

Современные методы анализа медицинских данных

А.И. Курочкина, Е. Н. Тимин
Институт хирургии им. А.В.
Вишневского РАМН
(директор – академик РАМН
В.Д. Федоров), Москва

The Contemporary Methods of Medical Data Analysis

A.I. Kurochkina, E.N. Timin
A.V. Vishnevsky Institute of
Surgery RAMSci. (director -
academician RAMSci.
V.D.Fedorov), Moscow

В настоящее время ни одно научное исследование, основанное на анализе результатов клинических наблюдений, невозможно представить себе без использования средств прикладной статистики. Более того, статистику часто рассматривают как метаязык описания действительности в науках, оперирующих данными, не поддающимися строгой формализации.

Прикладная статистика (Applied Statistics в англоязычной литературе) или анализ данных (L'analyse des donnees во франкоязычной) - это область математики, сложившаяся в последние десятилетия в самостоятельную научную дисциплину. Прикладная статистика определяется [1-3] как "научная дисциплина, разрабатывающая и систематизирующая понятия, приемы, математические методы и модели, предназначенные для организации сбора, стандартной записи, систематизации, обработки и анализа статистических данных с целью их удобного представления, интерпретации и получения научных и практических выводов". Прикладная статистика включает в себя математическую статистику (методы статистической обработки), методы, не опирающиеся на вероятностную природу обрабатываемых данных (теория измерений, кластер-анализ и др.), методологию их применения при решении прикладных задач, программное обеспечение, реализующее достаточно удобный и эффективный режим общения пользователя с ЭВМ в процессе решения задач.

Прикладная статистика, это совокупность рабочих инструментов, которые при осмысленном применении могут помочь исследователю упорядочить множество отдельных наблюдений, установить взаимосвязи и закономерности, обосновать или отклонить полученные при наблюдении за больным впечатления и субъективные представления, выявить и сделать наглядными обстоятельства и факты, которые ранее ускользали от внимания, способствовать обострению интуиции исследователя и выдвижению новых гипотез [5, 6]. Конечно, чем сложнее инструмент, тем опаснее его неквалифицированное использование. Известный биостатистик А.А. Любищев [8] заметил, что "всякое орудие тем опаснее при неразумном употреблении, чем оно совершеннее: автомобиль гораздо опаснее телеги, ружье гораздо опаснее топора, а в особенности дубины. Но вряд ли разумно из того, что имеет место много несчастных случаев с автомобилями и ружьями, отказаться от них...".

Широкое распространение персональных ЭВМ сопровождалось разработкой большого количества баз данных, системы управления которыми не имеют средств для анализа данных, и зачастую доказательность стала подменяться демонстрацией эффектных иллюстраций и ссылками на компьютеризированный сбор данных, а выводы остаются не доказанными. Проблема состоит и в том, что бурно возрастающие объемы информации, требующие грамотной статистической обработки и интенсивно развивающееся программное обеспечение находятся в дисбалансе с относительно медленной растущей численностью квалифицированных специалистов по прикладной статистике.

О рандомизации

Мысль о том, что для получения доказательных результатов эксперимента его следует правильно организовать имеет многолетнюю историю. Широко известна книга Фишера "Статистические методы для исследователей". М.: Госиздат, 1958. Фишер предложил и детально разработал идею рандомизации как универсальной основы планирования эксперимента. Под рандомизацией понимается случайное приписывание воздействия экспериментальным единицам. В медицине - это случайное назначение одного из вариантов лечения больным определенной категории. В последние годы, особенно после опубликования группой CONSORT рекомендаций по представлению данных рандомизированных исследований, использование рандомизации или оценки слепым методом исходов или факторов риска стало выдвигаться в качестве обязательного методологического критерия при отборе научных статей [10]. Основные способы формирования выборки при планировании научного исследования — это либо эксперимент (планируемое исследование), либо натурное наблюдение (непланируемое исследование). В основе большинства клинических исследований все-таки лежит наблюдение. Что, какие мысли, можно извлечь из совокупности наблюдений? Как анализировать множество наблюдений? Если есть сомнение в наличии каких-то эффектов, то можно ставить вопрос об организации спланированного эксперимента, используя идею рандомизации. Рандомизация - это только один из множества других способов подтвердить или отвергнуть наличие эффектов, выявленных в процессе анализа пассивного наблюдения.

Теоретически рандомизация обеспечивает устранение влияния систематических ошибок, однако при ограниченных объемах наблюдений простая рандомизация не всегда обеспечивает сопоставимость групп. Обширное рандомизированное исследование по диабету в США было прекращено, так как в процессе исследования выяснилось, что, несмотря на рандомизацию, в контрольной группе оказались более здоровые люди [11].

В рамках идей рандомизации часто требуется использование двойного слепого метода, когда врач (оценивающий результат исследования) и больной знают, какое проводится лечение. Мало сказать, что в большинстве клинических исследований это условие практически не может быть обеспечено. Одна из важных причин, по которым в клинических исследованиях идеи рандомизации до сих пор не принимаются безоговорочно - этическая и она состоит в том, что всякое исследование в клинической медицине является по существу экспериментом над людьми. Поэтому большинство клинических исследований с точки зрения точных наук имеют характер пассивного наблюдения, а не контролируемого эксперимента.

Конечно, интерпретация результатов исследований, в которых нарушались требования рандомизации и двойного слепого метода - сложная задача. Однако, развитие методов прикладной статистики дало возможность получать обоснованные выводы и из результатов пассивных наблюдений за счет усложнения процедур обработки данных.

Мысль о том, что обоснованность эксперимента не обязательно связана с идеей рандомизации развивалась в работах Кемпбелла и его учеников. Кемпбелл в рамках логики экспериментального метода трактует даже такое далекое от жестких рамок традиционного эксперимента понятие как "исследование единичного случая". В монографии [7] обсуждается понятие квазиэкспериментальных планов, позволяющих проверить влияние альтернативных гипотез наряду с эффектом изучаемого воздействия. Кемпбелл выделяет классы факторов, которые могут повлиять на результаты воздействия, если оно не является рандомизированным:

- фон (события, которые происходят в промежутке времени между измерениями до и после воздействия)
- естественное развитие заболевания, его длительность и пр.
- неэквивалентный по составу отбор групп, вызывающий появление систематической ошибки (например, разная тяжесть заболевания)
- отсев в ходе исследования (например, прекращение приема препарата) и др.

Исследование, естественно, усложняется, так как необходимо доказать, что наблюдаемый эффект не связан с влиянием перечисленных выше факторов, а лишь с изучаемым воздействием. Мы можем сослаться на блестяще проведенное исследование эффективности антиаритмических препаратов хинидина и кордарона на основании пассивного нерандомизированного наблюдения [3], проведенного в Факультетской терапевтической клинике ММА им. М.И. Сеченова.

Цитируя А.А. Любищева [7], можно резюмировать, что "принцип рандомизации при всей его простоте и пользе при правильном применении не является безусловно необходимым, но он не является и достаточным, так как непродуманное его применение от ошибок не гарантирует".

Об интерпретации корреляционной связи между признаками

Большое число ошибок связано с формальной интерпретацией корреляционной связи между признаками, когда приводят коэффициент корреляции и при значимом его отличии от нуля делают вывод о наличии связи между признаками. Все критерии значимости корреляционной связи предполагают нормальное (гауссово) распределение количественных признаков. Поэтому не имеет смысла интерпретировать корреляционную связь без графического отображения результатов измерений на плоскости, где по осям координат отложены значения исследуемых признаков (корреляционное облако). На всех графиках рис. 1 коэффициент корреляции один и тот же и равен 0.7. График А соответствует стандартной картине при нормальном распределении. Взаимосвязь на графике Б более тесная и надо исследовать причину появления нескольких выделяющихся наблюдений. На графиках В и Г имеет место неоднородность выборки, причем в случае В в каждой из двух подгрупп имеется тесная корреляционная связь, а в случае Г никакой корреляционной связи вообще нет.

О проверке однородности двух выборок

Типичным для клинических исследований является вычисление средних значений в двух группах и оценка достоверности их различий по t-критерию Стьюдента. При этом молчаливо предполагается, что выборка однородна и распределение нормально. Во многих случаях как исследователь, так и его оппоненты соглашались с этим далеко не очевидным предположением, поскольку его проверка, во-первых, трудна, а во-вторых, обе стороны интуитивно чувствуют, что в рассма-

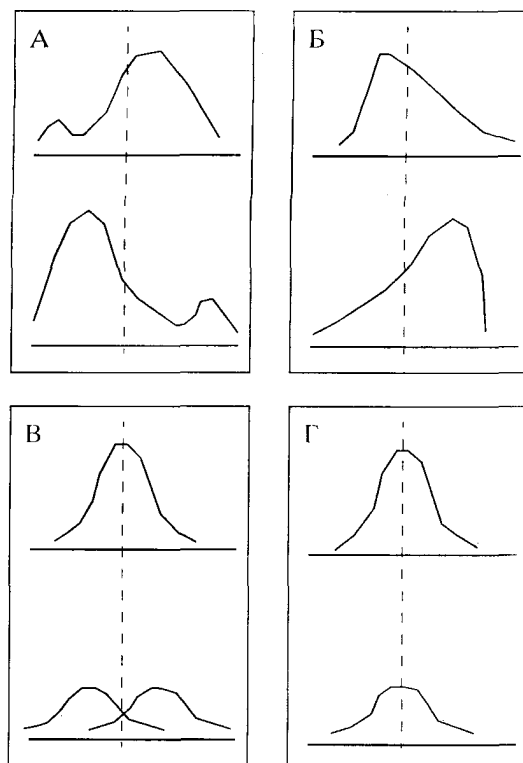


Рис. 1. Виды распределений с одинаковыми средними значениями. А – при наличии резко выделяющихся наблюдений; Б – при скошенности распределений; В – при неоднородности выборки; Г – при нормальном распределении с разной дисперсией.

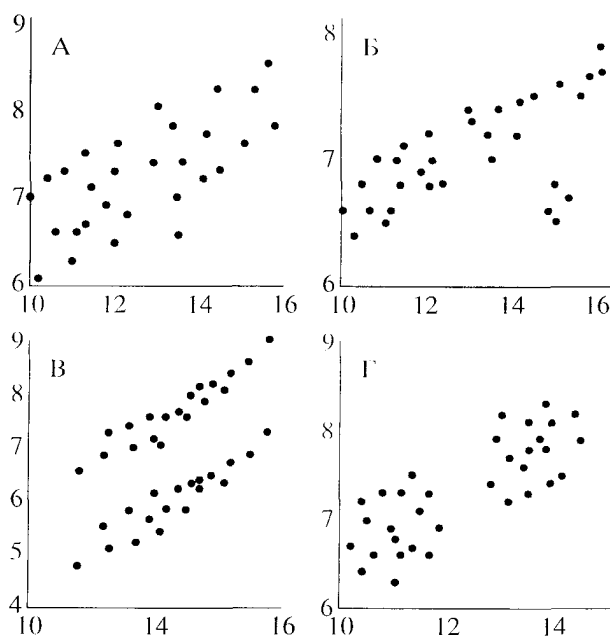


Рис. 2. Виды связи между признаками при одном и том же коэффициенте корреляции ($R = 0.7$). См. текст.

триваемой ситуации предположение скорее всего верно. Если распределения нормальны и их дисперсии равны, то задачи о сравнении распределений и о сравнении средних действительно эквивалентны. Однако, на рис. 2 приведены типичные примеры ситуаций, когда средние равны, а распределения существенно различны. При варианте А имеется сдвиг в сторону уменьшения значений параметра, однако наличие резко выделяющихся наблюдений приводит к равным средним. Особенно характерен случай В, у половины больных происходит уменьшение, у половины — увеличение значений параметра, однако среднее не изменится. Опираясь на стандартное применение t-критерия Стьюдента, был бы сделан вывод, что изменений не происходит, хотя в случаях А и Б это не так, а в случае В надо было бы поставить вопрос об уточнении показаний к операции из-за неоднородности реакций на воздействие. Вопросы проверки однородности двух выборок применительно к медикобиологическим исследованиям детально рассмотрены в работе А.И. Орлова [9].

Многомерные соотношения между признаками

Особенно эффективным применение методов прикладной статистики оказывается тогда, когда выясняется, что диагностическим признаком является не изменение измеренных показателей (они оказываются не достоверными), а изменение их соотношений. В подавляющем большинстве случаев интуиция исследователя не улавливает (за исключением случаев озарения) этих закономерностей. Много примеров, когда применение многомерных методов прикладной статистики дало возможность получить важные клинические выводы [5,6].

О программном обеспечении прикладной статистики

Пакеты программ по прикладной статистике предоставляют в распоряжение исследователя все богатство современных методов анализа данных. Один из наиболее мощных современных пакетов для анализа данных CSS: СТАТИСТИКА, создан ведущей американской компанией в области статистического программного обеспечения STATSOFT [12]. Для того, чтобы оценить возможности анализа данных, реализованные в этом пакете, приведем часть содержимого основного меню с небольшими пояснениями.

УПРАВЛЕНИЕ ДАННЫМИ - в наше распоряжение представлены средства манипуляции с данными — можно легко выделять подмножества признаков и наблюдений, упорядочивать наблюдения по любым совокупностям признаков, по разному группировать наблюдения, выполнять функциональные преобразования признаков, удалять или добавлять временно или постоянно (по желанию) любые группы наблюдений и признаков, осуществлять ввод данных или их импорт из баз данных.

ГРАФИЧЕСКИЙ АНАЛИЗ - создание множества типов графиков: гистограмм, круговых диаграмм, трехмерных и многих специальных видов графиков с привлечением соответствующих статистических тестов для проверки значимости. Графика доступна из всех разделов пакета.

Раздел **БАЗОВЫЕ СТАТИСТИКИ** включает в себя вычисление суммирующих показателей - средних, процентилей, доверительных интервалов, корреляций с графическим отображением, сравнение групп по t-критерию Стьюдента для зависимых и независимых выборок и с графическим отображением.

Раздел **НЕПАРАМЕТРИЧЕСКИЕ МЕТОДЫ** отличается от других аналогичных пакетов богатством реализованных там методов, что особенно важно для анализа качественных признаков, которых особенно много в медицине.

ТАБЛИЦЫ СОПРЯЖЕННОСТИ предназначены для анализа взаимосвязей между качественными признаками с выдачей таблиц с процентами как по столбцам, так и по строкам с проверкой гипотез по критерию χ^2 и многим другим.

ЛИНЕЙНАЯ РЕГРЕССИЯ И НЕЛИНЕЙНЫЕ МОДЕЛИ - для описания линейных и нелинейных регрессионных зависимостей с соответствующими статистическими критериями значимости.

ФАКТОРНЫЙ АНАЛИЗ-выявление факторов, которыми можно объяснить взаимосвязи между группами признаков. Одна из разновидностей факторного анализа — метод главных компонент дает возможность наглядно представить себе, как расположены наблюдения в многомерном пространстве признаков.

Два раздела предназначены для построения и анализа КРИВЫХ ВЫЖИВАЕМОСТИ.

КЛАСТЕР-АНАЛИЗ реализует методы автоматической классификации - с его помощью мы можем выделить группы похожих друг на друга наблюдений в виде графика-дендрограммы, которая отражает процесс группирования наблюдений по их сходству.

Заключение

В заключение нам хотелось бы остановиться на том обстоятельстве, что недооценка сути статистического анализа данных сказывается также в привычке привлекать статистические процедуры лишь по окончании процессов сбора данных, приводя до некоторой степени по завершении дела некое "статистическое благословение". В действительности оно часто оказывается несостоятельным, так как лежащий в основе вычислений материал не удовлетворяет требованиям методов статистической обработки. Представленные группы часто оказываются неоднородными, имеют выбросы. Можно также отметить определенную логическую незавершенность медицинских классификаций, лежащих в основе группирования данных. Поэтому необходимо обращение к специалистам по прикладной статистике еще на этапе формулировки темы до начала сбора данных для отработки плана и формы сбора данных таким образом, чтобы их затем можно было бы обрабатывать с помощью пакетов статистических программ и с тем, чтобы сбор данных и их анализ были звеньями единой системы.

Список литературы

1. Айвазян С.А., Енюков И.С., Мешалкин Л.Д. Прикладная статистика: Основы моделирования и первичная обработка данных. Справочное изд. М., Финансы и статистика, 1983.
2. Айвазян С.А., Енюков И.С., Мешалкин Л.Д. Прикладная статистика: Исследование зависимостей. Справочное изд. М., Финансы и статистика, 1985.
3. Айвазян С.А., Енюков И.С., Мешалкин Л.Д. Прикладная статистика: Классификация и снижение размерностей. Справочное изд. М., Финансы и статистика, 1989.
4. Алексеевская М.А. Методы объединения экспертных знаний и статистических выводов в задачах медицинского прогнозирования. Дис. канд. тех. наук. М., 1989.
5. Вольнский Ю.Д., Курочкина А.И. Многомерный анализ клинических данных. Кардиология. 1986. № 1. С. 84-93.
6. Вольнский Ю.Д., Курочкина А.И., Титова М.И. и др. Использование многомерного анализа данных при построении медицинских экспертных систем. Вестник АМН. 1988. № 8. С. 24-31.
7. Кемпбелл Д. Модели экспериментов в социальной психологии и прикладных исследованиях. Перевод английского. М.: Прогресс, 1980.
8. Любищев А.А. Дисперсионный анализ в биологии. М.: Изд. Московского Университета, 1986.
9. Орлов А.И. О применении статистических методов в медикобиологических исследованиях. Вестник АМН. 1987. № 1. С. 88-94.
10. CONSORT Group. Improving the quality of reporting of randomised controlled trials: the CONSORT statement. JAMA. 1996. V. 276. P. 637-639.
11. Cornfield J. The university group diabetes program: a further statistical analysis of mortality findings. J. American Medical Assoc. 1972. V. 217. P. 1676-1687.
12. STATISTICA. Copyright@Statsoft, 1994.