

А.Н. Герасимов

# МЕДИЦИНСКАЯ СТАТИСТИКА

**УЧЕБНОЕ ПОСОБИЕ**

*Рекомендовано Учебно-методическим объединением по медицинскому  
и фармацевтическому образованию вузов России в качестве  
учебного пособия для студентов медицинских вузов.*



Медицинское информационное агентство  
Москва  
2007

УДК 61  
ББК 5  
Г37

**Герасимов А.Н.**  
Г37 Медицинская статистика: Учебное пособие. — М.: ООО «Медицинское информационное агентство», 2007. — 480 с. : ил.

ISBN 5-89481-456-1

В книге, подготовленной заведующим кафедрой медицинской информатики и статистики ММА им. И.М. Сеченова, речь идет о статистической обработке медицинских данных с помощью пакетов статистических программ, то есть о теории вероятностей и математической статистике для медиков. Излагаются не только математические результаты и технические приемы обработки данных, но и некоторые вопросы планирования статистических исследований и интерпретации полученных результатов. Структура и последовательность изложения материала резко отличается от обычной для учебников по теории вероятностей и математической статистике. В них, в соответствии с внутренней логикой, вначале излагается теория вероятностей, затем на основе изложенного аппарата — математическая статистика, а в конце обсуждаются вопросы практического применения, в частности использования программного обеспечения. А в этой книге последовательность изложения материала соответствует последовательности действий при статистической обработке данных, поэтому книгу нужно осваивать параллельно с обработкой фактических данных. Книга состоит из двух параллельно излагаемых частей. При описании того, как нужно вводить данные и использовать конкретные методы обработки данных, дается точное описание, как пользоваться программой и какие команды нужно выполнять. При описании же того, как нужно понимать полученные результаты, даются необходимые элементы теории.

УДК 61  
ББК 5

ISBN 5-89481-456-1

© Герасимов А.Н., 2007  
© Оформление. ООО «Медицинское информационное агентство», 2007

# Содержание

|                                                                                    |    |
|------------------------------------------------------------------------------------|----|
| <i>Предисловие</i> .....                                                           | 4  |
| О названии этой книги.....                                                         | 4  |
| Зачем эта книга.....                                                               | 5  |
| Почему эта книга.....                                                              | 6  |
| Для кого эта книга.....                                                            | 7  |
| Как пользоваться этой книгой.....                                                  | 8  |
| <i>Введение</i> . Типовая схема статистического анализа<br>клинических данных..... | 10 |

## Часть 1

### Анализ вероятностей

|                                                                                                      |    |
|------------------------------------------------------------------------------------------------------|----|
| <i>Глава 1</i> . Назначение и возможности статистического<br>анализа.....                            | 16 |
| 1.1. Статистические и причинно-следственные<br>связи.....                                            | 16 |
| 1.1.1. Сцепленные факторы<br>и репрезентативные выборки.....                                         | 17 |
| 1.1.2. Смена групп наблюдения<br>и возрастные изменения.....                                         | 22 |
| 1.1.3. Причина или следствие.....                                                                    | 26 |
| 1.2. Доказательная медицина<br>и дизайн исследования.....                                            | 28 |
| 1.3. Статистическая обработка как<br>вид математической обработки.....                               | 32 |
| 1.4. Случайные и контролируемые условия.....                                                         | 37 |
| <i>Глава 2</i> . Планирование эксперимента и ввод данных.....                                        | 40 |
| 2.1. Этапы статистического исследования.....                                                         | 40 |
| 2.2. Кодирование и ввод данных.....                                                                  | 42 |
| 2.2.1. Общие замечания.....                                                                          | 42 |
| 2.2.2. Подготовка к работе с пакетом<br>статистических программ и работа<br>с файлами с данными..... | 47 |
| 2.2.3. Ввод данных в статпакете.....                                                                 | 53 |
| 2.2.4. Вычисляемые переменные.....                                                                   | 60 |
| 2.2.5. Поиск грубых ошибок.....                                                                      | 65 |

|                 |                                                                                                                                                                                                    |     |
|-----------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| <i>Глава 3.</i> | Анализ частот и оптимизация рисков .....                                                                                                                                                           | 66  |
| 3.1.            | Расчет частот .....                                                                                                                                                                                | 66  |
| 3.2.            | Частоты, риски и относительные риски .....                                                                                                                                                         | 75  |
| 3.3.            | Управление рисками и поиск наилучшего решения. Игры с противоположными интересами (задача о двух пивных) .....                                                                                     | 80  |
| 3.4.            | Игры с противоположными интересами (задачи о двух аспирантах и о трех разбойниках).....                                                                                                            | 84  |
| <i>Глава 4.</i> | Элементы теории вероятностей .....                                                                                                                                                                 | 90  |
| 4.1.            | Исторический экскурс. Случай равновероятных шансов и логические ловушки .....                                                                                                                      | 90  |
| 4.2.            | Современный подход. Аксиоматика Колмогорова .....                                                                                                                                                  | 100 |
| 4.3.            | Вычисление вероятностей. Условные, априорные и апостериорные вероятности .....                                                                                                                     | 106 |
| 4.4.            | Независимые события .....                                                                                                                                                                          | 111 |
| 4.5.            | Закон больших чисел. Расчет необходимого объема наблюдений .....                                                                                                                                   | 118 |
| 4.6.            | Функция распределения числовой случайной величины. Непрерывные и дискретные случайные величины. Распределения Бернулли, биномиальное, Пуассона, нормальное, «хи-квадрат», Стьюдента и Фишера ..... | 121 |
| <i>Глава 5.</i> | Отбор и поиск данных. Элементы математической логики .....                                                                                                                                         | 133 |
| 5.1.            | Отбор случаев при работе со статистическими программами .....                                                                                                                                      | 134 |
| 5.2.            | Вычисление истинности высказываний .....                                                                                                                                                           | 136 |
| 5.3.            | Теории и подтверждающие примеры .....                                                                                                                                                              | 142 |
| <i>Глава 6.</i> | Элементы математической статистики .....                                                                                                                                                           | 148 |
| 6.1.            | Статистические гипотезы и доверительная вероятность.....                                                                                                                                           | 148 |
| 6.2.            | Ошибки первого и второго рода. Выбор доверительной вероятности .....                                                                                                                               | 151 |
| 6.3.            | Сравнение частоты и вероятности. Доверительные вероятности.....                                                                                                                                    | 157 |

|        |                                                                                                                                                                        |     |
|--------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 6.3.1. | Расчет доверительных границ к частоте встречаемости .....                                                                                                              | 157 |
| 6.3.2. | Расчет доверительных границ к медиане и другим процентиллям .....                                                                                                      | 162 |
| 6.4.   | Сравнение набора частот с набором вероятностей и наборов частот с набором частот. Критерий $\chi^2$ («хи-квадрат») .....                                               | 168 |
| 6.4.1. | Вариант ручного счета — сравнение собственных результатов с литературными данными, сопоставление данных разных источников друг с другом, проверка на однородность..... | 168 |
| 6.4.2. | Сравнение наборов частот при помощи критерия $\chi^2$ . Условия применимости критерия $\chi^2$ .....                                                                   | 179 |
| 6.4.3. | Расчет частоты совместного распределения и определение достоверности различий при помощи критерия $\chi^2$ .....                                                       | 184 |
| 6.4.4. | Определение достоверности различий распределений в подгруппах при помощи критерия $\chi^2$ и критерия Колмогорова—Смирнова .....                                       | 187 |

## Часть 2

### Анализ параметров

|                 |                                                                                                   |     |
|-----------------|---------------------------------------------------------------------------------------------------|-----|
| <i>Глава 7.</i> | Параметры случайных величин .....                                                                 | 192 |
| 7.1.            | Что такое параметр.....                                                                           | 192 |
| 7.2.            | Лирическое отступление о статистической терминологии и трудностях перевода .....                  | 199 |
| 7.3.            | Расчет параметров в статистических пакетах...203                                                  |     |
| 7.4.            | Проверка гипотезы на принадлежность наблюдаемой случайной величины классу случайных величин ..... | 207 |
| <i>Глава 8.</i> | Содержательный анализ среднего арифметического.....                                               | 211 |
| 8.1.            | Расчет среднего арифметического и математического ожидания. Линейные                              |     |

|                  |                                                                                                             |     |
|------------------|-------------------------------------------------------------------------------------------------------------|-----|
|                  | свойства математического ожидания, дисперсии и среднеквадратичного ожидания .....                           | 211 |
| 8.2.             | Среднее арифметическое и нелинейность ущерба.....                                                           | 219 |
| 8.3.             | Выскакивающие варианты и среднее арифметическое.....                                                        | 224 |
| 8.4.             | Классификация переменных «scale», «ordinal» и «nominal» .....                                               | 228 |
| <i>Глава 9.</i>  | Точечные оценки параметров.....                                                                             | 232 |
| 9.1.             | Определение понятия оценки. Качество оценки .....                                                           | 233 |
| 9.2.             | Оценка моментов. Катастрофа неробастности старших моментов.....                                             | 238 |
| 9.3.             | Оценка дисперсии.....                                                                                       | 242 |
| 9.4.             | Построение оценок при помощи метода наибольшего правдоподобия.....                                          | 246 |
| <i>Глава 10.</i> | Интервальные оценки параметров. Определение достоверности различий.....                                     | 250 |
| 10.1.            | Центральная предельная теорема.....                                                                         | 251 |
| 10.2.            | Определение достоверности различия дисперсии .....                                                          | 257 |
| 10.3.            | Определение достоверности различия средних. Критерий Стьюдента .....                                        | 261 |
| 10.4.            | Расчет достоверности различия средних арифметических с помощью пакетов статистических программ .....        | 271 |
| 10.5.            | Расчет доверительных границ к математическому ожиданию.....                                                 | 278 |
| <i>Глава 11.</i> | Корреляционный анализ .....                                                                                 | 281 |
| 11.1.            | Определение коэффициента корреляции .....                                                                   | 282 |
| 11.2.            | Расчет коэффициента корреляции и анализ взаимосвязи двух переменных.....                                    | 288 |
| 11.3.            | Сопоставление расчета коэффициента корреляции с другими методами анализа взаимосвязи двух переменных.....   | 307 |
| 11.4.            | Расчет доверительных границ коэффициента корреляции и достоверности различий коэффициентов корреляции ..... | 309 |

|                                                                                                                                     |     |
|-------------------------------------------------------------------------------------------------------------------------------------|-----|
| <i>Глава 12.</i> Проверка корректности предположения о применимости центральной предельной теоремы и непараметрические методы ..... | 313 |
| 12.1. Случайное деление выборки на две части и сравнение результатов .....                                                          | 314 |
| 12.2. Отбрасывание высказывающей варианты .....                                                                                     | 319 |
| 12.3. Непараметрические методы .....                                                                                                | 322 |
| 12.3.1. Сравнение функции распределения .....                                                                                       | 323 |
| 12.3.2. Анализ связи рангов переменных .....                                                                                        | 325 |
| 12.3.3. Нелинейное шкалирование.....                                                                                                | 329 |

### **Часть 3**

#### **Прогнозирование**

##### **и некоторые дополнительные методы**

|                                                                                                                                                                |     |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| <i>Глава 13.</i> Постановка задачи прогнозирования и прогнозирование по одной случайной величине .....                                                         | 332 |
| 13.1. Достоверность связи и прогностическая сила связи .....                                                                                                   | 332 |
| 13.2. Прогнозирование и деление переменных на классы scale и nominal .....                                                                                     | 336 |
| 13.3. Прогнозирование ожидаемого значения и задача о наилучшем приближении .....                                                                               | 340 |
| 13.4. Линейный прогноз по одному коэффициенту корреляции. Уточнение коэффициента корреляции для того случая, когда переменная задана со случайной ошибкой..... | 342 |
| 13.5. Линейный прогноз по двум факторам .....                                                                                                                  | 347 |
| 13.6. Анализ согласованности тестов. Приложение к психологии и педагогике .....                                                                                | 349 |
| 13.7. Линейный прогноз по нескольким факторам (многофакторный анализ) .....                                                                                    | 353 |
| 13.8. Применение корреляционного анализа к динамическим рядам.....                                                                                             | 357 |
| 13.8.1. Анализ Фурье .....                                                                                                                                     | 358 |
| 13.8.2. Автокорреляционная функция .....                                                                                                                       | 362 |
| 13.8.3. Марковские случайные процессы.....                                                                                                                     | 363 |
| 13.8.4. Применимость методов статистического анализа для анализа динамических рядов.....                                                                       | 365 |

|                  |                                                                                                                                                                                                   |     |
|------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| <i>Глава 14.</i> | Повышение точности прогнозирования<br>и анализ его ценности.....                                                                                                                                  | 367 |
| 14.1.            | Линейный прогноз по одному фактору —<br>линеаризация фактора .....                                                                                                                                | 367 |
| 14.2.            | Многофакторное нелинейное прогнозирование<br>как потенциально мощный практически<br>неприменимый метод. Информационная<br>модель, неявно используемая<br>в многофакторном линейном прогнозе ..... | 370 |
| 14.3.            | Проверка качества прогноза.<br>Авторешаемость и скользящий экзамен .....                                                                                                                          | 373 |
| 14.4.            | Анализ практической ценности прогноза.<br>ROC-кривые и оптимальный выбор<br>соотношений чувствительность/<br>эффективность и надежность/<br>достоверность.....                                    | 379 |
| 14.5.            | Отбор факторов .....                                                                                                                                                                              | 386 |
| <i>Глава 15.</i> | Корреляционная адаптометрия .....                                                                                                                                                                 | 390 |
| 15.1.            | Проблема сравнения тяжести разных<br>патологических состояний .....                                                                                                                               | 390 |
| 15.2.            | Практическое применение .....                                                                                                                                                                     | 393 |
| <i>Глава 16.</i> | Анализ выживаемости .....                                                                                                                                                                         | 396 |
| 16.1.            | Расчет выживаемости по Каплан—Майер .....                                                                                                                                                         | 397 |
| 16.2.            | Определение факторов риска<br>для выживаемости при помощи<br>регрессии Кокса .....                                                                                                                | 402 |
| <i>Глава 17.</i> | Кластерный анализ и задачи<br>классификации .....                                                                                                                                                 | 406 |
| 17.1.            | Иерархический кластерный анализ .....                                                                                                                                                             | 407 |
| 17.2.            | Кластерный анализ по <i>K</i> средним .....                                                                                                                                                       | 411 |
| 17.3.            | Дискриминантный анализ.....                                                                                                                                                                       | 417 |
| <i>Глава 18.</i> | Морфометрия и статистический<br>анализ изображений .....                                                                                                                                          | 419 |
| 18.1.            | Анализ распределения объектов<br>на плоскости.....                                                                                                                                                | 420 |
| 18.2.            | Анализ распределения ориентации<br>объектов на плоскости .....                                                                                                                                    | 422 |
| 18.3.            | Изучение трехмерных объектов по<br>их срезам.....                                                                                                                                                 | 423 |

|                                         |     |
|-----------------------------------------|-----|
| 18.3.1. Сечения круглых сосудов .....   | 424 |
| 18.3.2. Сечения шаров.....              | 426 |
| 18.4. Краткое локальное заключение..... | 431 |

## Приложения

|                                                                                                                          |            |
|--------------------------------------------------------------------------------------------------------------------------|------------|
| <i>Приложение 1. Краткая шпаргалка<br/>по работе в SPSS .....</i>                                                        | <i>434</i> |
| <i>Приложение 2. Статистические таблицы .....</i>                                                                        | <i>440</i> |
| <i>Приложение 3. Использование электронной<br/>таблицы Excel для проведения<br/>типовых статистических расчетов.....</i> | <i>449</i> |
| <i>Список литературы с комментариями .....</i>                                                                           | <i>467</i> |