

А.А. Орловский

НЕСКОЛЬКО УДОБНЫХ ПРАКТИЧЕСКИХ ПРИЕМОВ ДЛЯ СТАТИСТИЧЕСКОЙ ОБРАБОТКИ ДАННЫХ В МЕДИКО-БИОЛОГИЧЕСКИХ ИССЛЕДОВАНИЯХ

Институт экспериментальной патологии, онкологии и радиобиологии им. Р.Е. Кавецкого НАН Украины, Киев

В статье рассмотрены четыре простых, быстрых и эффективных приема, дающих возможность использовать традиционные статистические методы и тем самым значительно повышать информативность медико-биологических исследований.

Ключевые слова: статистическая достоверность, выборка, критерий Стьюдента, медико-биологические исследования.

В медико-биологических исследованиях, как экспериментальных, так и клинических, часто складываются ситуации, в которых традиционные “учебниковые” методы статистического анализа не дают адекватного результата и не столько помогают в работе, сколько тормозят ее. Во многих подобных случаях можно помочь делу, если применять эти методы не “в лоб”, а в изобретательно составленных комбинациях.

В частности, предлагается использовать t -критерий Стьюдента [4] для разбиения малой выборки на моды (подгруппы или знаковые критерии):

- для расчета статистической достоверности различий между малыми выборками, в которых можно предполагать смещенные или полимодальные распределения;
- для выявления основных и дополнительных причин наблюдаемых явлений;
- для расчета статистической достоверности не статистических (сущностных) гипотез, а расчет коэффициентов корреляции — для выявления основной причины наблюдаемого явления и наличия маскирующих факторов.

Необходимо лишь следить за логической непротиворечивостью этих комбинаций и правильно интерпретировать результаты вычислений. Вот несколько часто встречающихся примеров из практики.

Пример 1. Имеется малая выборка из 5–10 вариант (значений измеряемого показателя у отдельных пациентов или лабораторных животных). Измерения во всех случаях проведены тщательно, реактивы проверены, аппаратура исправна, т.е. грубые промахи исключены. Тем не менее, в выборке присутствуют варианты, существенно

отклоняющиеся от среднего арифметического в ту или другую сторону.

Вместе с тем, во-первых, эти отклонения часто бывают недостаточно велики, чтобы отбросить соответствующие значения как промахи в соответствии с хорошо известной формулой для определения случайных “отскоков”; во-вторых, даже если отклонения и достаточны, отбросить соответствующие значения все равно нельзя, т.к., по условиям задачи, грубые промахи исключены.

Понятно, что такая ситуация указывает на резко смещенное или полимодальное распределение измеряемой величины. Однако хорошо известно, что для математически корректного построения распределения необходимо не менее 30 вариант, а у нас их 10 или того меньше, а дополнительный набор данных неминуемо растянется на недели или месяцы, а ответ нужно дать не просто сегодня, но в течение нескольких минут — иначе весь эксперимент пойдет прахом...

В описанном случае рекомендуется следующая последовательность действий.

1. Обработать всю выборку в целом (например, 8 вариант) по Стьюденту и записать результат в форме $M_0 \pm m_0$ [4].

2. “На глаз” выделить подгруппы (моды) наиболее близких друг к другу вариант (например, подгруппу 1 из 4 вариант, подгруппу 2 из 3 вариант и подгруппу 3 из 1 варианты).

3. Каждую из подгрупп, содержащих не менее 3 вариант, отдельно обработать по Стьюденту и записать результаты в виде (в данном случае) $M_1 \pm m_1$ и $M_2 \pm m_2$.

4. Интерпретировать результаты следующим образом. Если, несмотря на меньшее число вариант в каждой из подгрупп по сравнению с выборкой в целом, имеют место индикаторные соотношения $m_1 \leq m_0$ и $m_2 \leq m_0$, разделение на подгруппы выполнено правильно и в дальнейшем, возможно оценивать каждую из выделенных подгрупп в отдельности. Если хотя бы одна из величин m_1 и m_2 превосходит m_0 — попытаться выбрать другой состав подгрупп. Если при таких попытках оказалось невозможным добиться выполнения индикаторных соотношений —

интерпретировать имеющуюся выборку как единую (мономодальную) совокупность, не подлежащую разделению на подгруппы (моды).

Практика показывает, что каждый исследователь, применив такой прием хотя бы дважды или трижды, научится разделять выборку на подгруппы или интерпретировать ее как мономодальную совокупность практически безошибочно с первой попытки.

Пример 2. Хорошо известно, что распределения подавляющего большинства биологических величин в популяции (и, соответственно, в каждой из подопытных групп) весьма далеки от гауссианы, поэтому обработка биологических данных с помощью t -критерия Стьюдента (строго ориентированного на Гауссово распределение) в общем случае некорректна. Тем не менее, именно этот критерий наиболее часто используется в медико-биологических исследованиях и в большинстве случаев дает более или менее осмысленные результаты. Надежды исследователей на такой исход дела оправдываются до тех пор пока распределение, измеряемой величины в каждой из групп сравнения, мономодально, и, кроме того, либо:

а) смещение центров распределений в группах сравнения пренебрежимо мало,

б) распределения в контрольной и опытной группах смещены в противоположные стороны.

Если все распределения остаются мономодальными, но их смещения не малы и/или однонаправлены, пользуются одним из критериев для смещенных распределений — чаще всего, критерием χ^2 [2]. Если же распределение полимодально (как, например, это имеет место для числа метастазов и объема метастатических поражений, а также для множества других медико-биологических показателей), применение даже смещенных (но мономодальных) критериев чревато недопустимыми ошибками.

В таких случаях используют непараметрические (не зависящие от формы распределения) критерии — например, Вилкоксона-Манн-Уитни [4], Колмогорова или Колмогорова-Смирнова [3]. Такие критерии являются обычно ранговыми, т.е. требуют существования непрерывной функции распределения измеряемой величины (т.е., прежде всего, измеримости этой величины в конкретных физических единицах — метрах, вольтах, молях и т. п.), а иногда и одинакового вида этой функции в группах сравнения. Соответственно, их применение не вполне корректно, если функция распределения имеет разрывы (например, кусочно-непрерывна). Однако, например, функции распределения по количеству метастазов и объему метастатических поражений именно кусочно-непрерывны — в любой популяции существует несколько дискретных

групп организмов, внутри каждой из которых эти параметры могут принимать любое значение в пределах характерного интервала. Кроме того, такие критерии обычно требуют довольно большого числа вариантов, что не всегда осуществимо.

К числу непараметрических относят обычно и знаковые критерии (критерий знаков и близкий к нему точный метод Фишера), которые требуют не количественной измеримости, а лишь регистрации наличия (+) или отсутствия (-) некоторого признака (опухоли у данного животного, определенного гистохимического окрашивания в данной клетке и т.п.). Это — предельный случай непараметрического критерия, так как такой критерий не просто не зависит от параметров распределения — *само понятие параметров распределения для него не определено, не существует*. Потому, что в исходной форме их невозможно применить к величинам, измеряемым в конкретных физических единицах.

Однако, как и положено в диалектике, крайности сходятся. Чтобы пользоваться знаковыми критериями для сравнения любых групп по любому показателю, достаточно, как это успешно делается в эконометрии, ввести в совершенно непараметрический знаковый критерий параметрическую составляющую. Тогда критерий станет *полупараметрическим* (semiparametric). Это, казалось бы, “заумное” пожелание на самом деле исполнить очень просто. Для этого рекомендуется следующая процедура.

1. Вычислить среднее арифметическое значение измеряемой величины в контрольной (только в контрольной!) группе (CA_K).

2. В контрольной и опытной группах по отдельности определить число вариантов, равных или превышающих (не важно, на сколько именно) CA_K , и число вариантов, меньших (не важно, на сколько именно) CA_K .

3. Значения, равные или превышающие CA_K , отнести к категории (+), а меньшие CA_K — к категории (-).

4. Подставить количества знаков (+) и (-), найденные для каждой из групп сравнения, в формулу критерия знаков или точного метода Фишера.

Понятно, что, в зависимости от задачи исследования, вместо CA_K может быть использована другая нормировка — например, нормальное значение измеряемого показателя, взятое из справочника.

Таким способом, как показывает практика, в подавляющем большинстве случаев можно избежать как ложно-отрицательных, так и ложно-положительных выводов о существовании различия между опытной и контрольной группами.

Пример 3. Статистическая оценка достоверности сущностных гипотез (касающихся не статистической

достоверности различий между выборками, а механизмов наблюдаемого явления) в медико-биологической литературе практически не встречается. Сегодня для этой цели медикам и биологам предлагается единственный, Байесовский, подход который изначально оперирует понятиями априорной и апостериорной вероятности. Эти термины, вполне прозрачные для математика, у большинства естествоиспытателей вызывают инстинктивное неприятие.

Исправить описанное положение можно с помощью знаковых критериев. Для этого рекомендуется следующая процедура.

1. В серии из n опытов (т.е. полных экспериментов или отдельных измерений в пределах одного эксперимента), проведенных для проверки высказанной гипотезы, определить число опытов, не противоречащих гипотезе [категория (+)] и число опытов, противоречащих гипотезе [категория (-)].

2. Сформировать серию из n воображаемых опытов, в которой все n опытов противоречат гипотезе [т.е. все опыты относятся к категории (-)].

3. Сформировать серию из n воображаемых опытов, в которой категории (+) и (-) равномошны, т.е. ровно $\frac{n}{2}$ опытов не противоречат гипотезе, а остальные $\frac{n}{2}$ противоречат ей.

4. Подставив соответствующие значения в формулу одного из знаковых критериев, вычислить достоверность отличия реальной серии опытов от воображаемой серии, описанной в пункте 2.

5. Аналогичным образом вычислить достоверность отличия реальной серии опытов от воображаемой серии, описанной в пункте 3.

6. Интерпретировать результаты вычислений следующим образом. Если отличие реальной серии от воображаемых как по п. 2, так и по п. 3 высоко достоверно ($P \leq 0,05$) — считать механизм, указанный в проверяемой гипотезе, определяющей и, возможно, единственной причиной зарегистрированного явления. Если отличие реальной серии от воображаемой по п. 2 находится на грани достоверности ($P \leq 0,05$), а от воображаемой по п. 3, явно не достоверно ($P \geq 0,05$) — считать механизм, указанный в проверяемой гипотезе, реально существующим, но не определяющим и тем более не единственным. Если отличие реальной серии от воображаемых как по п. 2, так и по п. 3 явно не достоверно ($P \geq 0,05$) — считать механизм, указанный в проверяемой гипотезе, не существующим, а гипотезу, соответственно, ошибочной. Понятно, что в промежуточных случаях и интерпретации будут носить промежуточный характер.

Очевидно, что рекомендованная здесь процедура, по сути очень близка к Байесовской, однако

отличается от нее тем, что не оперирует явно понятиями априорной и апостериорной вероятности; позволяет выявить не только справедливость или ошибочность проверяемой гипотезы, но и существование дополнительных механизмов, участвующих в формировании изучаемого явления.

Столь же очевидно, что в реальной исследовательской практике далеко не всегда возможно применить вышеописанную процедуру полностью. В самом деле, для этого необходимо располагать серией из четного числа (причем не менее шести) опытов, направленных на проверку данной гипотезы. Для многочисленных случаев, когда это условие невыполнимо, можно рекомендовать упрощенный вариант процедуры, дающий лишь оценку (несколько огрубленную) правильности гипотезы, но не позволяющий выявить дополнительные механизмы явления. В упрощенном варианте процедуры выполняются лишь пункты 1, 2, 4 и 6 (последний, понятно, в упрощенном варианте). В таком исполнении процедура приобретает характер так называемой односторонней оценки. Для ориентировки читателя отметим, что в случае, когда из трех проверочных опытов все три не противоречат гипотезе, уровень достоверности последней, рассчитанный по точному методу Фишера в соответствии с упрощенной процедурой, составляет ровно 95% ($P=0,05$).

Пример 4. В любом руководстве по прикладной статистике содержатся предостережения, указывающие, что высокие значения коэффициента корреляции (линейной, нелинейной, ранговой) свидетельствуют лишь о наличии некоторой связи между явлениями, но отнюдь не о прямом причинно-следственном характере такой связи. Однако в тех случаях, когда в исследовании сравниваются между собой не менее шести — семи опытных и контрольных объектов (групп или индивидуумов), основанное на корреляциях суждение о причинно-следственной связи двух показателей можно сделать достаточно обоснованным с помощью простого приема, описанного в нижеследующем примере.

В институте экспериментальной патологии, онкологии и радиобиологии им. Р.Е. Кавецкого НАН Украины проводилось исследование связи между концентрацией клеток в опухолевом асците у мышей и объемным электрическим потенциалом единицы объема (1 мл) асцита. В эксперименте находилось семь групп животных. Были получены следующие данные (табл.).

На основании этих данных вычисляли *ранговый* коэффициент корреляции Спирмена (ρ). При расчете по всей совокупности данных получено: $\rho \approx 0,71$; $0,05 < P < 0,1$.

Таблица. Значения объемного потенциала и концентрации клеток в асците

Группа животных	Объемный потенциал 1 мл асцита, 10 ⁹ мВ	Концентрация клеток в асците, 10 ⁶ кл/мл
1	7,4	568
2	9,7	700
3	9,6	950
4	7,95	606
5	9,4	956
6	28,4*	2573*
7	3,9	469

*Примечание. При вычислении коэффициентов корреляции на основании данных, обработанных по Стьюденту и записанных в виде $M \pm m$, используются только значения M , а уровни статистической значимости различий между группами не учитываются. Поэтому в данной таблице приведены только значения M .

Из таблицы видно, что из общей закономерности выпадает группа 2. После ее исключения, для оставшихся шести групп было получено: $\rho \approx 0,902$; $P \approx 0,01$.

Далее, сравнивая группы 3 и 5, также замечаем небольшое несоответствие. При этом группа 5 находится в лучшем соответствии с оставшимися группами, нежели группа 3, с точки зрения *линейной* корреляции. Поэтому исключаем группу 3. После этого для оставшихся пяти групп получаем: $\rho = 1$; $P < 0,001$.

В результате проведенного корреляционного анализа был сделан следующий вывод: для групп 1,

4, 5, 6, 7 имеет место прямая причинно-следственная связь между объемным потенциалом единицы объема опухолевого асцита и концентрацией клеток в нем; в группах 2 и 3, причем в значительно большей степени в группе 2, эта связь замаскирована некоторыми дополнительными факторами.

Приведенные примеры показывают, что информативность данных медико-биологических исследований во многих случаях может быть не только количественно, но и качественно повышена с помощью математически и биологически корректной модификации традиционных методов статистической обработки.

ЛІТЕРАТУРА

1. *Вентцель Е.С.* Теория вероятностей. / Е.С. Вентцель. — М.: "Академия". 10-е изд. — 2005. — 576 с.
2. *Кендалл М., Стьюарт А.* Статистические выводы и связи. — М.: Наука. — 1973. — 443 с.
3. *Лемешко Б.Ю.* Об ошибках, совершаемых при использовании непараметрических критериев согласия / Б.Ю. Лемешко // Измерительная техника. 2004. — № 2. — С. 15–20.
4. *Медико-биологическая статистика / С. Гланц.* Пер. с англ. — М., Практика. — 1998. — 459 с.
5. *Методы множественного сравнения.* Режим доступа: http://www.statsoft.ru/home/portal/applications/medicine/multiple_comparisons.htm

ДЕКІЛКА ЗРУЧНИХ ПРАКТИЧНИХ ПРИЙОМІВ ДЛЯ СТАТИСТИЧНОЇ ОБРОБКИ ДАНИХ В МЕДИКО-БІОЛОГІЧНИХ ДОСЛІДЖЕННЯХ

А.А. Орловський

Інститут експериментальної патології, онкології і радіобіології ім. Р.Є. Кавецького НАН України, Київ
В статті розглянуто чотири простих, швидких та ефективних прийоми, що дають змогу застосувати традиційні статистичні методи з нетрадиційною метою і тим самим значино підвищувати інформативність медико-біологічних досліджень.

Ключові слова: статистична достовірність, вибірка, критерій Стьюдента, медико-біологічні дослідження.

SOME HANDY PRACTICAL TECHNIQUES FOR STATISTICAL DATA MANIPULATION IN MEDICAL AND BIOLOGICAL STUDIES

A.A. Orlovsky

R.E. Kavetsky institute of experimental pathology, oncology and radiobiology of NAS of Ukraine, Kiev
The paper deals with four of the simple, quick and effective techniques enabling one to use the traditional statistical methods to arrive at non-traditional goals and by this to enhance self-descriptiveness of medical and biological investigations.

Key words: statistical significance, sample t-test, medical and biological research.