

ГБОУ ВПО «ОРЕНБУРГСКАЯ ГОСУДАРСТВЕННАЯ
МЕДИЦИНСКАЯ АКАДЕМИЯ» МИНЗДРАВА РОССИИ

РУКОВОДСТВО

ПО ОБЕСПЕЧЕНИЮ РЕШЕНИЯ
МЕДИКО-БИОЛОГИЧЕСКИХ
ЗАДАЧ С ПРИМЕНЕНИЕМ
ПРОГРАММЫ STATISTICA 10.0



УДК: 311.17:61:004.42
ББК: 51.1 (2Рос), 0: 5 с 51
Р84

Рецензенты:

**Доктор технических наук, профессор,
заведующий кафедрой информационных систем и технологий
ФГБОУ ВПО «Оренбургский государственный университет»
В.И. ЧЕПАСОВ**

**Доктор медицинских наук, профессор,
заслуженный деятель науки РФ, заведующий кафедрой
общественного здоровья и здравоохранения
ГБОУ ВПО «Северо-Западный медицинский университет» Минздрава России
В.С. ЛУЧКЕВИЧ**

Боев В.М., Борщук Е.Л., Екимов А.К., Бегун Д.Н.

Р 84 Руководство по обеспечению решения медико-биологических задач с применением программы Statistica 10.0. Оренбург, ОАО «ИПК «Южный Урал», 2014 г. – 208 с.
ISBN 978-5-94162-092-6

В руководстве представлены характеристики большинства статистических методов, реализованных в программе STATISTICA 10.0, необходимых для решения научных задач, возникающих в процессе проведения медико-биологических исследований. Большое внимание уделено применению многомерного анализа.

Вместе с необходимой теоретической справкой, имеющейся в самой книге, на прилагаемом к ней CD-диске представлено подробное описание последовательности всех действий пользователя на примере конкретных задач. Приложены все исходные материалы для самостоятельной работы.

Книга ориентирована на аспирантов медицинских ВУЗов, научных медицинских работников, преподавателей, врачей, занимающихся научной работой.

УДК: 311.17:61:004.42
ББК: 51.1 (2Рос), 0: 5 с 51

ISBN 978-5-94162-092-6

© Боев В.М., 2014
© Борщук Е.Л., 2014
© Екимов А.К., 2014
© Бегун Д.Н., 2014
© ОАО «ИПК «Южный Урал», 2014

ОГЛАВЛЕНИЕ

Введение	5
Глава 1. Особенности анализа данных на современном этапе	6
§ 1. Основные особенности анализа данных	6
§ 2. Особенности использования программы STATISTICA	9
§ 2a. Использование «языка действий» для однозначного предписания пользователю последовательности действий при работе с программой STATISTICA	12
§ 3. Классы объектов, выборки, измерительные шкалы и адекватные математические методы	23
Глава 2. Анализ объектов, выраженных номинальными признаками	26
§ 4. Анализ одной выборки (качественные шкалы)	27
§ 5. Анализ двух выборок (качественные шкалы)	31
§ 6. Задачи установления наличия зависимости между признаками	38
Глава 3. Способы нахождения общих закономерностей оценки анализируемых объектов, оцениваемых несколькими признаками, измеренными в качественных шкалах	41
§ 7. Кластерный анализ	41
§ 8. Использование методов статистики для изучения особенностей выделенных кластеров	46
§ 9. Метод деревьев классификации	52
Глава 4. Анализ объектов, выраженных признаками, измеренными в количественных шкалах	61
§ 10. Анализ одной выборки	61
§ 11. Анализ двух выборок	66
§ 12. Сравнение средних нескольких выборок (дисперсионный анализ с односторонней классификацией)	75
§ 13. Анализ связей между несколькими количественными признаками (корреляционный анализ)	77
Глава 5. Методы анализа объектов, оцениваемых несколькими признаками, измеренными в количественных шкалах	84

§ 14. Факторный анализ	84
§ 15. Дискриминантный анализ	105
§ 16. Линейный регрессионный анализ	117
§ 17. Нелинейный регрессионный анализ	124
Глава 6. Методы анализа временных рядов	129
§ 18. Математические аспекты анализа временных рядов	129
§ 19. Анализ временных рядов с помощью программы STATISTICA	132
Глава 7. Использование нейронных сетей для обнаружения закономерностей в классе объектов, описываемых многими признаками	142
§ 20. Решение задачи классификации с помощью модуля «Нейронные сети» программы STATISTICA	143
§ 21. Обучение нейронной сети пользователем	155
§ 22. Решение задачи регрессии с помощью модуля «Нейронные сети» программы STATISTICA	157
Глава 8. Математические методы анализа, не включенные в программу STATISTICA 10	161
§ 23. Основные понятия теории вероятностей	162
§ 24. Метод неоднородной последовательной статистической процедуры (НПСП)	166
§ 25. Оценка анализируемых объектов с помощью цепей Маркова	183
Литература	202

Литература

1. *Breiman, L., Friedman, J.H., Olshen, R.A. & Stone, C.J. (1984)* Classification and Regression Trees. Wadsworth and Brooks/Cole, Monterey, California, 358 pp.
2. *Codd, E.F. (1970)*. «A Relational Model of Data for Large Shared Data Banks». Communications of the ACM 13 (6): 377–387.
3. *Fayyad, Piatetsky-Shapiro, Smyth, and Uthurusamy* Advances in Knowledge Discovery and Data Mining, (Chapter 1) AAAI/MIT Press 1996.
4. *Lim, T.-S., Loh, W.-Y., & Shih, Y.-S. (1997)*. An empirical comparison of decision trees and other classification methods. Technical Report 979, Department of Statistics, University of Wisconsin, Madison.
5. Manual of the international statistical classification of diseases, injuries, and causes of death, 9th revision, vol 1. Geneva: World Health Organization, 1977.
6. *Ripley, B. D. (1981)*. Spacial statistics. New York: Wiley.
7. *Айвазян С. А., Бухштабер В.М., Енюков Я. С., Мешалкин Л. Д.* Прикладная статистика: Классификация и снижение размерности. – М.: Финансы и статистика, 1989. – 607 с.
8. *Айвазян С. А., Енюков Я. С., Мешалкин Л. Д.* Прикладная статистика: Основы моделирования и первичная обработка данных. – М.: Финансы и статистика, 1983. – 471 с.
9. *Айвазян С.А., Енюков И. С., Мешалкин Л. Д.* Прикладная статистика: Исследование зависимостей. – М.: Финансы и статистика, 1985. – 487 с.
10. *Бартоломью Д.* Стохастические модели социальных процессов. – М.: Финансы и статистика, 1985. – 295 с.
11. *Бендат Дж., Пирсол А.* Применение корреляционного и спектрального анализа. – М.: Мир, 1983. – 312 с.
12. *Боев В. М.* Оценка демографических и социально-экономических показателей в системе социально-гигиенического мониторинга / В. М. Боев, Б. Л. Колесников, А. К. Екимов // Гигиена и санитария. 2008. – №3. – С. 92-94.
13. *Боев В.М., Воляник М.Н.* Антропогенное загрязнение окружающей среды и состояние здоровья населения Восточного Оренбуржья. – Екатеринбург: УрО РАН, 1995. – 126 с.
14. *Боев В.М., Екимов А.К., Руденко В.Н.* Функциональная организация корково-ретикулярной системы мозга при физических нагрузках. // Журнал высшей нервной деятельности. 1987, т. 37, №3, с. 515 – 521.
15. *Бокс Дж., Дженкинс Г.* Анализ временных рядов: Прогноз и управление. – М.: Мир, 1974. Вып.1. – 406 с.; Вып.2. – 224 с.
16. Большая Советская Энциклопедия. Изд. 3-е. М.: Советская энциклопедия, 1972, т. 10, – 592 с.
17. Большая Советская Энциклопедия. Изд. 3-е. М.: Советская энциклопедия, 1974, т. 17, – 616 с.
18. Большая Советская Энциклопедия. Изд. 3-е. М.: Советская энциклопедия, 1974, т. 30, – 632 с.
19. *Боровиков В.И.* STATISTICA. Искусство анализа данных на компьютере: Для профессионалов. 2-е изд. – СПб.: Питер, 2003. – 688 с.
20. *Вальтер Я.* Стохастические модели в экономике. – М.: Статистика, 1976. – 231 с.
21. *Вальд А.* Последовательный анализ. – М., Физматгиз, 1960. – 328 с.
22. *Канюков В.Н., Екимов А.К., Стрекаловская А.Д., Мурашов А.Д.* Вероятность и ин-

формация. Медицинский аспект // Современные информационные технологии в науке, образовании и практике. – 2005.

23. Лебедева И. В., Колесников Б. Л., Екимов А. К., Редина О. С. Влияние диспансерной работы на экономику региона. // Экономика здравоохранения 2005. – №1, с. 9 – 15.

24. Вознесенский В.А. Статистические методы планирования эксперимента в технико-экономических исследованиях. М.: Статистика, 1974. – 192 с.

25. Воляник О.В., Кирбятьева Е.Н., Сергиенко Л.А., Перевертова И.В., Багаутдинов А.Т., Екимов А.К. Изменение уровня интерферона- γ и фактора некроза опухоли- α у детей, больных бронхиальной астмой после курса баротерапии // Материалы Российской научно-практической конференции «Актуальные вопросы аллергологии и иммунологии в педиатрии», Оренбург, 20 – 21 мая 2003 г. – 108 с.; с. 28 – 30.

26. Горбань А.Н., Россиев Д.А. Нейронные сети на персональном компьютере. – Новосибирск: Наука, 1996. – 276 с.

27. Гублер Е.В. Вычислительные методы анализа и распознавания патологических процессов. – Л.: Медицина, 1978. – 296 с.

28. Дедова О.В., Воляник М.Н., Колосова Е.Г., Алеманова Г.Д., Багаутдинов А.Т., Екимов А.К. Функция цитовидной железы у детей с бронхиальной астмой под воздействием адаптации к гипобарической гипоксической стимуляции // Биохимия: от исследования молекулярных механизмов – до внедрения в клиническую практику и производство (Материалы межрегиональной конференции биохимиков Урала, Западной Сибири и Поволжья), Оренбург: ФГУП «ИПК «Южный Урал», 2003. – 527 с; с. 251 – 254.

29. Дмитриева Т.В., Алеманова Г.Д., Воляник О.В., Багаутдинов А.Т., Екимов А.К. Влияние адаптации к периодической гипобарической гипоксической стимуляции на цитокиновый профиль у мальчиков препубертатного возраста, больных бронхиальной астмой // Материалы Российской научно-практической конференции «Актуальные вопросы аллергологии и иммунологии в педиатрии», Оренбург, 20 – 21 мая 2003 г. – 108 с.; с.34 – 36

30. Дмитриева Т.В., Сергиенко Л.А., Перевертова И.В., Багаутдинов А.Т., Екимов А.К. Влияние баротерапии на динамику некоторых цитокинов у мальчиков пубертатного возраста, страдающих бронхиальной астмой // Материалы Российской научно-практической конференции «Актуальные вопросы аллергологии и иммунологии в педиатрии», Оренбург, 20 – 21 мая 2003 г. – 108 с.; с.33 – 34.

31. Дубров А.М. Мхитарян В.С. Многомерные статистические методы для экономистов и менеджеров – М.: Финансы и статистика, 2000. – 352 с.

32. Дьяконов В. МАТНСАД 8/2000: специальный справочник. – СПб: Питер, 2001. – 592 с.

33. Дьяконов В. П. Maple 9 в математике, физике и образовании. М.: СОЛОН-Пресс, 2004. – 688 с.

34. Дюран Б., Одел П. Кластерный анализ. – М.: Статистика, 1977. – 128 с.

35. Екимов А. К., Дрошнев В. В. Математическое моделирование: основа экономического прогнозирования функционирования системы здравоохранения города // Экономика здравоохран., 1999. – №7, 8. – с. 43 – 45.

36. Екимов А.К. Анализ системы здравоохранения с использованием методов марковских процессов. // Реформирование системы здравоохранения города Оренбурга: Состояние, Перспективы развития. Механизмы реализации, Оренбург, 1998, с. 37 – 45.

37. Екимов А.К. Методика оценки прямых затрат и косвенных потерь, связанных с заболеваемостью работников региона // Международный форум информатизации процессов охраны здоровья населения. М.: 2001, с. 69 – 71.

38. Екимов А.К. Экономическая оценка работника в состоянии «здоров» и «болен» // Экономика здравоохранения, 1 (60), 2002 с. 41 – 44.

39. Екимов А.К., Дрошнев В.В. Возможности применения метода марковских про-

цессов для анализа и моделирования работы системы здравоохранения. Страхование в условиях формирования рыночных отношений. Тезисы II Всероссийской научно-практической конференции. Часть 1. 1999, с. 25 – 27.

40. *Екимов А.К., Дрошнев В.В., Саньков А.Н.* Клинико-экономические и социально-демографические аспекты медикаментозной терапии беременных женщин. // Новые технологии в акушерстве и гинекологии, М., 1999, с. 156 – 157.

41. *Екимов А.К., Естеев В.М., Комаров Н.Н.* Современные подходы к управлению в здравоохранении. – Оренбург: ОАО «ИПК Южный Урал», 2006. – 400 с.

42. *Екимов А.К., Перепелкина Н.Ю.* Экономический анализ прямых затрат ЛПУ и косвенных потерь общества, связанных с болезнями мочевой системы у детей // Бюллетень НИИ социальной гигиены, экономики и управления здравоохранением им. Н.А. Семашко РАМН (Материалы научно-практической конференции «Здоровье населения, политика и технология развития здравоохранения»). Вып. 5, М., 2003. с. 120 – 122.

43. *Екимов, А. К.* Применение математических методов для прогнозирования результатов лечения больных с приобретенной миопией / А. К. Екимов, А. Е. Апрелев // Вестник Оренбургского Государственного Университета (специальный выпуск). – 2010. – №12. – С. 64-67.

44. *Естеев В.М., Комаров Н.Н., Екимов А.К., Брижан П.В., Екимов А.А., Сидоров С.В.* Использование идей объектно-ориентированного программирования в управлении здравоохранением. // Бюллетень НИИ социальной гигиены, экономики и управления здравоохранением им. Н.А. Семашко РАМН, вып. 10., М.: 2003, с. 81 – 86.

45. *Жамбю М.* Иерархический кластер-анализ и соответствия. – М.: Финансы и статистика, 1988. 342 с.

46. *Закс Л.* Статистическое оценивание. – М.: Статистика, 1976. – 598 с.

47. *Иберла К.* Факторный анализ. – М.: Статистика, 1980. – 398 с.

48. *Канюков В.Н., Афанасьев В.Н., Екимов А.К.* Использование методов моделирования для выработки оптимального экономического поведения офтальмологического центра. Современные информационные технологии в науке, образовании и практике. Материалы региональной научно-практической конференции. – Оренбург, ИПК ОГУ, 2002. – с. 336 - 344.

49. *Канюков В.Н., Афанасьев В.Н., Екимов А.К.* Способ интегральной оценки врачей-офтальмологов. Новые технологии микрохирургии глаза: (XIII Российская ежегодная научно-практическая конференция). Оренбург, 2002 г. – 200 с; с. 150 – 156.

50. *Канюков В.Н., Афанасьев В.Н., Екимов А.К., Чеснокова Е.Ф.* Использование математических методов для снижения уровня послеоперационных осложнений при операциях катаракты. Современные информационные технологии в науке, образовании и практике. Материалы региональной научно-практической конференции. – Оренбург, ИПК ОГУ, 2002. – 436 с., с. 327 – 335.

51. *Канюков В.Н., Борщук Е.Л., Трубина О.М., Екимов А.К.* Основы региональной системы здравоохранения: Руководство. – Оренбург, 2012. – 317 с.

52. *Канюков В.Н., Екимов А.К.* Моделирование в управлении здравоохранением. – Оренбург: ОГУ, 2010. – 564 с.

53. *Канюков В.Н., Екимов А.К., Мецержакова Г.Ф., Щербанов В.В.* Использование методов моделирования для повышения эффективности диспансерной работы // Современные информационные технологии в науке, образовании и практике. – Оренбург, ИПК ОГУ ОГУ, 2005. – 450 с., с. 346 – 349

54. *Канюков В.Н., Екимов А.К., Щербанов В.В.* Математический анализ в офтальмологии. Оренбург: ОАО «ИПК «Южный Урал», 2005. – 240 с.

55. *Каратыгин С.А., Тихонов А.Ф., Тихонова Л.Н.* VisualFoxPro 7. – М.: Бином-Пресс, 2003 г. – 656 с.

56. *Кемени Дж., Снелл Дж.* Конечные цепи Маркова. – М.: Наука, 1970.
57. *Ким Дж.-О., Мюллер Ч.У., Клекка У.Р. и др.* Факторный, дискриминантный и кластерный анализ. М: Финансы и статистика 1989. – 215 с.
58. Классификация и кластер. – М.: Мир, 1980. – 390 с.
59. *Клир Дж.* Системология. Автоматизация решения системных задач. – М.: Радио и связь, 1990. – 544 с.
60. *Колесников Б.Л., Екимов А.К., Редина О.С.* Оценка диспансерной работы на основе процессов Маркова – Оренбург, Вестник ОГУ №6, 2006, том 2, с. 117 – 127.
61. *Комаров Н. Н.* К вопросу о создании системы автоматизированного поиска управленческих решений в здравоохранении / Н. Н. Комаров, А. К. Екимов // Экономист лечебного учреждения. – 2009, №2. – с. 12 – 17.
62. *Коц Я.И., Екимов А.К., Сермягина Н.Ф.* Математический подход к проблеме ранней диагностики сердечной недостаточности. Ж. Кардиология, 1989, т. 29, № 12, с. 57 – 60.
63. *Лебедева И.В., Колесников Б.Л., Екимов А.К., Редина О.С.* Влияние диспансерной работы на экономику региона// Ж. «Экономика здравоохранения», 1, 2005 с. 9 – 15.
64. *Лебедева И.В., Семивеличенко Т.Н., Екимов А.К., Богомолова Е.Н.* Использование методов математического моделирования для прогнозирования экономических аспектов реформирования системы здравоохранения города. Сб. «Реформирование системы здравоохранения города Оренбурга: Состояние, Перспективы развития. Механизмы реализации», Оренбург, 1998, с. 30 - 36.
65. *Литтл Р. Дж. А., Рубин Д.Б.* Статистический анализ данных с пропусками/ пер. с англ. – М.: Финансы и статистика, 1990. – 336 с.
66. *Малиновский Л. Г.* Классификация объектов средствами дискриминантного анализа – Л.: Наука, 1979. – 260 с.
67. *Мандель И.Д.* Кластерный анализ. – М.: Финансы и статистика, 1988. – 176 с.
68. Международная статистическая классификация болезней и проблем, связанных со здоровьем: МКБ-10. Т. 1 (ч. 1, 2); Т. 2.; Т. 3. 10-й пересмотр. 1995.
69. *Миркин Б.Г.* Анализ качественных признаков и структур. – М.: Статистика, 1980. – 320 с.
70. *Орел В.И., Чолоян С.Б., Екимов А.К.* Использование методов сетевого планирования для экономического анализа деятельности отделений и служб ЛПУ.// Бюллетень НИИ социальной гигиены, экономики и управления здравоохранением им. Н.А. Семашко РАМН (Материалы научно-практической конференции «Здоровье населения, политика и технология развития здравоохранения»). Вып. 8, М., 2003. с. 102 – 105.
71. Основы геоинформатики/ Под ред. Тикунова В.С.. В 2 кн. – М.: Издательский центр «Академия», 2004, кн. 1. – 352 с.; кн. 2. – 480 с.
72. *Пережудов Ф.И., Тарасенко Ф.П.* Введение в системный анализ. – М.: Высшая школа, 1989. – 367 с.
73. *Поллард Дж.* Справочник по вычислительным методам статистики. – М.: Финансы и статистика, 1982. – 344 с.
74. *Пфанцгль И.Г.* Теория измерений. – М.: Мир, 1976. – 248 с.
75. *Рассел С., Норвиг П.* Искусственный интеллект: современный подход. – Вильямс, 2006. – 1408 с.
76. *Реброва О.Ю.* Статистический анализ медицинских данных. Применение пакета прикладных программ STATISTICA. – М.: МедиаСфера, 2002. 312 с.
77. *Себер Дж.* Линейный регрессионный анализ. – М.: Мир, 1980. – 456 с.
78. *Феллер В.* Введение в теорию вероятностей и ее приложения. – М.: Мир, 1984. т.1. – 528 с.