



АЛГОРИТМЫ ВЫЧИСЛИТЕЛЬНОЙ СТАТИСТИКИ В СИСТЕМЕ

А. Г. Буховец
П. В. Москалев



**А. Г. БУХОВЕЦ,
П. В. МОСКАЛЕВ**

АЛГОРИТМЫ ВЫЧИСЛИТЕЛЬНОЙ СТАТИСТИКИ В СИСТЕМЕ R

Издание второе, переработанное и дополненное

ДОПУЩЕНО

*УМО по образованию в области прикладной информатики
в качестве учебного пособия для студентов вузов,
обучающихся по направлению «Прикладная информатика»*



ЛАНЬ®
САНКТ-ПЕТЕРБУРГ · МОСКВА · КРАСНОДАР
2023

ББК 22.18я73

Б 68

Буховец А. Г., Москалев П. В.

Б 68 Алгоритмы вычислительной статистики в системе R: Учебное пособие. — 2-е изд., перераб. и доп. — СПб.: Издательство «Лань», 2023. — 160 с.: ил. — (Учебники для вузов. Специальная литература).

ISBN 978-5-8114-1802-2

В учебном пособии в краткой форме излагается теоретический материал и приводятся примеры решения практических задач по разделам: линейная алгебра, теория вероятностей, методы оценивания и проверки гипотез, метод главных компонент, регрессионный и кластерный анализ с применением свободной системы статистической обработки данных и программирования R. В приложениях к настоящему пособию содержатся сведения по установке и использованию системы R, а также листинги программ, которые могут быть использованы в учебном процессе.

Учебное пособие предназначено для студентов, обучающихся по направлению «Прикладная информатика», программа которых предусматривает изучение современных средств и методов вычислительной статистики.

ББК 22.18я73

Рецензенты:

Ю. Ю. ТАРАСЕВИЧ — доктор физико-математических наук, профессор, зав. кафедрой прикладной математики и информатики Астраханского государственного университета;

М. Г. МАТВЕЕВ — доктор технических наук, профессор, зав. кафедрой информационных технологий управления Воронежского государственного университета.

Обложка
Е. А. ВЛАСОВА

© Издательство «Лань», 2023
© А. Г. Буховец, П. В. Москалев, 2023
© Издательство «Лань»,
художественное оформление, 2023

Оглавление

Введение	5
Глава 1. Элементы линейной алгебры	7
1.1. Векторное пространство	7
1.2. Базис векторного пространства	9
1.3. Скалярное произведение векторов	11
1.4. Матрицы	13
1.5. Транспонирование, произведение, ранг	15
1.6. Определители и собственные значения	18
Глава 2. Сведения из теории вероятностей	25
2.1. Случайное событие и вероятность	25
2.2. Условная вероятность событий	27
2.3. Одномерные случайные величины	27
2.4. Многомерные случайные величины	29
2.5. Числовые характеристики случайных величин	31
2.6. Некоторые распределения	34
Глава 3. Методы оценивания и проверки гипотез	51
3.1. Генеральная и выборочная совокупности	51
3.2. Точечные оценки параметров распределения	52
3.3. Интервальные оценки параметров распределения	58
3.4. Проверка статистических гипотез	63
Глава 4. Метод главных компонент	84
4.1. Постановка задачи	84
4.2. Вычисление главных компонент	85
4.3. Основные свойства главных компонент	87
Глава 5. Начала регрессионного анализа	93
5.1. Парная линейная регрессия	93
5.2. Множественная линейная регрессия	101

Глава 6. Основы кластерного анализа	114
6.1. Содержательная постановка задачи	114
6.2. Формальная постановка задачи	115
6.3. Алгоритмы кластерного анализа	118
Литература	126
Приложение А. Введение в систему R	127
А.1. Принципы взаимодействия с R	127
А.2. Основные возможности языка R	131
Приложение Б. Листинги программ	139
Б.1. Визуализация законов распределения	139
Б.2. Методы оценивания и проверки гипотез	140
Б.3. Метод главных компонент	145
Б.4. Начала регрессионного анализа	145
Б.5. Основы кластерного анализа	147