

## Рекомендации по использованию и описанию методов статистики

1. При использовании выражений типа « $M \pm m$ » и « $M \pm \delta$ » необходимо указывать, какие величины соединены знаком « $\pm$ » [1].
2. Нельзя использовать выражения типа  $M \pm m$  для балльных признаков [1].
3. При применении t-критерия Стьюдента необходимо проверять два ограничительных условия – нормальность распределения и равенство дисперсий. В противном случае его использование недопустимо. Ошибочное использование t-критерия Стьюдента увеличивает вероятность выявить несущественные различия, то есть подтвердить недостоверную гипотезу. Например, вместо признания равноэффективными двух методов лечения один из них может быть безосновательно объявлен «лучшим» [1, 2, 3].
4. Недопустимо применение t-критерия Стьюдента для сравнения средних более чем в двух группах, без проверки двух ограничительных условий [1].
5. В среде пакета EXCEL нет процедуры проверки нормальности распределения, поэтому применение t-критерия Стьюдента без дополнительной проверки нормальности распределения недопустимо [1].
6. «Достоверность параметров» с помощью критерия Стьюдента проверить нельзя, так как он предназначен для проверки статистической гипотезы о равенстве средних при наличии двух независимых выборок (если нет указания на применение парного критерия) [4].
7. Однофакторный дисперсионный анализ (ANOVA) для несвязанных выборок может применяться только при выполнении двух ограничительных условий: не менее трех градаций фактора и не менее двух испытуемых в каждой градации; результативный признак (значение зависимой переменной) должен быть нормально распределен в исследуемой выборке [1].
8. «Неправильное» и «правильного» распределения данных не существует [4].
9. Описание данных с помощью средних арифметических и стандартных отклонений не рекомендуется делать в случае, если они не подчиняются закону нормального распределения [4].
10. Нельзя применять t-критерий Стьюдента к дискретным, балльным признакам [1].
11. При использовании в тексте статей конкретных значений достигнутого уровня статистической значимости необходимо указывать статистические критерии, для которых они вычислялись [1].
12. Утверждение, что уровень значимости принимался более 5 % – некорректно. Фактически такая формулировка означает принятие ложных выводов как истинных. Корректным будет выражение типа «Критический уровень значимости принимался равным 5 %» [1].
13. Утверждение, что уровень значимости принимался более 95 % некорректно [1].
14. Выражение « $p < 0,05$ » может быть получено только при использовании конкретных статистических критериев, например t-критерия Стьюдента, критерия  $s^2$  и т.д. Поэтому их нужно указывать [1].

15. Необходимо указывать, что подразумевается под величиной “р” – уровень значимости или доверительной вероятности. Если  $P > 0,95$  (0,99, 0,999) или  $p < 0,05$  (0,01, 0,001), то различия статистически значимы, если  $P < 0,95$  (0,99, 0,999) или  $p > 0,05$  (0,01, 0,001), то статистически не значимы [1].

16. Не рекомендуется использовать слово «достоверно» применительно к результатам проверки статистических гипотез, поскольку в ГОСТ Р 50779.10–2000 «Статистические методы. Вероятность и основы статистики. Термины и определения» нет такого термина как «достоверность». Следует использовать выражения "статистически значимые различия", "статистически значимо" или "статистически не значимо" [1, 2, 4].

17. При упоминании о применении многомерных методов статистического анализа необходимо конкретизировать эти методы и описание результатов их использования [1].

18. При представлении данных корреляционного анализа нужно конкретизировать название используемого коэффициента корреляции [1].

19. Нельзя использовать выражения «непараметрические данные», «непараметрические признаки», «параметрические признаки» [4].

20. В разделе «Материал и методы» необходимо указывать только те методы статистического анализа, результаты использования которых приводятся в тексте статьи [1].

21. Вероятности более 1 не существует, поэтому выражение « $p > 1$ », в котором величина «р» является вероятностью, просто бессмысленно [1].

22. Нужно указывать, что подразумевается под стандартными (общепринятыми) методами и программами обработки информации. Недопустимо описание статистических методов и критериев в форме шаблонных клише с весьма расплывчатым содержанием (“Результаты обработаны статистически”, “Результаты опытов обработаны статистически общепринятым способом, различие считали при уровне вероятности не менее 95 %” и т.д.) [1, 4].

23. Для читателя ссылка на определенную фамилию (Лакин, Урбах, Стрелков) ни о чем не говорит, поскольку в книгах данных авторов описано много разнообразных методов. В случае применения достаточно известных статистических методов (например, t-критерия Стьюдента) отсутствие подобной ссылки не так уж важно. Однако если автор использует не столь популярный метод или критерий, нужно указывать литературу и конкретные страницы [1].

24. Использование выражения «обработано на персональном компьютере» излишне. Нужно указывать конкретную программу [1].

25. Без вариации нет статистики, поэтому от термина «вариационная статистика» было рекомендовано отказаться еще в 1973 г. [4]. 26. Целесообразнее указывать конкретное значение достигнутого уровня значимости для использованного статистического критерия. Например, достигнутый уровень значимости  $p=0,04$ ;  $p=0,0004$  [1].

### **Литература:**

1. Леонов В. П. Статистика в кардиологии. 15 лет спустя // Медицинские технологии. Оценка и выбор. 2014. № 1. С. 17-28. [Электронный ресурс]. URL: [http://www.ebm.am/Portals/0/Publications/ru/Leonov\\_2.pdf](http://www.ebm.am/Portals/0/Publications/ru/Leonov_2.pdf) (дата обращения: 06.04.2018).

2. Леонов В.П., Ижевский П.В. Применение статистики в статьях и диссертациях по медицине и биологии. Часть I. Описание методов статистического анализа в статьях и диссертациях // Международный журнал медицинской практики. 1998. № 4. С. 7-12. [Электронный ресурс]. URL: <https://www.mediasphera.ru/journals/mjimp/98/4/r4-98-1.htm> (дата обращения: 06.04.2018).

3. Хлопова А.Е., Щербакова И.В. Условия применимости t-критерия Стьюдента в медицине // Всероссийская научно-практическая интернет-конференция студентов и молодых учёных с международным участием «YSRP-2014». [Электронный ресурс]. URL: <https://medconfer.com/node/4411> (дата обращения: 06.04.2018).

4. Гржибовский А. М. Использование статистики в российской биомедицинской литературе // Экология человека. 2008. № 12. С. 55-64. [Электронный ресурс]. URL: <https://cyberleninka.ru/article/n/ispolzovanie-statistiki-v-rossiyskoy-biomeditsinskoy-literature> (дата обращения: 06.04.2018).

Наибольшее количество работ, посвященных оценке качества публикаций, опубликовано В. П. Леоновым и В. В. Власовым, причем если второй больше акцентирует внимание на дизайн исследований, то первый детально описывает применение статистических методов, а также ошибки при их использовании или описании, причины ошибок, историю использования статистики в биомедицинских исследованиях в СССР и России, а также возможные пути выхода из сложившейся ситуации. Многие из публикаций В. П. Леонова представлены на созданном им сайте доказательной биологии и медицины «Биометрика» <http://www.biometrica.tomsk.ru/index.htm>. Пропагандой доказательной медицины, грамотного планирования исследований, обработки данных и представления результатов в Украине активно занимается К. П. Воробьев, чьи работы находятся в свободном доступе в интернете на сайте <http://www.vkp.dsip.net/Main.htm>. Читателям, особенно аспирантам, будет полезно узнать, что на сайте <http://www.biometrica.tomsk.ru/index.htm> существует специальный раздел «Кунсткамера», в котором В. П. Леонов «выставляет» наиболее неграмотные диссертации с детальным анализом ошибок обработки и представления данных. Ознакомление с этим разделом рекомендуется всем будущим авторам без исключения, хотя бы с целью профилактики попадания в него, а также как стимул для ознакомления со специальной статистической литературой или приглашения биостатистика для совместного планирования исследования и обработки данных для будущей диссертации.

**Шесть особенностей, связанных с вероятностью того, что опубликованные результаты (подразумеваются значимые результаты) не являются истинными (по J. Ioannidis // PLoS Medicine. 2005. Vol. 2. P. 696–701):** 1) чем меньше исследование, тем выше вероятность того, что полученный результат будет ложным; 2) чем меньше эффект изучаемого воздействия на изучаемый исход, тем выше вероятность получения ложных результатов; 3) чем больше взаимосвязей проверяется в ходе исследования, тем выше вероятность получения ложных зависимостей; 4) чем больше неопределенности в определении изучаемого состояния, а также способов его изучения, тем выше вероятность получения ложного результата; 5) чем выше финансовая заинтересованность исследователей, а также чем прочнее устоявшееся мнение по поводу изучаемой проблемы, тем больше вероятность получить ложные результаты и 6) чем более актуальной и злободневной является тема исследования, тем больше вероятность получения ложного результата.

## Литература:

Гржибовский А. М. Использование статистики в российской биомедицинской литературе // Экология человека. 2008. № 12. С. 55-64. [Электронный ресурс]. URL: <https://cyberleninka.ru/article/n/ispolzovanie-statistiki-v-rossiyskoy-biomeditsinskoy-literature> (дата обращения: 06.04.2018).

## Принципы описания статистики в публикациях:

1. **описание должно иметь цель.** Например, приводя результаты **АНАЛИЗА**, мы можем использовать эти результаты для целей **ОПИСАНИЯ**.

2. **описание должно соответствовать контексту.** Согласно этому же принципу необходимо тщательно согласовывать между собой всю терминологию присущую именно той модели, о которой идет речь.

3. ("**принцип трех ножек табуретки**"): **описания разных этапов должны взаимно дополнять друг друга.** Если в списке использованных методов не упомянут некий конкретный метод и цель и его использования, то нет смысла и обсуждать полученный с его помощью результат, и наоборот.

4. **степень детализации и объем описания статистики в публикации должны быть адекватны той роли, которую использованные методы сыграли в получении обсуждаемых в работе выводов.**

5. **помните, что читателей во много раз больше, чем авторов публикаций.** Из этого следует, что рано или поздно найдется опытный читатель, который обнаружит ваши ошибки или неточности описания.

## Описание использованных методов статистики (общая часть)

Следует различать собственно описание использованных методов от описания результатов их применения. Обычно краткое описание использованных методов статистики состоит из перечисления этих методов в традиционном разделе (главе) "Материалы и методы". Здесь же приводятся данные о размере (объеме) выборки, а также соответствующих групп и подгрупп, название статистических методов и критериев и значение критического уровня значимости. Уместно дать и обозначение наиболее часто встречающихся значений параметров, как выборочных, так и генеральных. Поскольку до сих пор биомедицинская отрасль не имеет каких-либо стандартов в обозначении таких параметров, в отличие от физико-математических и технических отраслей, где такие стандарты де-юре и де-факто имеются, следует привести названия и обозначения таких параметров. В частности, в большинстве отраслей знания принято генеральные параметры в популяции обозначать буквами греческого алфавита, тогда как выборочные параметры соответствующими буквами латинского алфавита.

Приводим ниже пример такого описания, отражающего личную точку зрения автора, и никоим образом не претендующего на "истину в последней инстанции".

"В работе анализировалась выборка объемом 360 наблюдений. Каждое наблюдение содержало 56 переменных, из которых 26 признаков являлись дискретными, качественными, и 30 признаков – непрерывными, количественными. Доля пропущенных значений составила менее 1%. Использовались следующие методы статистического

анализа: проверка нормальности распределения количественных признаков с использованием критерия Колмогорова-Смирнова с поправкой Лиллиефорса и критерия Шапиро-Уилка; проверка равенства генеральных дисперсий с помощью критериев Фишера и Кохрэна; анализ таблиц сопряженности; ранговый корреляционный анализ; непараметрический дисперсионный анализ Краскела-Уоллиса и Фридмана; лог-линейный анализ; логистическая регрессия для мультиномиальных откликов. Выборочные параметры, приводимые далее в таблицах, имеют следующие обозначения:  $M$  – среднее,  $s$  – стандартное (среднеквадратичное) отклонение,  $m$  – ошибка среднего,  $n$  – объем анализируемой подгруппы,  $p$  – достигнутый уровень значимости. Критическое значение уровня значимости принималось равным 5%. Анализ данных производился с помощью пакетов программ SAS 9 и SPSS 14."

Обратите внимание, что в приведенном выше примере нет перечисления всех остальных параметров и показателей, которые могут быть получены в отдельных конкретных видах анализа. Это вызвано тем, что более уместно о них сказать именно в контексте обсуждения результатов этого анализа. Например, приводя результаты рангового корреляционного анализа можно сказать: "Наблюдалась положительная корреляционная связь между признаками . . . . Коэффициент корреляции Спирмена  $r_s=0,79$  ( $p=0,002$ )."

При описании качественных признаков следует указать их природу – ранговые, порядковые это признаки, или же номинальные, классификационные. Привести значения каждой градации, уровня такого признака. Очевидно, что если в массиве данных много признаков, то такое описание может занять довольно много места. Весьма желательно показать соотношения отдельных подгрупп наблюдений по группирующим признакам, например по полу. Обычно это отражают в таблицах, давая значения доли в процентах для наблюдений по каждой подгруппе.

При описании количественных признаков для характеристики исследуемых популяций приводят выборочные оценки соответствующих генеральных параметров, в частности мер положения (центральной тенденции), мер рассеяния и мер формы распределения. Чаще всего это выборочные средние  $M$ , медиана  $Me$ , мода  $Mo$ , стандартное отклонение  $s$ , ошибка среднего  $m$ , и межквартильный (интерквартильный) размах. Говоря о законе распределения, и в частности о результатах проверки гипотезы о нормальности распределения, можно пойти двумя путями. В первом случае ограничиться только тем, что привести значения использованных критериев проверки нормальности и значений " $p=...$ " с соответствующим комментарием. Во втором же случае помимо этой информации привести еще и графики. Это может быть либо гистограмма с линией ожидаемой нормальной функции плотности или функции распределения, либо же так называемый график распределения на "вероятностной бумаге".

Автор рекомендует остановиться на последнем графике, поскольку помимо соответствия (несоответствия) нормальному распределению здесь можно наблюдать также возможные группировки наблюдений, аномальные выбросы, наличие симметрии и другие нюансы, заслуживающие дальнейшего обсуждения. Предположим, что по результатам проверки нормальности нулевая гипотеза была отвергнута. График, о котором идет речь выше, позволяет предположить, что мы имеем дело с лог-нормальным распределением. После проведения соответствующего преобразования можно вновь привести аналогичный график как свидетельство правильности такого предположения.

Нередко автор публикации стоит перед выбором: использовать ли для описания выборочных параметров обозначение  $(M \pm m)$  или же  $(M \pm s)$ , где  $m$  – ошибка среднего,  $s$  –

среднеквадратичное отклонение. В некоторых рекомендациях отдается предпочтение первому варианту, в других – второму. Учитывая, что "m" и "s" связаны между собой соотношением  $m=s/\sqrt{n}$ , ( $\sqrt{n}$  – корень квадратный из объема выборки "n"), спор об этом представляется достаточно схоластическим. Выбор того или иного варианта должен в большей степени определяться контекстом. Например, если речь идет о сравнении групповых средних, то логичнее использовать первый вариант. Если же просто приводятся выборочные характеристики отдельных групп, то разумнее использовать второй вариант. Важно, чтобы как для первого, так и для второго случая всегда приводилось значение объема выборки "n". Тогда зная, например, "m" всегда можно вычислить и "s", и наоборот.

В последнее время часть авторов использует доверительные интервалы для описания точности оценок, например среднего. Такие оценки обычно представляют собой обычно двусторонние 95%-ные интервалы, задаваемые выражением  $M \pm 1,96m$ . Однако авторы при этом забывают, что при этом необходимо указать величину доверительной вероятности. В ряде случаев доверительный интервал представляет самостоятельный интерес, в частности, при анализе частотных соотношений. В этом случае достаточно информативен такой признак, как ЧБНЛ – "Число Больных, которых Необходимо Лечить определенным методом в течение определенного времени, чтобы достичь определенного эффекта или предотвратить неблагоприятный исход". Данный показатель находится как обратная величина для одной из границ доверительного интервала для доли, пропорции.

### **Описание использованных методов статистики (специальная часть)**

Помните, что содержание описания выводов о применении любого статистического метода помимо собственно информации о полученных результатах, всегда дает читателю информацию и об уровне вашего владения этим методом. Поэтому старайтесь описывать только то, что сами хорошо понимаете. Чтобы такого не было, организуйте свое исследование так, чтобы анализ данных проводил профессионал в данной области, который и даст вам описание этого этапа. Если же вы сомневаетесь в правильности или корректности собственных формулировок, то обратитесь за консультацией к специалисту в этой области.

**рекомендации по описанию более 50 методов и критериев**, наиболее часто используемых в статистическом анализе биомедицинских данных – <http://www.biometrica.tomsk.ru/recom.pdf>

### **Литература:**

Леонов В. Три "Почему ..." и пять принципов описания статистики в биомедицинских публикациях (исправленная версия от 10.12.06) [Электронный ресурс]. URL: <http://www.biometrica.tomsk.ru/principals.htm> (дата обращения: 06.04.2018).