей среды. При этом регистрироваться может либо обилие видов, либо показатели их жизнеспособности. В отличие от кластерного анализа регрессионный анализ предполагает обработку данных по каждому виду в отдельности.

При работе методом регрессионного анализа переменная обилия или любая другая характеристика видов (например, проективное покрытие, степень развития крон и т.д.) является зависимой (объясняемой переменной), а переменные факторы среды – независимыми (объясняющими переменными или предикторами).

В экологических исследованиях регрессионный анализ используют для решения следующих задач:

- 1) оценки экологических параметров, например оптимума и экологической амплитуды вида;
- 2) определения (посредством проверки статистической значимости) того фактора среды, который вносит наибольший вклад в реакцию отклика вида, и тех факторов, которые не имеют большого значения;
- 3) прогнозирования обилия или присутствия/отсутствия видов в местообитаниях по наблюдаемым значениям одного и более факторов среды;
- 4) прогнозирования значений факторов среды в местообитаниях по наблюдаемым значениям одного или более видов (калибровка).

Зависимость между переменными величинами  $Y$  и  $X$  можно выразить аналитически с помощью формул и уравнений и графически в виде геометрического места точек в системе прямоугольных координат (диаграмма рассеяния). В регрессионном анализе программа строит линию регрессии так, чтобы минимизировать квадраты отклонений этой линии от реально наблюдаемого разброса точек. В практических исследованиях с целью возможности прогнозирования возникает необходимость описать диаграмму рассеяния математическим уравнением.

При рассмотрении влияния двух независимых переменных (в двумерном пространстве) прямая линия на плоскости задается уравнением

$$Y = a + b \cdot X,$$

где  $Y$  – зависимая переменная (обилие или присутствие вида),  $a$  – константа (свободный член регрессии),  $b$  – угловой коэффициент (регрессионный),  $X$  – независимая переменная.

В случае рассмотрения влияния множества факторов (многомерный вариант), когда имеется более одной независимой переменной, процедура

множественной регрессии будет оценивать параметры линейного уравнения вида

$$Y = a + b_1 \cdot X_1 + b_2 \cdot X_2 + \dots + b_p \cdot X_p.$$

Здесь регрессионные коэффициенты ( $b_1, b_2, b_p$ ) представляют собой вклады каждой независимой переменной в предсказание зависимой переменной. И в том, и в другом случае зависимость между переменными описывается линейным уравнением, поэтому регрессионному анализу подвергаются данные, зависимость между которыми носит линейный характер.

Линия регрессии выражает наилучшее предсказание зависимой переменной ( $Y$ ) по независимым переменным ( $X$ ). Однако на практике очень редко природные явления бывают полностью предсказуемыми, поэтому на графике наблюдается существенный разброс точек относительно идеально построенной прямой. Отклонение отдельной точки от линии регрессии называется остатком. Чем меньше разброс значений остатков около линии регрессии по отношению к общему разбросу значений, тем более точным является прогноз.

## 4.2. Обработка данных и представление результатов

При работе методом регрессионного анализа данные в электронную таблицу *Statistica* вводятся следующим образом: в первый столбец вводятся количественные значения зависимой переменной ( $Y$ ), в остальные – значения измеренных параметров, на основе которых предполагается прогнозирование поведения зависимой переменной (табл. 10).

Таблица 10

Пример расположения данных в электронной таблице *Statistica* при работе методом регрессионного анализа

Обилие изучаемого вида	Параметр 1	Параметр 2	Параметр 3	Параметр 4
100,2	11,4	9,47	13	13,8
450,6	60,8	64,3	62,3	56,1
320,5	29,8	23,6	25,5	58,4
50,1	1,3	2,95	2,1	2,1
40,6	1,1	2,2	1,9	2,6
40,1	1,1	1,9	1,5	0,9
32,5	0,8	1,6	1,2	0,7
480,9	62,4	36,4	2,2	2,4
25,5	1,2	1,9	1,2	1,3
20,4	0,9	0,1	0,1	0,8

В ходе применения анализа множественной регрессии на экран выводится информационное окно, в котором содержатся количественные значения промежуточных расчетов коэффициентов и полная информация об анализе (табл. 11).

Таблица 11

**Информационное окно регрессионного анализа программы *Statistica***

Множественная регрессия	
Dependent: Обилие	
Multiple R = 0,99932414	F = 923,8120
	R <sup>2</sup> =0,99864874 df =4,5
No. of cases: 10	adjusted R <sup>2</sup> =0,99756773 p=0,000000
Standard error of estimate: 9,170113679	
Intercept: 24,858626449	Std.Error: 3,690032
t(5) = 6,7367	p = 0,0011
<b>Параметр 1 beta=0,616</b> <b>Параметр 2 beta=0,551</b> <b>Параметр 3 beta=-0,57</b> <b>Параметр 4 beta=0,393</b>	

Примечание. Dependent – наименование зависимой переменной (в данном случае это обилие); No. of cases – число измерений каждого параметра, по которым построена регрессия (в данном примере – 10); Multiple R – коэффициент множественной корреляции; R<sup>2</sup> – коэффициент детерминации; adjusted R<sup>2</sup> – квадрат коэффициента детерминации; Standard error of estimate – стандартная ошибка оценки; Intercept – оценка свободного члена регрессии (константа а в уравнении); Std.Error – стандартная ошибка свободного члена; t(5) and p – t-критерий и уровень значимости (t-критерий используется для проверки гипотезы о равенстве нулю свободного члена регрессии); beta – стандартизированный коэффициент регрессионного уравнения. Значимые стандартизированные коэффициенты выделены полужирным шрифтом.

Анализируем значение коэффициента детерминации (R<sup>2</sup>), который показывает, какой процент вариации зависимой переменной объясняют независимые переменные, и выявляем параметры, которые в дальнейшем войдут в регрессионное уравнение (значения стандартизированных коэффициентов *beta* для них выделены полужирным шрифтом).

В данном случае R<sup>2</sup> = 0,99. Значит, 99% изменений обилия изучаемого вида может быть объяснено анализируемыми параметрами 1, 3 и 4.

Конечные результаты регрессионного анализа выдаются в виде таблицы, в которой представлены количественные значения всех коэффициентов регрессии (табл. 12).

Таблица 12

**Пример итоговой таблицы регрессионного анализа  
в программе *Statistica***

	<i>Beta</i>	Std.Err. of <i>Beta</i>	<i>b</i>	Std.Err. of <i>b</i>	p-level
Intercept (a)			<b>24,85863</b>	<b>3,690032</b>	<b>0,001093</b>
Параметр 1	<b>0,615589</b>	<b>0,210812</b>	<b>4,54899</b>	<b>1,557823</b>	<b>0,033010</b>
Параметр 2	0,551497	0,312302	4,83775	2,739526	0,137671
Параметр 3	<b>-0,573242</b>	<b>0,181405</b>	<b>-5,41548</b>	<b>1,713757</b>	<b>0,025095</b>
Параметр 4	<b>0,392735</b>	<b>0,064443</b>	<b>3,15095</b>	<b>0,517028</b>	<b>0,001721</b>

Примечание. Intercept (a) – свободный член регрессии; *Beta* – стандартизированный коэффициент регрессионного уравнения (коэффициент при независимой переменной); Std.Err. of *Beta* – стандартная ошибка *Beta*; *b* – угловой коэффициент для каждой переменной; Std.Err. of *b* – стандартная ошибка углового коэффициента; p-level – уровень значимости; выделены значимые коэффициенты при уровне значимости  $p \leq 0,05$ .

Из этой итоговой таблицы в уравнение регрессии берутся данные по количественным значениям константы и *b*-коэффициентов с учетом их знака. При этом в уравнение включаются только те параметры, для которых уровень значимости  $p \leq 0,05$ . Итоговое уравнение регрессии имеет вид

$$\text{Обилие} = 24,8 + 4,55 \text{ Параметр 1} - 5,41 \text{ Параметр 3} + 3,15 \text{ Параметр 4}$$

По итогам применения регрессионного анализа можно сделать вывод, что из 4 проанализированных независимых переменных вклад в предсказание поведения зависимой переменной (обилие) вносят 3 параметра, объясняя 99% вариации зависимой переменной.

Для дальнейшей возможности достоверного прогнозирования поведения зависимой переменной при изменении независимых переменных составление регрессионного уравнения бывает недостаточным. Зачастую необходима проверка прогностической модели на выборках большего размера с возможным исключением коррелирующих между собой параметров.

### 4.3. Варианты заданий для самостоятельного анализа

#### ВАРИАНТ 1

На территории скверов изучалось состояние крон хвойных пород деревьев. С этой целью было проведено измерение протяженности и диаметра кроны у разновозрастных елей.

Диаметр кроны, м: 1,1; 3,1; 1,6; 2,1; 4,6; 3,7; 5,6; 6,3; 4,0; 4,1; 3,6; 5,6; 6,1; 6,0; 4,4; 4,6; 1,0; 1,2; 1,6; 2,1; 2,3.

Протяженность кроны, м: 7,0; 10,5; 7,2; 7,0; 9,5; 13,0; 10,0; 7,5; 7,4; 9,4; 12,8; 12,0; 12,1; 8,1; 8,0; 6,0; 6,0; 6,2; 7,2; 7,3.

Возраст дерева, лет: 21, 37, 35, 36, 42, 46, 44, 41, 45, 45, 46, 43, 44, 44, 42, 42, 21, 22, 35, 35, 35.

Для имеющихся данных напишите уравнение зависимости диаметра кроны от ее протяженности и возраста дерева, предполагая, что связь носит линейный характер.

Сколько процентов вариации зависимой переменной объясняют независимые переменные, включенные в модель?

Чему будет равен диаметр кроны при ее протяженности 10 м и возрасте дерева 100 лет?

#### ВАРИАНТ 2

Для исследования возможности применения в городском озеленении одного из видов камнеломок изучалась успешность роста данного вида в зависимости от различных характеристик субстрата. В качестве показателя успешности роста использовалась величина проективного покрытия растения на 15 опытных участках.

Проективное покрытие камнеломки, %: 20, 30, 10, 90, 50, 60, 20, 30, 50, 60, 50, 20, 10, 60, 80.

Влажность субстрата, %: 50, 20, 30, 5, 6, 6, 35, 20, 6, 4, 4, 18, 30, 6, 4.

Твердость субстрата, кг/см<sup>2</sup>: 10,6; 10,3; 10,8; 10,1; 11,2; 12,2; 10,7; 13,9; 14,2; 10,1; 11,1; 13,8; 11,9; 14,1; 11,2.

Воздухопроницаемость субстрата, %: 10, 20, 30, 80, 40, 40, 10, 20, 50, 45, 35, 15, 10, 50, 70.

По итогам регрессионного анализа выявите наличие зависимости проективного покрытия камнеломки от характеристик субстрата. Напишите уравнение регрессии.

Сколько процентов вариации зависимой переменной объясняют независимые переменные, включенные в модель?

Чему будет равно проективное покрытие камнеломки при значении воздухопроницаемости субстрата 95?



## 5. КЛАСТЕРНЫЙ АНАЛИЗ

### 5.1. Цель и задачи

Кластерный анализ – группа методов статистической обработки данных, которая включает различные классификации объектов на основе их сходства. Кластерный анализ «сжимает» информацию, уменьшая число рассматриваемых объектов, поскольку исходная совокупность объектов разбивается на кластеры – группы объектов, обладающие сходными свойствами. При постановке задач классификации число кластеров, на которое должно быть разбито исходное множество объектов, может задаваться заранее или выявляться в процессе решения.

Кластерный анализ решает следующие задачи:

- 1) разработки типологии или классификации;
- 2) исследования полезных концептуальных схем группирования объектов;
- 3) представления гипотез на основе исследования данных;
- 4) проверки гипотез для определения, действительно ли типы (группы), выделенные тем или иным способом, присутствуют в имеющихся данных.

В экологических исследованиях кластерный анализ применяется для решения широкого спектра задач, но чаще всего его используют именно с целью классификации. Назначение кластерного анализа состоит в объединении объектов, характеризующихся большим количеством признаков, в достаточно большие кластеры, с учетом некоторой меры сходства или расстояния между объектами.

Выделяют две группы методов кластерного анализа: иерархические и неиерархические. Основными методами иерархического кластерного анализа являются метод ближнего соседа, метод полной связи, метод средней связи и метод Варда. Неиерархических методов больше, хотя работают они на одних и тех же принципах. Между собой методы различаются выбором начальной точки, правилом формирования новых кластеров и правилом остановки.

### 5.2. Обработка данных и представление результатов

Анализ данных методом кластерного анализа начинается с введения их в электронную таблицу *Statistica*. Данные вводятся следующим образом: по горизонтали располагаются объекты, которые необходимо классифицировать, по вертикали – количественные значения признаков этих объектов (табл. 13).

**Расположение данных в электронной таблице *Statistica*  
при работе методом кластерного анализа**

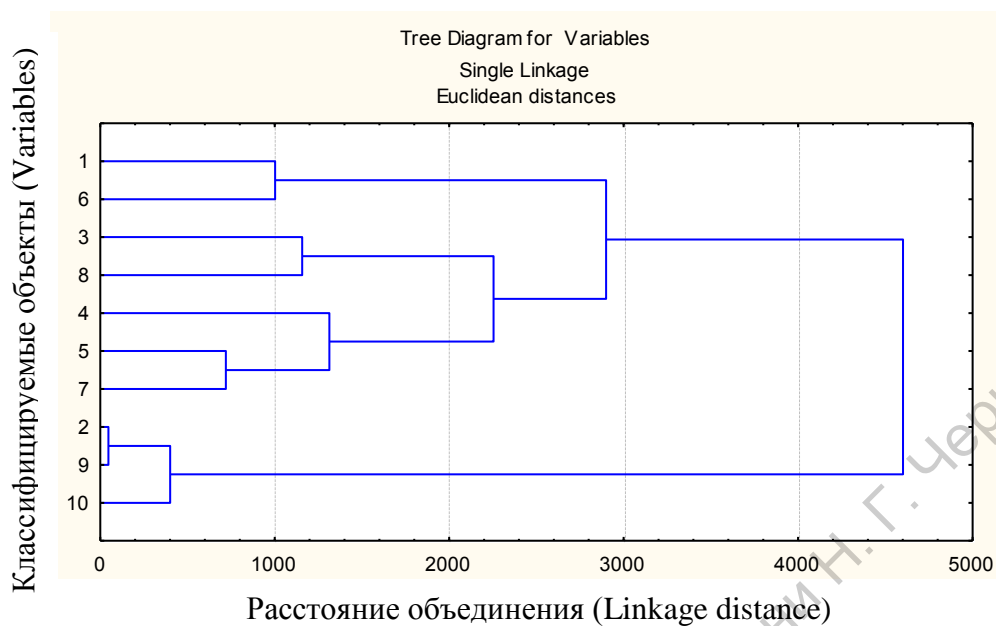
Характеристики классифицируемых объектов	Объект 1	Объект 2	Объект 3	Объект 4	Объект 5	Объект 6	Объект 7	Объект 8
Признак 1	7,6	11,4	9,47	13	13,8	12,3	15,1	17,5
Признак 2	63,9	60,8	64,3	62,3	56,1	62,4	54,9	48,9
Признак 3	11,9	29,8	23,6	25,5	58,4	36,4	62,4	70
Признак 4	1,9	1,3	2,95	2,1	2,1	2,2	3,1	3,5
Признак 5	2,6	1,1	2,2	1,9	2,6	2,4	2,59	2,9
Признак 6	0,7	1,1	1,9	1,5	0,9	0,9	1,2	1,9
Признак 7	0,6	0,8	1,6	1,2	0,7	1	1,5	1,1
Признак 8	12,3	62,4	36,4	2,2	2,4	0,9	1,0	0,3
Признак 9	0,3	1,2	1,9	1,2	1,3	4,5	1,2	1,3
Признак 10	0,3	0,9	0,1	0,1	0,8	0,9	0,5	0,4

При помощи модуля "Анализ кластера" задаются необходимые параметры кластеризации (метод построения кластера, мера расстояния объединения, характер диаграммы).

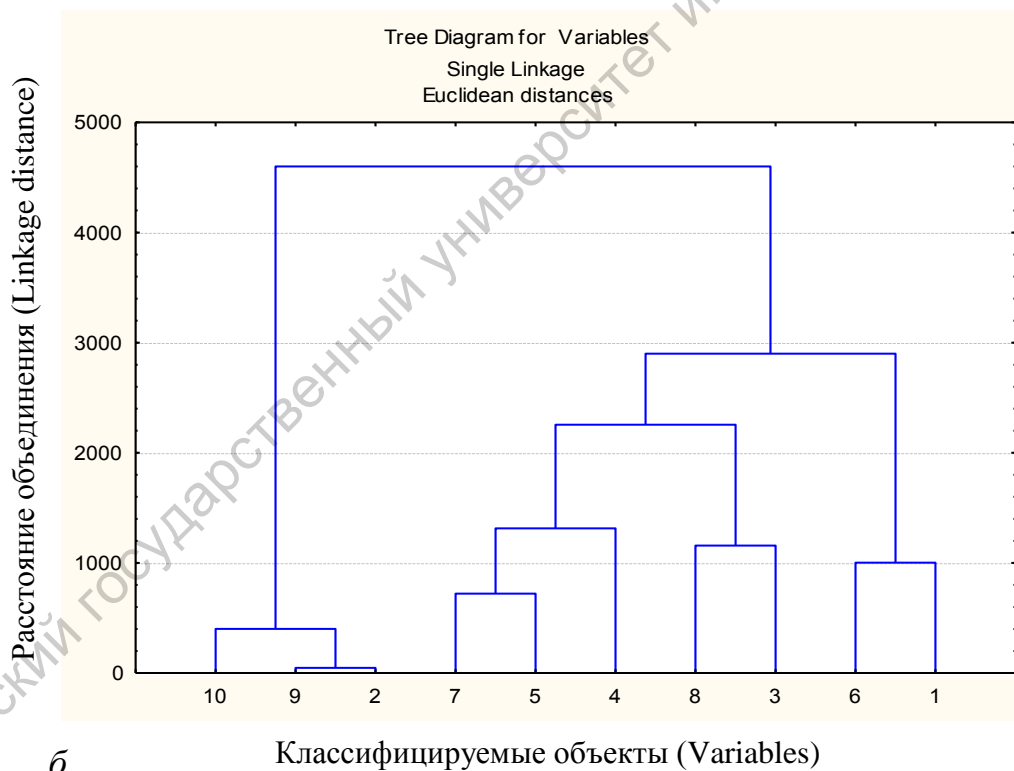
Результатом процедуры кластеризации является построение дендрограммы, которая является графическим изображением результатов процесса последовательной кластеризации, осуществляемой посредством матрицы расстояний. Существует большое количество методов построения дендрограмм и множество возможных мер измерения расстояний между классифицируемыми объектами. Наиболее часто используемыми методами объединения объектов являются метод ближнего и метод дальнего соседа.

Различают горизонтальную и вертикальную древовидные диаграммы (рис. 3). Построение диаграммы начинается с разбиения совокупности объектов на несколько классов и определения первого объекта в классе. При присоединении каждого следующего объекта снижается критерий его уникальности, таким образом, каждый последующий объект по совокупности анализируемых признаков все менее и менее похож на первый.

В процессе объединения связывается всё большее число объектов и объединяется все больше кластеров, состоящих из все сильнее различающихся элементов. На последнем шаге все объекты объединяются вместе. В результате анализа появляется возможность обнаружить кластеры и интерпретировать их. В экологических исследованиях не рекомендуется объединять в одной диаграмме более 30 объектов, поскольку это значительно затрудняет интерпретацию результатов и не дает возможности наглядного представления данных.



а



б

Рис. 3. Графическое изображение результатов кластерного анализа в программе *Statistica*: а – горизонтальная дендрограмма; б – вертикальная дендрограмма. В верхней части графика указывается тип диаграммы (древовидная диаграмма – Tree Diagram for Variables), метод построения кластера (одиночная связь – Single Linkage), мера объединения (Евклидово расстояние – Euclidean distances)

### 5.3. Методика объяснения результатов

Интерпретация результатов классификации начинается с выделения на диаграмме визуально четко обособленных кластеров и нахождения межкластерного расстояния между ними. В пределах крупного кластера на основе содержательных соображений выделяются подгруппы близких между собой объектов, и отмечается межкластерное расстояние, дающее представление о степени сходства или различия между парами объектов.

При наличии в кластере объектов, не вошедших ни в один из крупных кластеров, проводится описание расстояния их отхождения от основных кластеров и поиск (путем анализа количественных значений, характеризующих объекты признаков) возможного критерия уникальности этих объектов, не дающего возможности говорить о сходстве их с остальными классифицируемыми объектами.

### 5.4. Пример обработки данных методом кластерного анализа

В пределах разных по породному составу растительных сообществ проведено измерение девяти различных фитоценологических параметров (табл. 14). Необходимо определить, какие сообщества являются наиболее сходными по структурным характеристикам, а какие обладают наибольшими отличиями.

Таблица 14

Среднее значение фитоценологических параметров различных вариантов растительных сообществ

Растительное сообщество	Древостой			Подрост		Подлесок		Травостой	
	Высота, м	Диаметр стволов, см	Сомкнутость крон, %	Высота, м	Диаметр стволов, см	Высота, м	Диаметр стволов, см	Высота, м	Покрытие, %
Дубрава остепненная	7,6	63,9	11,9	1,9	2,6	0,7	0,6	0,32	25,8
Дубрава дубравно-мятликовая	11,4	60,8	29,8	1,3	1,1	1,1	0,8	0,27	30
Дубрава разнотравная	9,47	64,3	23,6	2,95	2,2	1,9	1,6	0,47	45,9
Дубрава ландышевая	13,0	62,3	25,5	2,1	1,9	1,5	1,2	0,62	75,3
Липо-дубрава ландышевая	13,8	56,1	58,4	2,1	2,6	0,9	0,7	0,6	63,4
Липо-дубрава крапивная	12,3	62,4	36,4	2,2	2,4	0,9	1,0	0,3	32,7
Липняк снытевый	15,1	54,9	62,4	3,1	2,59	1,2	1,5	0,58	66,4
Липо-кленовник снытевый	17,5	48,9	70,0	3,5	2,9	1,9	1,1	0,25	32,5
Березняк разнотравный	16,0	50,1	62,5	3,0	1,9	1,6	1,2	0,24	70,8
Осинник ландышевый	15,0	49,3	30,2	1,5	1,2	1,1	0,3	0,36	60,2
Сосняк мертвопокровный	10,0	49,3	30,0	0,5	0,3	1,1	0,3	0,25	5,01

В результате применения кластерного анализа (использовался метод ближнего соседа, мера расстояния между объектами – евклидово расстояние) была получена следующая дендрограмма (рис. 4.).

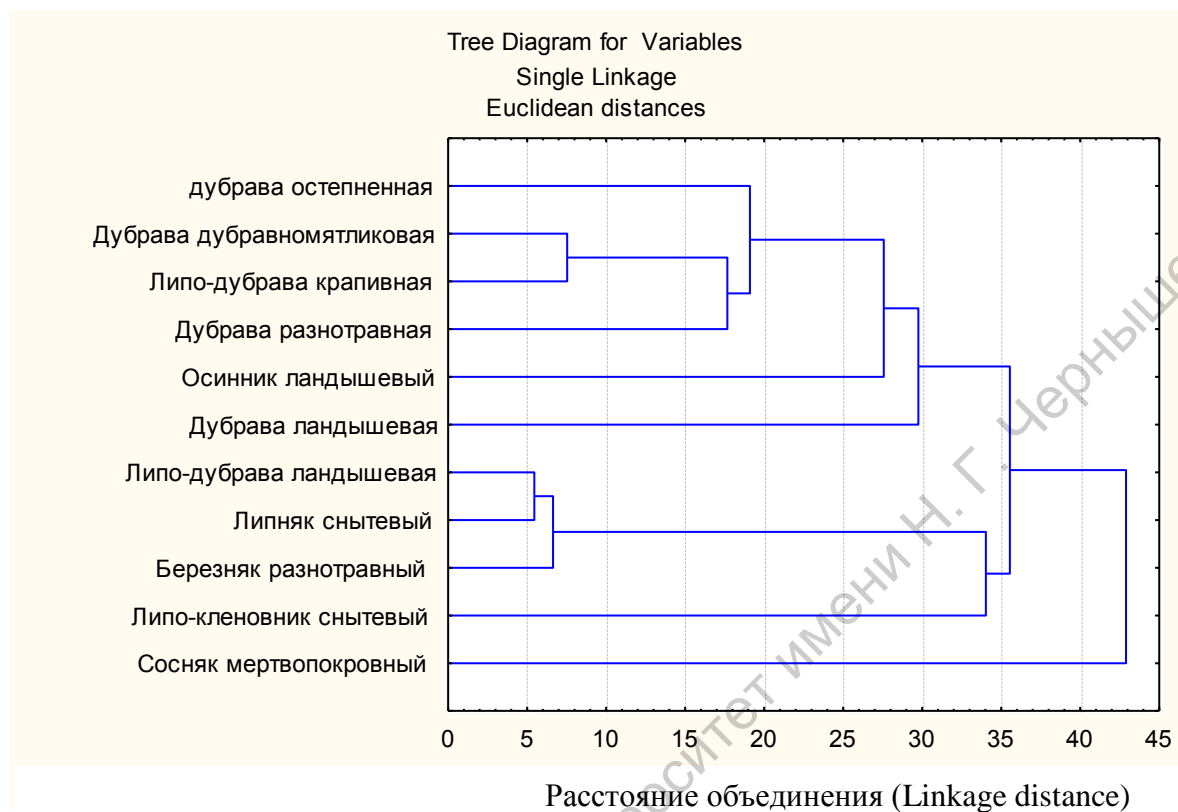


Рис. 4. Вертикальная кластерная диаграмма, объединяющая в группы растительные сообщества по сходству фитоценологических параметров

На кластерной диаграмме выделяются два кластера. Первый образуют дубравы остепненная, дубравномятликовая, разнотравная и липо-дубрава крапивная. Эти сообщества являются сходными по структуре между собой в большей степени, чем с любыми другими сообществами на диаграмме. Во второй кластер вошли липо-дубрава ландышевая, липняк снытевый и березняк разнотравный. Они являются наиболее похожими по комплексу фитоценологических характеристик (межкластерное расстояние равно 7). Осинник и дубрава ландышевые в некоторой степени схожи с сообществами первой группы. Наиболее отличающимся по структурным особенностям от остальных сообществ является сосняк мертвopoкpoвный (расстояние отхождения от остальных объектов равно 43). Такое положение данного растительного сообщества на кластерной диаграмме связано, вероятно, с очень низкими значениями высоты и диаметра подростa и проективного покрытия травостоя. Таким образом, из всего многообразия изученных лесных сообществ четко выделяются две группы фитоценозов, члены которых по структурным особенностям в наибольшей степени сходны. Оставшиеся четыре сообщества по комплексу структурных характеристик не имеют аналогов среди изученных фитоценозов.

## 5.5. Варианты заданий для самостоятельного анализа

### Общее задание к занятию

На территории Ботанического сада проводились исследования по успешности акклиматизации разных видов растений из различных климатических зон. По результатам акклиматизации необходимо было выяснить, какие условия являются оптимальными для развития изучаемых видов растений. Для этого было проведено измерение морфологических признаков растений каждого вида из 10 разных природных популяций, развивающихся в различных климатических условиях. Основными диагностическими признаками успешности акклиматизации служили высота и диаметр куста растений, количество вегетативных и генеративных побегов, облиственность побега, длина и ширина листьев, диаметр соцветия. Кроме того, в зависимости от особенностей того или иного вида растений проводились некоторые дополнительные промеры.

Используя метод кластерного анализа, требуется определить, растения каких популяций являются наиболее сходными по совокупности морфологических признаков, какие популяции отличаются в наибольшей степени.

Сделать вывод.

### Последовательность работы

- занести в таблицу данные по измеряемым морфологическим признакам, обозначив номер популяции и наименование измеряемого признака;
- с помощью многомерных статистик построить кластерную диаграмму одним из предложенных методов, используя указанную меру измерения расстояния между объектами;
- интерпретировать полученные результаты, ответив на вопросы:
  1. Какой метод построения кластеров используется?
  2. Сколько крупных кластеров выделяется на диаграмме?
  3. Какие объекты являются наиболее похожими?
  4. Чему равно межкластерное расстояние между ними?
  5. Какие объекты обладают наименьшим сходством со всеми остальными?
  6. Чему равно межкластерное расстояние между ними?
  7. Можно ли говорить о том, что независимо от климатических условий морфологические признаки растений одного вида в разных популяциях являются стабильной величиной?
  8. Наблюдается ли сходное развитие морфологических признаков растений изучаемого вида независимо от разницы условий произрастания?

## ВАРИАНТ 1

В табл. 15 представлены количественные значения морфологических признаков энотеры миссурийской. Построить кластерную диаграмму методом дальнего соседа (мера объединения – квадрат евклидова расстояния).

Таблица 15

### Количественные значения морфологических признаков энотеры миссурийской

Признаки	Номер популяции									
	1	2	3	4	5	6	7	8	9	10
Диаметр куста, см	58,7	47,5	46,2	47,3	49,3	58,8	52,1	38,4	41,6	32,9
Высота растения, см	47,0	31,6	29,1	36,5	29,0	25,3	29,6	21,8	25,1	32,6
Число вегетативных побегов, шт.	5,0	19,0	6,8	12,5	5,2	5,9	6,0	2,1	4,2	3,7
Облиственность побега, шт.	19,7	10,9	10,6	11,9	11,2	16,5	18,2	11,5	16,7	12,3
Длина среднего листа, см	12,5	15,8	11,2	19,2	10,8	9,6	11,2	11,2	13,4	10,0
Ширина среднего листа, см	2,4	1,7	3,1	1,9	3,1	2,8	2,8	4,1	2,1	2,3

## ВАРИАНТ 2

В табл. 16 представлены количественные значения морфологических признаков очитка большого. При построении кластерной диаграммы использовать метод ближнего соседа (мера объединения – евклидово расстояние).

Таблица 16

### Количественные значения морфологических признаков очитка большого

Признаки	Номер популяции									
	1	2	3	4	5	6	7	8	9	10
Высота растения, см	42,0	43,4	45,8	46,8	47,6	47,6	47,4	50,2	45,4	49,1
Число побегов, шт.	39,9	36,6	42,3	39,1	45,1	42,9	42,5	38,6	39,6	42,1
Диаметр растения, см	34,6	33,8	36,6	35,8	38,3	38,2	41,8	42,4	53,8	43,6
Число листьев на побеге, шт.	14,7	15,3	17,3	17,3	17,3	17,3	17,3	17,3	15,9	15,1
Длина листа, см	5,2	5,3	5,75	5,5	3,3	2,5	3,3	2,5	5,3	3,2
Ширина листа, см	2,5	2,6	3,1	3,9	1,6	1,4	1,7	1,8	5,6	1,8

### ВАРИАНТ 3

В табл. 17 представлены значения морфологических признаков очитка гибридного. При построении кластерной диаграммы использовать метод ближнего соседа (мера объединения – квадрат евклидова расстояния).

Таблица 17

#### Количественные значения морфологических признаков очитка гибридного

Признаки	Номер популяции									
	1	2	3	4	5	6	7	8	9	10
Высота растения, см	22,0	33,4	25,8	46,8	51,6	47,6	47,4	50,2	45,4	49,1
Число побегов, шт.	52,9	36,6	29,3	35,1	45,1	42,9	42,5	38,6	39,6	42,1
Диаметр растения, см	29,6	33,8	25,6	35,8	31,3	38,2	28,8	49,4	53,8	43,6
Число листьев на побеге, шт.	14,7	15,3	17,3	19,3	17,3	17,3	16,3	13,3	15,9	15,1
Длина листа, см	5,2	5,3	5,75	5,5	6,3	2,5	2,3	2,5	5,3	3,2
Ширина листа, см	2,5	2,6	3,1	3,9	1,6	1,9	2,7	1,8	5,6	2,8

### ВАРИАНТ 4

В табл. 18 представлены количественные значения промеров длины и ширины листьев алтея армянского. При построении кластерной диаграммы использовать метод дальнего соседа (мера объединения – квадрат евклидова расстояния).

Таблица 18

#### Количественные значения морфологических признаков алтея армянского

Признаки		Номер популяции									
		1	2	3	4	5	6	7	8	9	10
Верхний лист	длина, см	10,2	11,2	10,9	8,9	11,6	12,1	9,6	9,2	9,8	11,4
	ширина, см	9,2	7,0	8,6	8,9	9,1	10,0	8,5	8,2	7,1	8,9
Средний лист	длина, см	11,6	13,4	13,6	12,9	12,6	12,6	12,1	13,1	13,5	12,0
	ширина, см	11,0	9,4	9,4	11,2	9,6	10,0	9,5	9,1	9,9	9,0
Нижний лист	длина, см	12,2	11,5	12,5	12,0	13,0	11,5	11,8	13,4	14,0	12,9
	ширина, см	8,5	9,4	10,5	10,2	12,1	10,2	7,1	9,6	9,8	9,4



## ВАРИАНТ 5

В табл. 19 представлены количественные значения промеров длины и ширины листьев алтея лекарственного. При построении кластерной диаграммы использовать метод взвешенного попарного среднего значения (мера объединения – расстояние Манхеттена).

Таблица 19

**Количественные значения морфологических признаков алтея лекарственного**

Признаки		Номер популяции									
		1	2	3	4	5	6	7	8	9	10
Верхний лист	длина, см	9,8	9,2	10,2	8,9	11,6	12,1	9,6	9,2	10,0	10,1
	ширина, см	8,5	8,2	7,6	8,9	9,1	9,6	8,5	8,2	7,1	8,9
Средний лист	длина, см	13,0	12,9	13,6	12,9	12,6	12,6	12,1	13,1	13,5	12,0
	ширина, см	9,7	9,4	9,4	9,4	9,6	10,0	9,5	9,1	9,9	9,0
Нижний лист	длина, см	12,2	11,5	12,5	12,9	13,0	11,5	11,8	13,4	14,0	12,9
	ширина, см	9,8	9,4	10,5	10,2	12,1	10,2	9,5	9,6	9,8	9,4

## ВАРИАНТ 6

В табл. 20 представлены количественные значения морфологических признаков лаванды лекарственной. Построить кластерную диаграмму методом дальнего соседа (мера объединения – расстояние Чебышева).

Таблица 20

**Количественные значения морфологических признаков лаванды лекарственной**

Признаки	Номер популяции									
	1	2	3	4	5	6	7	8	9	10
Диаметр куста, см	41,2	47,5	46,9	47,3	49,3	58,9	52,1	38,4	41,6	32,9
Высота растения, см	19,6	31,6	21,6	36,5	29,0	25,3	29,6	21,8	25,1	32,6
Число вегетативных побегов, шт.	5,0	19,0	6,8	12,5	5,2	5,9	6,0	2,1	4,2	3,7
Облиственность побега, шт.	15,8	10,9	10,6	11,9	11,2	16,5	19,5	11,5	16,7	12,3
Длина среднего листа, см	13,9	15,8	11,2	19,2	10,8	9,6	11,2	11,2	13,4	11,1
Ширина среднего листа, см	1,2	1,7	3,1	1,9	3,1	2,8	2,8	4,1	2,1	2,3

## 6. ФАКТОРНЫЙ АНАЛИЗ

### 6.1. Цель и задачи

Факторный анализ – это многомерный метод статистического анализа данных, объединяющий методы оценки значимости множества наблюдаемых переменных посредством исследования структуры корреляционных матриц. Основное предположение факторного анализа заключается в том, что исходные показатели являются линейной комбинацией небольшого числа объективно существующих, но не поддающихся непосредственному измерению факторов, детерминирующих различия между объектами. Таким образом, корреляционные связи между большим числом наблюдаемых переменных определяются существованием меньшего числа гипотетических ненаблюдаемых переменных или факторов. Некоторые из этих факторов являются общими для двух и более переменных, другие характерны для каждого признака в отдельности.

Целью факторного анализа является построение классификации объектов путем сокращения числа объясняющих переменных. Основные задачи факторного анализа:

- 1) снижение размерности описания объектов путем замены значительного количества исходных показателей небольшим числом обобщенных факторов, сохраняющих достаточную информацию об исследуемых объектах;
- 2) определение структуры взаимосвязи между переменными.

Факторный анализ применяется к объектам, которые содержат большое количество характеризующих их переменных. Поэтому затруднена предварительная интерпретация тех переменных, которые вносят наибольший вклад в различия между объектами. Процедура трактовки результатов в факторном анализе состоит из двух этапов: оценки факторной структуры (числа факторов, необходимых для объяснения корреляционной связи между переменными) и факторной нагрузки, а затем установления самих факторов по результатам наблюдения.

### 6.2. Обработка данных и представление результатов

Ввод данных в электронную таблицу *Statistica* осуществляется следующим образом: по горизонтали располагаются переменные, характеризующие объекты, а по вертикали – классифицируемые объекты. При помощи модуля "Многомерные исследовательские методы → Анализ фактора" задаются необходимые параметры построения, выявляется количество факторов, вносящих наибольший вклад в дисперсию объектов, рассчитывается факторная нагрузка по каждому из значимых факторов.

Результатом процедуры факторного анализа является построение таблицы, отражающей значения факторной нагрузки на каждую из рассматриваемых переменных. По результатам выделения наиболее значимых

факторов, вносящих наибольший вклад в дисперсию объектов, проводится построение графика, отражающего распределение анализируемых объектов в пространстве выделенных значимых факторов. Как правило, построение графика ведется в системе XY координат, где по осям располагаются значения 2-х факторов, объясняющих наибольший процент изменчивости объектов (рис. 5). Но иногда различия между объектами обусловлены влиянием 3-х и более значимых факторов. В этом случае график строится в системе XYZ, где по осям откладываются значения трех наиболее значимых факторов.

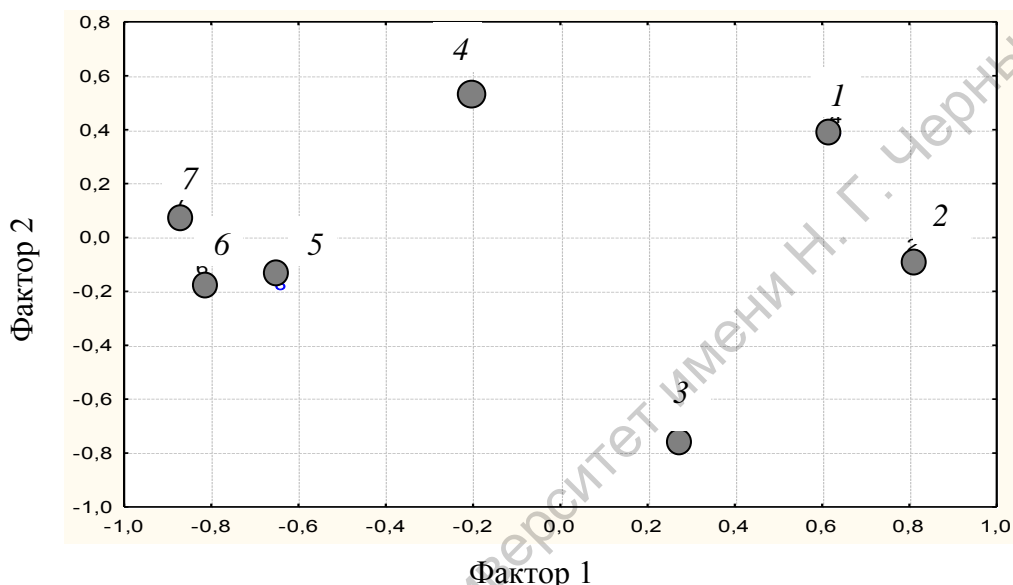


Рис. 5. Пример распределения классифицируемых объектов (1–7) в пространстве двух значимых факторов

### 6.3. Методика объяснения результатов

Интерпретация результатов факторного анализа начинается с рассмотрения таблицы факторной нагрузки (табл. 21).

Таблица 21

Пример распределения факторной нагрузки по переменным

Переменные	Factor 1	Factor 2	Factor 3
Переменная 1	0,109128	0,401590	0,123500
Переменная 2	-0,202718	0,634403	0,003251
Переменная 3	<b>0,832814</b>	-0,090545	0,012591
Переменная 4	0,241705	<b>-0,777051</b>	-0,002547
Переменная 5	0,649253	0,316393	0,021431
Переменная 6	<b>-0,797061</b>	-0,115767	0,325141
Переменная 7	<b>-0,879587</b>	0,090363	0,658213
Дисперсия	0,421140	0,246676	0,10236

В первую очередь анализируют значения дисперсии для каждого из выделенных факторов. Факторы, обладающие долей дисперсии менее 0,1, считаются не значимыми и в дальнейший анализ не включаются. В нашем случае таким фактором является Factor 3, доля дисперсии которого составляет всего 0,1. Таким образом, остается два значимых фактора – Factor 1 и Factor 2.

Затем по каждому из оставшихся значимых факторов отдельно анализируем факторные нагрузки, выделяя те, собственные значения которых по модулю превышают 0,7 (см. табл. 18). Для Factor 1 переменными, вносящими наибольший вклад в образование фактора, являются переменные 3, 6, 7 (собственные значения которых равны 0,83; – 0,79 и – 0,87 соответственно). В образование Factor 2 наибольший вклад вносит переменная 4 (собственное значение – 0,77).

По итогам факторного анализа делается вывод о количестве значимых факторов, объясняющих наибольший процент дисперсии изучаемых объектов, строится графическое распределение объектов в пространстве выделенных факторов, дается наименование этим факторам.

#### 6.4. Пример обработки данных методом факторного анализа

С целью изучения закономерностей структурной организации степных фитоценозов было проведено измерение 7 параметров в 40 растительных сообществах. В результате обработки данных методом факторного анализа получились следующие результаты (табл. 22).

Таблица 22

##### Факторная нагрузка по структурным параметрам фитоценозов

Параметры	Фактор 1	Фактор 2
Проективное покрытие травостоя	0.670	–0.221
Средняя высота травостоя	<b>0.711</b>	–0.204
Вертикальная плотность	–0.207	<b>0.794</b>
Горизонтальная неоднородность	<b>–0.734</b>	–0.179
Вертикальная неоднородность	<b>0.784</b>	–0.084
Коэффициент вариации высоты	–0.189	<b>–0.773</b>
Коэффициент вариации вертикальной плотности	–0.328	–0.552
<b>Дисперсия</b>	<b>0,41</b>	<b>0,20</b>

Анализ таблицы факторных нагрузок показал, что два первых фактора в сумме объясняют 61% изменений параметров растительности. Первый фактор объясняет 41% вариаций в матрице данных, второй – 20% вариаций.

Наивысшие положительные значения факторной нагрузки по фактору 1 установлены для двух переменных структуры (высота и вертикальная неоднородность фитоценоза). А отрицательную нагрузку имеет показатель горизонтальной неоднородности.

Таким образом, первый фактор отражает градиент от фитоценозов с горизонтально неоднородным растительным покровом к территориям с высоким и густым покровом, характеризующимся большими значениями вертикальной неоднородности.

В отношении второго фактора наибольшую нагрузку испытывают такие показатели, как вертикальная плотность травостоя (в положительном спектре), коэффициент вариации высоты травостоя (в отрицательном спектре), т.е. можно выделить градиент от фитоценозов, отличающихся большими колебаниями высоты растений, к фитоценозам с максимальными значениями вертикальной плотности травостоя.

На рис. 6 представлено расположение конкретных фитоценозов в пространстве этих двух факторов, что дает возможность визуальной оценки их сходства между собой по данным критериям.

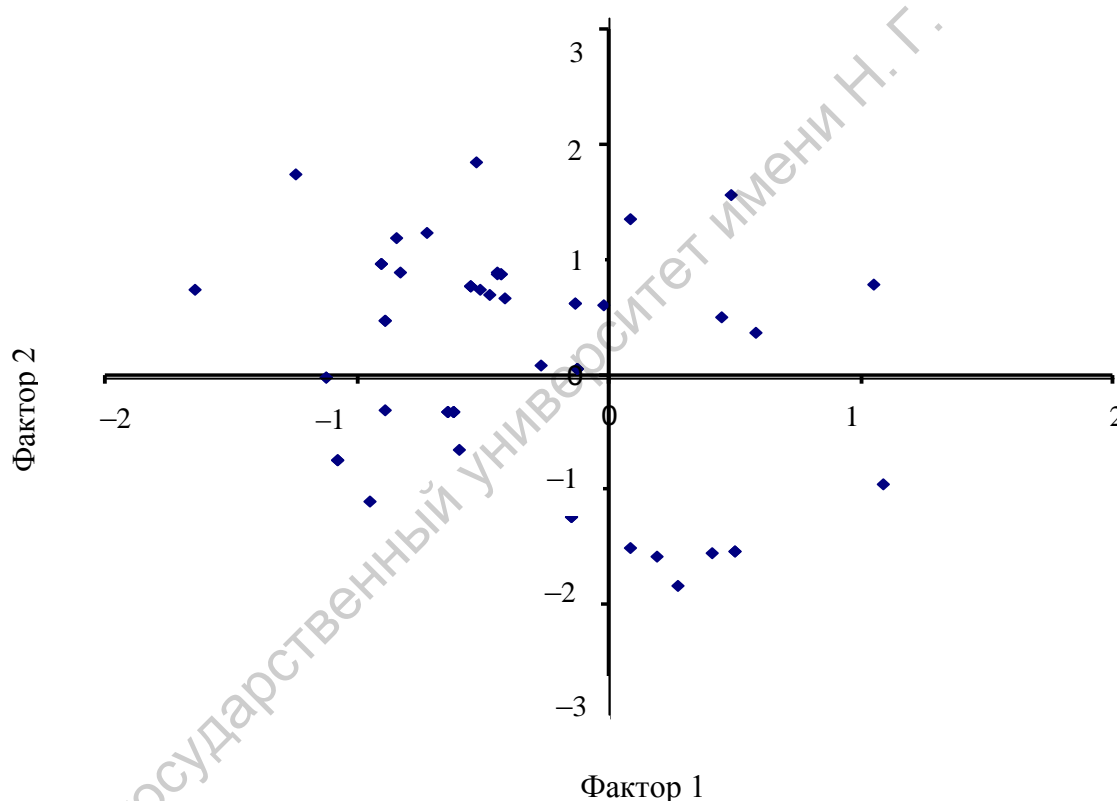


Рис. 6. Распределение классифицируемых объектов в пространстве двух значимых факторов

По результатам факторного анализа структурных параметров различных степных фитоценозов можно заключить, что изученные растительные сообщества в наибольшей степени различаются по признаку вертикальной неоднородности и вертикальной плотности травостоя. Все разнообразие структуры укладывается в рамки двух факторов, объясняющих 61% изменчивости строения фитоценозов.

## 6.5. Варианты заданий для самостоятельного анализа

### Общее задание к занятию

На территории лесных и степных участков проводилось изучение особенностей местообитаний различных видов растений и животных. С целью выявления наиболее значимых для них параметров среды обитания было проведено измерение ряда фитоценологических параметров.

Необходимо было выявить переменные, определенное количественное сочетание которых позволяет изучаемому виду поселяться на той или иной территории с достаточно высокой плотностью.

В табл. 23–26 представлены количественные данные по характеристике фитоценозов на участках обитания разных видов растений или животных. Используя процедуру "Факторный анализ":

1) определите количество факторов, в которые могут быть трансформированы исходные переменные;

2) определите переменные, вносящие наибольший вклад в образование каждого фактора;

3) постройте график распределения участков обитания растений или животных в пространстве двух ведущих факторов;

4) опишите общее направление изменения фитоценологических параметров при продвижении вдоль оси каждого фактора.

5) сделайте вывод о толерантности изучаемого вида по отношению к значениям факторов.

### Последовательность работы

1. Занести исходные данные в электронную таблицу.
2. Дать наименование столбцам таблицы.
3. Провести процедуру факторного анализа и проанализировать таблицу факторных нагрузок.
4. Получить таблицу координат участков обитания изучаемого вида в пространстве выделенных факторов.
5. Скопировать полученную таблицу координат и вставить ее в исходную таблицу данных.
6. Построить график распределения участков в пространстве выделенных значимых факторов.
7. Проанализировать полученные результаты.
8. Дать наименование выделенным значимым факторам.
9. Сделать вывод.

## ВАРИАНТ 1

На территории лесного массива, находящегося в зоне постоянного рекреационного использования, проводилось изучение местообитаний редких видов растений. В ходе исследований было выявлено 39 участков обитания рябчика русского и в их пределах проведено описание фитоценологических параметров (табл. 23). Необходимо было установить, насколько толерантен данный вид растений к изменению структурных особенностей лесных фитоценозов.

Проанализировав результаты факторного анализа ответьте на вопрос: однородны ли по структуре участки, на которых может существовать этот вид?

Таблица 23

### Количественные значения фитоценологических параметров на участках обитания рябчика русского

№ участка	Высота древостоя, м	Сомкнутость крон древостоя, %	Высота подроста, м	Высота подлеска, м	Покрытие травостоя, %	Число видов травостоя	Покрытие подлеска, %
1	16,6	72	3,36	1,81	86,7	7	35
2	9,7	60	3,72	1,33	60	3	30
3	7,2	60	4,77	1,31	59	4	20
4	6,1	79	6,42	1	52,4	3	20
5	6,1	79	6,42	1	52,4	3	30
6	11,9	91,3	4,39	1	23,5	6	30
7	6,1	91,3	4,4	1	52,5	6	20
8	6,1	54,2	4,31	1,1	52,2	6	25
9	11,9	54,2	2,27	1,1	23,5	2	25
10	11,9	87,9	2,27	1,1	23,5	2	40
11	10,7	87,9	4,43	1,56	78,4	3	20
12	17,3	60,5	5,09	2,17	97,2	8	30
13	13,1	28,2	1,48	1,46	62,3	4	20
14	5,6	60,5	1,8	1	19,6	2	30
15	13	94,2	1,48	1,47	62,3	4	40
16	11,4	53,9	3,47	1,99	70,5	8	45
17	22,1	87,9	2,19	1,37	40,9	3	60
18	21,6	87,9	3,21	1,39	71,4	4	80
19	21,6	87,9	3,21	1,39	71,4	3	80

20	21,6	80,9	3,21	1,39	71,4	3	50
21	20,6	67,2	2,7	1,62	74,3	3	50
22	20,6	50,9	2,57	1,79	74,3	4	25
23	32,6	39,9	2,11	1,55	47,8	2	60
24	16,1	49,5	2,35	2,43	26,9	2	60
25	21,7	50	2,94	2,92	10	2	40
26	43,9	85,5	3,07	3,4	20,9	8	60
27	34,6	68,3	1,7	3,12	25,7	8	50
28	31,3	57	1,84	2,89	39,9	7	80
29	47,6	56,7	1,79	2,88	35,5	13	60
30	36,3	56,9	1,53	2,54	46	15	50
31	42,9	47,8	1,61	2,83	40,6	9	50
32	32,7	42,1	1,52	3,41	32,7	3	70
33	36,6	37,7	1,76	2,15	28,5	6	60
34	36,5	39,3	1,1	2,65	51,6	8	80
35	23,2	37,7	2,01	2,26	59,1	8	60
36	19,7	32,5	1,76	1,92	7	6	40
37	23,1	37,7	2,32	2,6	29,1	8	50
38	25,3	34	2,42	2,6	34,1	7	55
39	36,6	29	1,45	2,05	36,2	5	40

## ВАРИАНТ 2

В подзоне опустыненных степей проводилось изучение структуры местообитаний разных видов жаворонков. В ходе исследований было выявлено 30 участков обитания полевого жаворонка. На этих участках, расположенных в пределах разных по видовому составу растительных сообществ, проведено измерение ряда фитоценологических параметров (табл. 24).

Проанализировав результаты факторного анализа, ответьте на вопрос: выбирает ли этот вид для гнездования структурно однородные участки или они характеризуются значительным варьированием структурных характеристик фитоценозов?



Таблица 24

## Значения фитоценологических параметров на участках обитания полевого жаворонка

№ участка	Гориз. неоднородность	Высота доминанта, см	Высота травостоя, см	Коэфф. вариации высоты травостоя	Густота травостоя, экз./10 см <sup>2</sup>	Проективное покрытие травостоя, %	Проективное покрытие доминанта, %
1	2,73	1	6	29,3	2	50	20
2	1,44	3	37,5	23,2	11,8	47	65
3	2,72	1,78	18,7	88,2	8,1	100	40
4	2,5	1,46	13,2	81,1	3,4	61	80
5	2,5	1,46	13,2	81,1	3,5	60	60
6	1,68	3,39	32,2	32,9	10,9	45,8	70
7	2,31	1,17	10,6	85,9	1,9	58	30
8	3,75	1,29	10,8	99,1	1,6	56	20
9	3,61	1,89	23,5	71,5	7,5	96	30
10	3,61	1,89	23,5	71,5	7,5	96	30
11	1,97	2,62	39,5	30,5	7,7	60	70
12	2,47	1,9	31,3	54,7	4,5	80	40
13	2,46	2,42	23,4	72,2	7,6	98	30
14	1,05	2,47	22,3	57,4	7,6	30	30
15	1,79	2,25	31,1	54,3	5,8	59	60
16	2,3	2,62	43,5	51	7,1	91	70
17	3,4	1,87	27,2	66,5	2,9	97	50
18	3,3	1,47	17,7	55,9	2,9	82	40
19	2,98	1,82	22,7	74,4	3,9	82	40
20	2,8	1,8	23,2	71,9	3,6	75	60
21	1,9	2	26,4	77,6	5,2	78	40
22	2,33	1,37	17,8	20,8	3,8	65	50
23	3,85	2,22	28,6	22,3	8,4	59	60
24	1,82	2,08	28,7	40,2	6,3	40	70
25	1,51	3,22	36,9	27,1	9,2	39	70
26	2,04	2,35	17,3	79,1	7,2	48	60
27	1,76	1,79	24,3	51,4	2,8	58	40
28	0,78	2,56	28,6	22,3	8,9	39	80
29	1,83	1,71	21,6	85	7,1	78	60
30	1,83	1,72	21,6	85	7,2	78	50

### ВАРИАНТ 3

В байрачных дубравах проводилось изучение разнообразия видового состава травяного яруса. Было установлено, что на 50 лесных участках доминантом травостоя является мятлик дубравный. С целью выявления толерантности данного вида растений к изменению структурной сложности дубрав были проведены фитоценотические измерения участков его обитания (табл. 25).

Проанализировав результаты факторного анализа, ответьте на вопрос: может ли этот вид существовать на участках с достаточно широкой амплитудой изменения структурных характеристик фитоценозов?

Таблица 25

#### Количественные значения фитоценотических параметров на участках обитания мятлика дубравного

№ площадки	Густота древостоя, экз./га	Высота древостоя, м	Сомкнутость крон, %	Густота подроста, экз./га	Высота подроста, м	Густота подлеска, экз./га
1	625	16,5	64	900	1,76	0
2	533	16,8	47	1155	2,3	1422
3	600	16,8	35	1850	1,6	0
4	311	15,8	49	1200	1,7	0
5	355	19,76	57	666	2,3	0
6	900	10,3	70	1700	2,6	0
7	833	14	48	933	3,05	0
8	428	17,1	32	619	2,5	238
9	666	18	55	1288	2,3	366
10	519	16,4	46	1418	2,1	620
11	933	16,5	82	800	3	0
12	675	16,2	24	1100	1,9	0
13	845	17	39	434	3	0
14	640	15,6	37	220	3,1	0
15	448	17,9	46	288	3,2	0
16	350	13	46	1925	1,7	275
17	675	15	42	375	1,8	906
18	321	16	46	867	0,6	586
19	400	17,1	45	350	2,2	1632
20	426	12,3	39	213	6,9	566

21	453	15,6	66	453	2,5	366
22	888	19,8	53	800	1,4	280
23	491	17,4	54	1943	2,4	580
24	900	17,3	57	400	1,5	1325
25	700	17,1	47	333	4,7	450
26	1020	18,4	67	440	1,4	420
27	342	18,9	59	600	3,7	0
28	950	17,5	36	225	5,6	0
29	725	18	16	575	4,8	0
30	340	16	50	300	2,7	0
31	600	16,8	35	1850	1,6	0
32	311	15,8	49	1200	1,7	0
33	355	19,76	57	666	2,3	0
34	900	10,3	70	1700	2,6	0
35	833	14	48	933	3,05	0
36	428	17,1	32	619	2,5	238
37	666	18	55	1288	2,3	366
38	448	17,9	46	288	3,2	0
39	350	13	46	1925	1,7	275
40	675	15	42	375	1,8	906
41	321	16	46	867	0,6	586
42	400	17,1	45	350	2,2	1632
43	833	14	48	933	3,05	0
44	428	17,1	32	619	2,5	238
45	666	18	55	1288	2,3	366
46	491	17,4	54	1943	2,4	580
47	900	17,3	57	400	1,5	1325
48	700	17,1	47	333	4,7	450
49	1020	18,4	67	440	1,4	420
50	342	18,9	59	600	3,7	0

#### ВАРИАНТ 4

На территории Приволжской возвышенности проводилось изучение местообитаний различных видов птиц. В ходе этих исследований было выявлено 35 участков обитания обыкновенной овсянки, расположенных в разных по структурной сложности растительных сообществах. В табл. 26 приведены количественные значения измеренных параметров фитоценозов на данных участках.

Проанализировав результаты факторного анализа, ответьте на вопрос: насколько толерантна обыкновенная овсянка по отношению к выделенным факторам?

Таблица 26

**Количественные значения фитоценологических параметров на участках обитания обыкновенной овсянки**

№ участка	Диаметр стволов кустарников, см	Высота кустарников, м	Диаметр стволов деревьев, см	Сомкнутость крон, %	Число видов травостоя	Общее покрытие травостоя, %	Покрывание доминанта, %
1	1,45	2,05	36,6	36,2	5,9	29	55
2	3,53	1	3,3	45,4	5,8	76,1	40
3	2,73	1	6	29,3	2	50	20
4	1,44	3	37,5	23,2	11,8	47	65
5	2,72	1,78	18,7	88,2	8,1	100	40
6	2,5	1,46	13,2	81,1	3,4	61	80
7	2,5	1,46	13,2	81,1	3,5	60	60
8	1,68	3,39	32,2	32,9	10,9	45,8	70
9	2,31	1,17	10,6	85,9	1,9	58	30
10	3,75	1,29	10,8	99,1	1,6	56	20
11	3,61	1,89	23,5	71,5	7,5	96	30
12	3,61	1,89	23,5	71,5	7,5	96	30
13	1,97	2,62	39,5	30,5	7,7	60	70
14	2,47	1,9	31,3	54,7	4,5	80	40
15	2,46	2,42	23,4	72,2	7,6	98	30
16	1,05	2,47	22,3	57,4	7,6	30	30
17	1,79	2,25	31,1	54,3	5,8	59	60
18	2,3	2,62	43,5	51	7,1	91	70
19	3,4	1,87	27,2	66,5	2,9	97	50

Окончание табл. 26

20	3,3	1,47	17,7	55,9	2,9	82	40
21	2,98	1,82	22,7	74,4	3,9	82	40
22	2,8	1,8	23,2	71,9	3,6	75	60
23	1,9	2	26,4	77,6	5,2	78	40
24	2,33	1,37	17,8	20,8	3,8	65	50
25	3,85	2,22	28,6	22,3	8,4	59	60
26	1,82	2,08	28,7	40,2	6,3	40	70
27	1,51	3,22	36,9	27,1	9,2	39	70
28	2,04	2,35	17,3	79,1	7,2	48	60
29	1,76	1,79	24,3	51,4	2,8	58	40
30	0,78	2,56	28,6	22,3	8,9	39	80
31	3,75	1,29	10,8	99,1	1,6	56	20
32	3,61	1,89	23,5	71,5	7,5	96	30
33	3,61	1,89	23,5	71,5	7,5	96	30
34	1,97	2,62	39,5	30,5	7,7	60	70
35	2,47	1,9	25,9	54,7	4,5	30	40
36	2,46	2,42	23,4	72,2	7,6	98	30
37	1,05	2,47	22,3	57,4	7,6	30	30
38	1,79	2,25	31,1	54,3	5,8	59	50
39	3,75	1,4	16,9	99,1	1,6	56	20
40	2,15	1,89	23,5	71,5	9,7	60	30
41	3,61	1,89	23,5	71,5	7,5	96	30
42	1,97	1,69	39,5	30,5	5,9	60	70
43	2,47	1,9	31,3	54,7	4,5	80	40
44	2,46	2,42	23,4	72,2	7,6	56	30
45	2,36	2,47	36,5	57,4	7,6	30	35
46	3,12	2,25	31,1	54,3	5,8	59	60

## 7. ДИСКРИМИНАНТНЫЙ АНАЛИЗ

### 7.1. Цель и задачи

Дискриминантный анализ используется для принятия решения о том, какие переменные различают (дискриминируют) две или более возникающие совокупности (группы).

Основной целью дискриминантного анализа является классификация объектов на основании большого числа их количественных характеристик. Для осуществления данного анализа необходимо иметь одну качественную характеристику объекта и несколько количественных (не менее 5). Например, если проводится работа по изучению морфометрии какого-то вида растений, то в роли качественных характеристик могут выступать конкретный сорт, цвет, местообитание популяции; количественными критериями будут служить измерения длины, ширины листьев, высоты растения, диаметр куста и т.д.

В результате применения метода дискриминантного анализа строится классификационная матрица, в которой каждый объект отнесен к определенной группе и представлена оценка правильности этих отнесений. Анализируя полученную таблицу, можно судить об уникальности тех или иных количественных характеристик, строго соответствующих определенным качественным, или наоборот, о невозможности четкого разделения объектов на основе количественных параметров. Кроме того, дискриминантный анализ предоставляет возможность относить новый объект к той или иной группе на основе его характеристик.

### 7.2. Обработка данных и представление результатов

Ввод данных в электронную таблицу осуществляется следующим образом: по горизонтали – признаки объектов, по вертикали – классифицируемые объекты, в таком порядке, чтобы сначала располагались объекты, относящиеся к первой группе, потом – ко второй и т.д. (табл. 27).

Таблица 27

**Пример таблицы ввода данных при работе методом дискриминантного анализа**

№ группы	Объекты	Признак 1	Признак 2	Признак 3	Признак 4	Признак 5
I	1					
	2					
	3					
II	4					
	5					

Результатом применения процедуры дискриминантного анализа является таблица дискриминантной функции, в которой отражены линейные коэффициенты по всем переменным. На основе анализа таблицы выбираются переменные, имеющие наибольшие (по модулю) значения линейных коэффициентов и вносящие наибольший вклад в разделение групп. Обычно выделяют три переменные.

С учетом знака линейных коэффициентов записывается линейное уравнение дискриминантной функции ( $y = kx+b$ ) для каждой группы отдельно. Затем на экран выводится сводная таблица отнесений каждого объекта к той или иной группе (табл. 28). На ее основе оценивается процент правильных отнесений объектов к каждой группе. Если процент высок, то можно говорить о высокой предсказательной способности дискриминантной функции.

Таблица 28

**Пример классификационной матрицы результатов дискриминантного анализа**

Группы	Percent correct *	Группа 1 p = 0,02	Группа 2 p = 0,04
1	96,0000	24	1
2	100,0000	0	15
<b>Всего</b>	92,5000	24	16

Percent correct \* – процент правильных отнесений.

Из таблицы видно, что из 25 объектов первой группы 24 объекта были отнесены именно к этой группе, а 1 объект попал в группу 2. Процент правильных отнесений составляет 96%. Все 15 объектов второй группы были корректно отнесены именно к своей группе (процент правильных отнесений 100%). На основе этих данных можно сделать вывод, что дискриминантная функция дает возможность предсказания принадлежности объектов к группе 1 или группе 2 с высокой степенью точности.

### 7.3. Пример обработки данных методом дискриминантного анализа

На участках обитания двух экологически близких видов птиц – обыкновенного соловья и зарянки – изучалась структура растительных сообществ. Необходимо выяснить, разделение по каким параметрам фитоценозов позволяет видам сосуществовать на одной территории. Всего было описано 15 участков обитания соловья и 15 участков обитания зарянки.

По результатам дискриминантного анализа наибольшее расхождение видов в пространстве структуры растительных сообществ наблюдается по высоте подроста, проективному покрытию и высоте травостоя (табл. 29).

**Значение классификационной функции для 2-х групп объектов  
(1 – участки соловья, 2 – участки зарянки)**

Переменные	Участки соловья	Участки зарянки
Густота древостоя	0,0161	0,018
Высота древостоя	2,2258	0,267
Покрытие крон	0,8335	1,060
Густота подроста	0,0080	-0,001
Высота подроста	<b>7,7541*</b>	<b>-7,5126</b>
Густота подлеска	0,0054	0,004
Высота подлеска	-1,0047	-1,923
Высота травостоя	<b>-11,4254</b>	<b>-14,3215</b>
Покрытие травостоя	<b>5,6125</b>	<b>6,1045</b>
Constant	-92,5	-84,9

\* Наибольшие количественные значения линейных коэффициентов

По полученным данным уравнение дискриминантной функции имеет вид

$$y = -84,9 - 14,3x_1 + 7,5x_2 + 6,1x_3 - \text{для зарянки;}$$

$$y = -92,5 - 11,4x_1 + 7,7x_2 + 5,6x_3 - \text{для соловья;}$$

где  $x_1$  – высота травостоя;  $x_2$  – высота подроста;  $x_3$  – проективное покрытие травостоя.

Классификационная матрица, построенная по результатам дискриминантного анализа, дает возможность определения принадлежности участков обитания с известными параметрами растительности к тому или иному виду птиц с высокой степенью точности (100% – для соловья и 93,3% – для зарянки) (табл. 30). Из 15 изученных участков зарянки один участок был отнесен к группе участков соловьев. Для соловья все участки отнесены к своей группе.

Таблица 30

**Классификационная матрица *Statistica* по результатам дискриминантного анализа для двух групп объектов: участков соловья и участков зарянки**

Группы	Percent correct	Участки зарянки p=0,05	Участки соловья p=0,05
Участки зарянки	93,3333	14	1
Участки соловья	100,0000	0	15
Всего	96,6667	14	16

Затем на основе анализа первичных данных выявляем, что по сравнению с участками обитания зарянок, местообитания соловьев характе-



ризуются значительно более высокими показателями проективного покрытия и высоты травостоя, высоты подроста.

Таким образом, применение дискриминантного анализа показало, что участки обитания соловьев отличаются от участков обитания зарянок по трем основным параметрам. Знание количественных значений этих параметров на любом другом участке леса позволяет предсказывать возможность поселения здесь либо соловья, либо зарянки с высокой степенью точности.

#### 7.4. Варианты заданий для самостоятельного анализа

##### Общее задание к занятию

В табл. 32–35 представлены данные по средним значениям параметров структуры лесных растительных сообществ на участках обитания экологически близких видов птиц, предъявляющих сходные требования к структуре местообитаний.

Используя дискриминантный анализ, определите:

- 1) отличаются ли по структуре участки обитания изучаемых видов;
- 2) разделение по каким структурным параметрам позволяет данным видам сосуществовать на одной территории.

##### Последовательность работы

1. Занести в электронную таблицу данные по характеристикам структуры на участках разных видов птиц по следующей схеме (табл. 31).

Таблица 31

##### **Форма таблицы *Statistica* для занесения данных при работе методом дискриминантного анализа**

Виды птиц	Параметр 1	Параметр 2	Параметр 3	Параметр 4	Параметр 5
Вид 1					
Вид 1					
Вид 1					
Вид 1					
Вид 1					
Вид 2					
Вид 2					
Вид 2					
Вид 2					

2. Провести дискриминантный анализ и определить, какие параметры структуры фитоценозов вносят наибольший вклад в отличия между участками разных видов.
3. Написать уравнение дискриминантной функции для каждого вида.

4. Определить предсказательные возможности дискриминантной классификации для каждого вида (в %).
5. Сделать вывод.

#### ВАРИАНТ 1

В водораздельных липняках Приволжской возвышенности проводилось изучение структурных особенностей местообитаний большой синицы и обыкновенной лазоревки. Для этого были проведены количественные измерения некоторых фитоценологических параметров на участках обитания указанных видов птиц. Средние значения параметров по каждому из участков представлены в табл. 32–33.

По результатам применения дискриминантного анализа ответьте на вопрос: насколько сходны по структурным характеристикам участки обитания этих видов птиц в пределах одного растительного сообщества.

Таблица 32

**Количественные значения фитоценологических параметров на участках обитания большой синицы**

№ участка	Высота древостоя, м	Сомкнутость крон, %	Густота подроста, экз./га	Высота подроста, м	Густота подлеска, экз./га	Высота травостоя, м	Покрытие травостоя, %
1	16,5	64	900	1,76	0	0	0,25
2	16,8	60	1155	2,3	1422	2,2	0,44
3	16,8	50	1850	1,6	0	0	0,31
4	15,8	69	1200	2,8	0	0	0,24
5	19,76	57	666	2,3	0	0	0,31
6	10,3	70	1700	2,6	0	0	0,24
7	14	48	933	3,05	0	0	0,21
8	17,1	62	619	2,5	238	1,3	0,28
9	18	71	1288	2,3	366	0,96	0,32
10	16,4	65	1418	2,1	620	1,4	0,3
11	16,5	82	800	3	0	0	0,17
12	16,2	68	1100	2,6	0	0	0,31
13	17	70	434	3	0	0	0,23
14	15,6	75	220	3,1	0	0	0,3
15	17,9	80	288	3,2	0	0	0,3

**Количественные значения фитоценологических параметров  
на участках обитания лазоревки**

№ участка	Высота древостоя, м	Сомкнутость крон, %	Густота подроста, экз./га	Высота подроста, м	Густота подлеска, экз./га	Высота травостоя, м	Покровные травостоя, %
1	13	62	1925	1,7	275	1,7	0,7
2	15	42	375	1,8	906	1,2	0,6
3	16	46	867	0,6	586	0,9	0,6
4	17,1	45	350	2,2	1632	1,6	0,9
5	12,3	39	213	1,3	566	1,4	0,9
6	15,6	42	453	1,2	366	2,9	0,6
7	19,8	53	800	1,4	280	1,2	0,3
8	17,4	54	1943	1	580	3,4	0,9
9	17,3	35	400	1,5	1325	3,4	1,1
10	17,1	47	333	1,3	450	3,4	0,6
11	18,4	67	440	1,4	420	2,5	0,7
12	18,9	32	600	3,7	120	1,9	1,2
13	17,5	36	225	1,5	60	0,9	1,1
14	18	16	575	1,6	150	1,2	0,9
15	16	50	300	1,2	0	0	0,5

### ВАРИАНТ 2

На территории дубрав Приволжской возвышенности проведено описание местообитаний черного и певчего дроздов (табл. 34–35).

Проанализировав результаты дискриминантного анализа ответьте на вопросы:

Знание количественных значений каких фитоценологических параметров дает возможность предсказания поселения на данном участке леса черного или певчего дрозда?

Могут ли данные виды поселяться на одинаковых по структурным особенностям участках леса, или каждый вид имеет четкие предпочтения?

Таблица 34

**Количественные значения фитоценологических параметров  
на участках обитания черного дрозда**

№ участка	Высота древостоя, м	Сомкнутость крон, %	Густота подроста, экз/га	Высота подроста, м	Густота подлеска, экз/га	Высота травостоя, м	Покрывие травостоя, %
1	14,3	93,8	2500	2,7	5100	2,1	78
2	16,3	55,8	3700	2,6	2900	1,21	48,6
3	11	59,7	2300	2,5	4100	6	75
4	14,7	56,2	4300	1,9	2100	1,7	76
5	12,1	49,6	4800	0,9	4300	1,5	80
6	10	46,5	3200	3,7	5600	2,3	56
7	13,3	46,3	5100	3,8	2300	1,6	50
8	14,1	52,3	2900	2,6	3200	1,7	51
9	17,5	55,1	4100	1,9	6500	2,6	60
10	16,5	47,1	2100	1,8	5900	1,1	69
11	16,5	41,2	4300	1,2	6500	0,9	65
12	15,5	64,7	5600	2,5	4500	1,5	54
13	17	42,1	2300	3	6200	1,9	59
14	16,5	32,8	3200	3,6	5500	2,1	68

Таблица 35

**Количественные значения фитоценологических параметров  
на участках обитания певчего дрозда**

№ участка	Высота древостоя, м	Сомкнутость крон, %	Густота подроста, экз/га	Высота подроста, м	Густота подлеска, экз/га	Высота травостоя, м	Покрывие травостоя, %
1	9,7	37	6500	3	5600	0,9	71
2	13,9	59	5900	2,7	3600	1,2	70
3	14,9	72	6500	3,1	3100	1	63
4	16,9	37	4500	2,4	4200	1,2	51,7

Окончание табл. 35

5	13,1	12	6200	2,7	3200	1,1	46,3
6	15,5	25	5500	2,5	1200	0,5	51,2
7	16,5	12,5	1300	2,5	10000	0,5	21,5
8	15,5	60,1	1400	2,5	9600	0,5	28
9	12,3	53,9	1300	2,1	8000	0,6	37
10	13,1	41,2	6200	3,5	4500	0,4	62
11	15,5	59	8200	2,7	4200	0,6	70
12	15,5	60,1	2500	2,5	5300	0,5	28
13	14,2	40,1	7300	2,5	3100	0,4	54

Саратовский государственный университет имени Н. Г. Чернышевского

## СПИСОК РЕКОМЕНДУЕМОЙ ЛИТЕРАТУРЫ

- Афифи А., Эйзен С.* Статистический анализ: Подход с использованием ЭВМ. М.: Мир, 1982. С. 141–379.
- Бейли Н.* Статистические методы в биологии. М.: Мир, 1985. 271 с.
- Боровиков В.П., Боровиков И.П.* STATISTICA. М.: Информ.-изд. дом «ФИЛИН», 1997. 608 с.
- Голикова Т.И., Никитина Е.П., Терехин А.Т.* Математическая статистика.: Учеб. пособие для студентов-биологов. М.: Изд-во МГУ, 1981. 185 с.
- Джонгман Р.Г., Тер Браак С.Д.* Анализ данных в экологии сообществ и ландшафтов. М.: Финансы и статистика, 1999. 306 с.
- Дубровин А.М., Мхитрян В.С., Трошин Л.И.* Многомерные статистические методы. М.: Финансы и статистика, 2000. 352 с.
- Животовский Л.А.* Популяционная биометрия. М.: Наука, 1991. Гл. 1. С. 7–69, Гл. 4. С. 203–253.
- Иберла К.* Факторный анализ. М.: Статистика, 1980. Гл. 1–3. С. 12–162.
- Компьютерная биометрика / Под ред. В.Н.Носова. М.: Изд-во МГУ, 1990. 210 с.
- Лакин С.Ф.* Биометрия: Учеб. пособие для биол. вузов. М.: Высш. школа, 1990. 352 с.
- Любищев А.А.* Дисперсионный анализ в биологии. М.: Изд-во МГУ, 1996. 200 с.
- Малета Ю.С., Тарасов В.В.* Непараметрические методы статистического анализа в биологии и медицине. М.: Изд-во МГУ, 1982. 180 с.
- Маслов А.А.* Количественный анализ горизонтальной структуры лесных сообществ. М.: Наука, 1990. 159 с.
- Маслов А.А.* О совместном применении метода блоков и метода главных компонент для анализа мозаичности лесных сообществ // Бюл. МОИП. 1985. Т. 90. Вып. 4. С. 107–117.
- Мегарран Э.* Экологическое разнообразие и его измерение. М.: Мир, 1992. 184 с.
- Павлинов И.Я.* Кластический анализ. М.: Изд-во МГУ, 1990. 118 с.
- Пузаченко Ю.Г.* Математические методы в экологических и географических исследованиях. М.: Изд. центр «Академия», 2004. 416 с.
- Справочник по прикладной статистике: В 2 т. /Под ред. С.А. Айвазяна, Ю.Н. Тюрина. М.: Финансы и статистика, 1990. Т.1, 2.
- Шаталкин А.И.* Биологическая систематика. М.: Изд-во МГУ, 1988. 182 с.
- Эфрон Б.С.* Нетрадиционные методы многомерного статистического анализа. М.: Финансы и статистика, 1988. 263 с.

Саратовский государственный университет имени Н. Г. Чернышевского

Учебное издание

*Давиденко Татьяна Николаевна,  
Давиденко Ольга Николаевна,  
Пискунов Владимир Валериевич,  
Болдырев Владимир Александрович*

**МНОГОМЕРНЫЕ МЕТОДЫ  
СТАТИСТИЧЕСКОГО АНАЛИЗА ДАННЫХ  
В ЭКОЛОГИИ**

*Учебное пособие для студентов  
биологического факультета, обучающихся  
по специальностям 013100 «Экология», 011600 «Биология»*

Редактор В.А. Трушина  
Технический редактор Л.В. Агальцова  
Корректор Е.Б. Крылова  
Оригинал-макет подготовила Т.Н. Давиденко

---

Подписано в печать 8.11.2006.  
Формат 60x84 1/16. Бумага офсетная. Гарнитура Таймс.  
Печать офсетная. Усл. печ. л. 3,25 (3,5) Уч.-изд. л. 3,1  
Тираж 200. Заказ 182.

---

Издательство Саратовского университета.  
410012, Саратов, Астраханская, 83.  
Типография Издательства Саратовского университета.  
410012, Саратов, Астраханская, 83.



Саратовский государственный университет имени Н. Г. Чернышевского