

УДК 519.2:303.732.4

UDC 519.2:303.732.4

01.00.00 Физико-математические науки

Physics and mathematical sciences

**СТАТИСТИЧЕСКИЕ МОДЕЛИ В
МЕДИЦИНЕ****STATISTICAL MODELS IN MEDICINE**

Орлов Александр Иванович
д.э.н., д.т.н., к.ф.-м.н., профессор
РИНЦ SPIN-код: 4342-4994
*Московский государственный технический
университет им. Н.Э. Баумана, Россия, 105005,
Москва, 2-я Бауманская ул., 5, prof-orlov@mail.ru*

Orlov Alexander Ivanovich
Dr.Sci.Econ., Dr.Sci.Tech., Cand.Phys-Math.Sci.,
professor
*Bauman Moscow State Technical University,
Moscow, Russia*

Статья посвящена обсуждению организации клиничко-статистических исследований и экспериментов. Рассмотрены примеры применения статистических методов в научных медицинских исследованиях. Под клиничко-статистическим исследованием понимают специально организованный сбор и анализ медицинских данных о течении заболеваний у пациентов, о динамике объективных и субъективных показателей их состояния, о реакции на те или иные лечебные воздействия. Исследуются одна, две или более групп лиц (больных или здоровых), выводы делаются по группам в целом, а не по каждому конкретному пациенту. Цель исследований - перенести выводы, сделанные для выборки, на генеральную совокупность», т.е. клиничко-статистическое исследование ориентировано на получение полезных рекомендаций, касающихся тех пациентов, которые попадут в поле зрения врачей после окончания исследования. Существует два основных вида исследований - ретроспективные и проспективные. Первые связаны с анализом прошлого пациентов, вторые - с наблюдением за течением их заболеваний в будущем. Рассмотрены типовые ошибки при организации клиничко-статистических исследований. При планировании исследования, обычно выделяют экспериментальную и контрольную группы, одинаковые или близкие по всем показателям, кроме изучаемого фактора (воздействия). Обсуждаются различные варианты слепого метода. Рассмотрено применение статистических моделей и методов в научных медицинских исследованиях. Разобраны примеры доверительного оценивания доли (вероятности) и проверки однородности вероятностей. Для статистического моделирования в случае малых вероятностей используют распределение Пуассона. С его помощью проанализированы статистические данные о заболеваемости описторхозом

The article is devoted to the discussion of the organization of clinical-statistical studies and experiments. We have considered the examples of the application of statistical methods in scientific medical research. Under the clinical-statistical research we understand specially organized collection and analysis of medical data about the course of disease in patients, research of the dynamics of objective and subjective indicators of the state of reaction to these or other therapeutic effects. We study one, two or more groups of individuals (patients or healthy), conclusions are drawn on the whole group, but not for each individual patient. The purpose of research - to transfer the conclusions reached for the sample to the general population, i.e., clinical and statistical study focused on the production of useful recommendations concerning those patients who fall into the field of view of doctors after the end of the study. There are two main types of research - prospective and retrospective. The first related to the analysis of the last patients, the second - to monitoring the course of their disease in the future. We have considered typical mistakes in the organization of clinical-statistical studies. When planning a research, we usually distinguish the experimental and control groups, which are identical or similar in all respects except for the studied factors (exposure). We discuss the various options for blind methods and consider the application of statistical models and methods in scientific medical research. We have analyzed examples of confidence estimation of proportion (probability) and the homogeneity test of probabilities. For statistical modeling we use the Poisson distribution in the case of small probability. With its help, we analyze statistical data on the opisthorchiasis

Ключевые слова: СТАТИСТИЧЕСКИЕ МОДЕЛИ, МЕДИЦИНА, МАТЕМАТИЧЕСКАЯ СТАТИСТИКА, КЛИНИКО-СТАТИСТИЧЕСКИЕ ИССЛЕДОВАНИЯ, ДОВЕРИТЕЛЬНОЕ

Keywords: STATISTICAL MODELS, MEDICINE, MATHEMATICAL STATISTICS, CLINICAL-STATISTICAL RESEARCH, CONFIDENCE ESTIMATION OF PROBABILITY,

ОЦЕНИВАНИЕ ВЕРОЯТНОСТИ, ПРОВЕРКА
ОДНОРОДНОСТИ ВЕРОЯТНОСТЕЙ,
РАСПРЕДЕЛЕНИЕ ПУАССОНА, ОПИСТОРХОЗ

HOMOGENEITY PROBABILITIES TESTING,
POISSON DISTRIBUTION, OPISTHORCHIASIS

Doi: 10.21515/1990-4665-124-063

1. Введение. Клинико-статистические исследования

Обсудим организацию клинико-статистических исследований и экспериментов. Рассмотрим примеры применения статистических методов в научных медицинских исследованиях.

Под клинико-статистическим исследованием понимают специально организованный сбор и анализ медицинских данных о течении заболеваний у пациентов, о динамике объективных и субъективных показателей их состояния, о реакции на те или иные лечебные воздействия. Исследуются одна, две или более групп лиц (больных или здоровых), выводы делаются по группам в целом, а не по каждому конкретному пациенту. Цель исследований - перенести выводы, сделанные для выборки, на генеральную совокупность», т.е. клинико-статистическое исследование ориентировано на получение полезных рекомендаций, касающихся тех пациентов, которые попадут в поле зрения врачей после окончания исследования. Таким образом, имеется потенциальное противоречие интересов практикующего врача и научного работника, проводящего клинико-статистическое исследование. Первый заинтересован оказать наилучшую возможную помощь каждому своему пациенту. Второй разрабатывает рекомендации для будущих больных, а для этого иногда, например, при использовании плацебо (о чем подробно рассказано ниже), вынужден лишить пациентов лечебного воздействия.

2. Сбор данных и «карта больного»

По каждому отдельному пациенту подобная информация обычно содержится в его истории болезни. Поэтому при проведении клинико-

статистического исследования сбор данных зачастую состоит в заполнении специальной формы (карты больного) на основе истории болезни. Такая работа сильно облегчается, если в медицинской организации ведутся автоматизированные истории болезни со стандартными процедурами заполнения. В идеале сбор данных для конкретного клинико-статистического исследования может быть заменен на выделение соответствующей подбазы данных из общей базы медицинской организации, содержащей автоматизированные истории болезни.

Карта больного - один из видов анкет. Начальные вопросы позволяют идентифицировать пациента (фамилия, имя, отчество, адрес, паспортные данные), другие нацелены на получение информации, используемой при лечении. Обычно фиксируют:

пол, возраст (дату рождения), профессию (вид деятельности);

данные о перенесенных заболеваниях (анамнез), вредных привычках (курение, употребление иных наркотических веществ);

рост, вес, пульс, температура, диастолическое и систолическое артериальное давление, результаты ЭКГ, других обследований и анализов при поступлении в медицинское учреждение и в другие моменты времени;

течение заболевания и назначенное лечение;

даты начала заболевания, госпитализации, операции, выписки, смерти и т.д.

Внешне карта больного похожа на личный листок по учету кадров. Отличие в том, что она предназначена для получения выводов по группе пациентов, а не отдельно по каждому из них. Поэтому записи в каждой графе должны вестись так, чтобы на их основе можно было формулировать выводы по группе в целом. Другими словами, записи должны быть либо числами, либо кодами градаций из заранее составленных списков.

С формальной точки зрения карта больного напоминает маркетинговые и социологические анкеты, а также анкет-коды,

применяемые при статистическом анализе исторических событий. Только вопросы в них другие, соответствующие предметным областям.

3. Виды клинико-статистических исследований

Существует два основных вида исследований - ретроспективные и проспективные. Первые связаны с анализом прошлого пациентов, вторые – с наблюдением за течением их заболеваний в будущем.

Исходный материал для ретроспективных исследований - истории болезней. Чего в них нет - восстановить нельзя. В прошлом нельзя ни повторно опросить больных, ни провести дополнительные анализы и обследования. Это – недостаток ретроспективных исследований.

Преимущества ретроспективных исследований:

- нет необходимости ждать развития заболевания и проявления результата лечебного воздействия, в связи с чем клинико-статистическое исследование может быть проведено быстро (при хорошо поставленной системе автоматизированных историй болезни - за один день);

- отсутствие этических проблем, связанных, например, с разделением пациентов на группу вмешательства и группу сравнения.

При планировании проспективного исследования можно предусмотреть сбор всей необходимой информации, проведения всех требуемых врачу анализов и обследований.

Преимуществом является также возможность обеспечить сопоставимость группы вмешательства и группы сравнения, например, применить рандомизацию или «двойной слепой метод». С одной стороны, это дает возможность статистически достоверно (т.е. с вероятностью, близкой к единице) обосновать наличие (отсутствие) эффекта. С другой стороны, часть пациентов будет получать плацебо, а не лекарство, т.е. будет лишена лечебного воздействия, что не всегда корректно с этической точки зрения.

К недостаткам проспективных исследований относится большая продолжительность исследования.

4. Формулировки вопросов

Как и в маркетинговых и социологических опросах, в научных медицинских исследованиях в «картах больного (пациента)» используют три типа вопросов - закрытые, открытые и полужакрытые, они же полуоткрытые.

Если в исследовании практически для всех вопросов варианты ответов можно перечислить заранее, то целесообразно широко использовать закрытые вопросы. И в общем случае целесообразно применять в основном закрытые и полужакрытые вопросы. Как показывает опыт проведения научных медицинских исследований, такой подход обычно является правильным - лишь в небольшом числе карт оказываются вписаны свои варианты ответов. Вместе с тем демонстрируется уважение к мнению врача-исследователя, - он может добавить свое суждение, но обычно редко пользуется этой возможностью (как правило, не более чем в 5% случаев).

5. Типовые ошибки при организации клинико-статистических исследований

Как показывает практика консультаций специалиста по статистическим методам, оказываемых медикам, начинающие врачи-исследователи иногда недостаточно продумывают процедуры сбора и анализа информации.

Например, начинают бессистемно выписывать те или иные сведения, которые, по их мнению, представляют интерес для клинико-статистического исследования. Затем приходят к специалисту по статистическим методам анализа данных и просят обработать. Тот дает

заключение: материал обработке не поддается. Масса пропусков - для одних пациентов информация по определенному вопросу карты больного есть, для других – нет, следовательно, ответы для разных пациентов не сопоставимы. Наборы анализов, обследований, лечебных воздействий - у каждого свои. В результате обработать удается примерно 1% собранной информации. (Например, обычно удается установить численность изученной группы пациентов.)

Другая типичная ошибка - в начале работы начинающий врач-исследователь использует одно определение термина, потом постепенно от него отходит.

Иногда разные научно-практические коллективы, в том числе в медицине, по-разному определяют понятия. Например, в 1950-х и 1960-х гг. психиатры отметили разительное несовпадение уровней госпитализации с диагнозами шизофрении и маниакально-депрессивного психоза в Великобритании и США. Так, уровень госпитализации с маниакально-депрессивными психозами в возрастной группе 55...64 лет в Великобритании был в 20 раз выше, чем в США. И наоборот, шизофрения в Нью-Йорке оказалась в 2 раза более распространенной, чем в Лондоне. Причина различий оказалась простой - разные традиции диагностики, т.е. разное понимание терминов. Одно и то же заболевание в США может диагностироваться как шизофрения, а в Великобритании - как маниакально-депрессивный психоз [1].

Наиболее глубокая причина ошибок - несовершенство наших знаний. Хорошо известно, что обнаружить причину различия групп пациентов зачастую весьма трудно. Например, в середине XIX в. по уровню заболевания холерой различались богатые и бедные лондонцы. В Англии осенью и зимой холера не переходила в эпидемию, а в Шотландии - наоборот. Оказалось (и это было большое научное достижение), что заболеванию холерой способствует питье сырой воды, зараженной

холерными вибрионами [1]. Сырую воду бедные лондонцы пили гораздо чаще, чем богатые. По традиции сырая вода широко употреблялась осенью и зимой в Шотландии, в то время в Англии в это время пили чай. Представим теперь, что в то время проводилось бы исследование, посвященное заболеваемости холерой. Были бы взяты группа англичан и группа шотландцев, одинаковые во всем, кроме места жительства. И результат исследования состоял бы в том, что зимняя заболеваемость объясняется географическим положением пациентов. Аналогично было бы «доказано», что богатство приводит к снижению заболеваемости.

6. Планирование и организация сбора данных в научных медицинских исследованиях

Опыт показывает, что первое, с чего следует начинать научное медицинское исследование - формулировка цели исследования.

В типовой формулировке цели научного медицинского исследования выделяются две составляющие:

- 1) влияние какого фактора (воздействия) изучается;
- 2) какая совокупность пациентов рассматривается в исследовании.

В большинстве случаев выводы, полученные по сравнительно небольшому числу пациентов, распространяют на всех пациентов с подобными заболеваниями, т.е. выводы с выборки переносят на генеральную совокупность. При таком взгляде становится ясной важность тщательного описания генеральной совокупности. Например, генеральная совокупность пациентов поликлиники (начальные стадии заболевания, легкое течение, как следствие - амбулаторное лечение) сильно отличается от другой генеральной совокупности - пациентов больницы.

Другой пример из практики. Сельский врач был уверен в эффективности определенного антибиотика, поскольку практически все больные с пневмонией, поступавшие в ту больницу, где он работал, быстро

и хорошо этим антибиотиком вылечивались. В то же время врач областной больницы вполне обоснованно считал, что этот антибиотик бесполезен: он не смог вылечить им ни одного из больных с пневмонией, кто поступил в его отделение, и был вынужден давать этим больным другие лекарства.

В чем причина такого расхождения мнений? Дело в том, что эти два врача лечили больных из двух разных генеральных совокупностей. Сельский врач основывал свои выводы на анализе лечения больных из генеральной совокупности жителей своего участка. В областную больницу привозили тех и только тех больных, которым не смогли помочь сельские врачи, т.е. больных, для которых обсуждаемый антибиотик оказался неэффективным. Это – совсем иная генеральная совокупность. Доля таких пациентов в сельских больницах была невелика, но все они после неэффективного лечения на местах собирались в областной больнице – потому ее врач так критически об этом антибиотике и высказывался. И ведь оба были правы – на основании собственного опыта и опыта своих больниц.

7. Материалы и методы

После формулирования цели описывают материалы и методы исследования. Материалы – описание групп пациентов, прежде всего по полу, возрасту, анамнезу. Под методами понимают способы изучения течения заболевания.

Все научные исследования, в том числе и медицинские, и особенно клинико-статистические, проводятся по одним и тем же правилам. Рассмотрим типовую схему экспериментальной научной работы [2].

8. Базовая модель эксперимента

Цель эксперимента - эмпирическое подтверждение или опровержение гипотезы исследования и(или) справедливости теоретических результатов.

Рассмотрим следующую базовую модель эксперимента. Пусть имеется некоторый медицинский объект, изменение состояния которого исследуется в ходе эксперимента. В качестве объекта может выступать отдельный индивид, группа, коллектив и т.д., например, множество пациентов, лечение которых осуществляется по новой методике. Состояние объекта описывается показателями по критериям, отражающим его существенные свойства (исход заболевания, продолжительность лечения, число осложнений и т.д.).

Эксперимент заключается в целенаправленном воздействии на объект, призванном изменить его определенным образом. Собственно, это воздействие – его состав, структура, свойства и т.д. – и есть результат теоретического исследования. Примеры воздействия: применение новых лекарств и видов лечебного воздействия.

Следовательно, при проведении эксперимента необходимо обосновать изменения состояния объекта, причем в требуемую сторону. Однако этого недостаточно. Необходимо установить, что изменения произошли именно в результате произведенного воздействия.

Например, в России в 1990-е гг. резко снизилась обращаемость пациентов в лечебные заведения по поводу ряда заболеваний, особенно легких, не требующих госпитализации. Однако из этого не следует, что снизилась и заболеваемость, здоровье населения стало лучше. Судя по резкому возрастанию смертности в 1990-е гг. (см., например, [3, разд.8.3]), дело обстояло прямо противоположным образом - здоровье населения ухудшилось. А снижение обращаемости определялось социально-экономическими причинами - появлением безработицы, снижением уровня жизни населения, ухудшением работы учреждений здравоохранения. Все

эти причины приводят к тому, что население при легких заболеваниях обращается не к врачу, а к «домашним» средствам, люди продолжают работать во время течения заболевания.

Чтобы выделить в явном виде результат целенаправленного воздействия на исследуемый объект, необходимо взять аналогичный объект и посмотреть, что происходит с ним в отсутствии воздействий.

9. Экспериментальная и контрольная группы

При планировании исследования обычно выделяют экспериментальную и контрольную группы, одинаковые или близкие по всем показателям, кроме изучаемого фактора (воздействия) [4]:

Экспериментальная группа (группа лечения, группа вмешательства) - группа, подвергающаяся вмешательству (лечению) в ходе исследования.

Контрольная группа (группа сравнения) - группа испытуемых, получающих обычное лечение, или не получающих лечения, или получающих плацебо. Результаты измерений показателей состояния испытуемых в контрольной группе сравниваются с результатами аналогичных измерений в экспериментальной группе для оценки эффекта исследуемого метода лечения».

На рис. 1 приведена типовая структура медицинского эксперимента. Сплошными стрелками показаны переходы групп из начального в конечное состояние. Штриховыми стрелками и римскими цифрами отмечены процедуры сравнения показателей объектов. С помощью рис. 1 обсудим все составляющие базовой модели медицинского эксперимента.

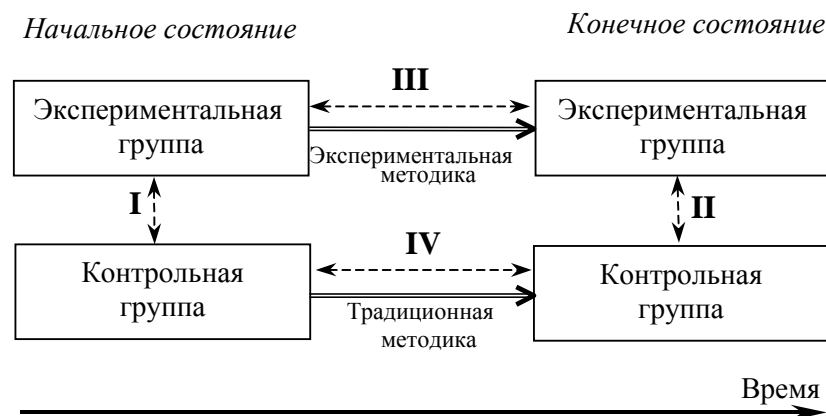


Рис. 1. Структура медицинского эксперимента

Констатации (в результате процедуры сравнения III – см. рис. 1) различий начального и конечного состояний (динамики) экспериментальной группы недостаточно, так как аналогичные изменения могут происходить и с состояниями пациентов контрольной группы, что можно установить процедурой сравнения IV. Поэтому *алгоритм* действий исследователя состоит из следующих шагов:

- 1) на основании процедуры сравнения I установить совпадение начальных состояний экспериментальной и контрольной групп;
- 2) реализовать воздействие на экспериментальную группу¹;
- 3) на основании процедуры сравнения II установить различие конечных состояний экспериментальной и контрольной групп.

Выполнение перечисленных шаг косвенным образом позволяет реализовать процедуру сравнения III, исключая влияние общих для экспериментальной и контрольной группы условий и воздействий.

¹ Необходимо быть уверенным, что экспериментальная и контрольная группы находятся в одинаковых условиях, за исключением условий, целенаправленно изменяемых исследователем.

Роль статистических методов заключается в том, чтобы корректно и достоверно обосновать совпадение или различие состояний контрольной и экспериментальной группы.

10. Многообразие схем эксперимента и статистические технологии

Эксперимент может проходить и по более сложной, но укладывающейся в рамки описанной методологии схеме – показатели пациентов контрольных и экспериментальных групп могут измеряться и сравниваться неоднократно в различные моменты времени.

Применяют и иные схемы клинико-статистических исследований. Например, в исследовании по схеме «до - после» отсутствует контрольная группа. Изучается течение заболевания в одной группе пациентов, подвергающихся изучаемому вмешательству (в терминах рис. 1 - проводится процедура сравнения III). Обычно такое исследование предназначено для первоначального анализа ситуации, является базой для организации дальнейшего исследования по схеме, приведенной на рис. 1. Иногда при проведении исследования по схеме «до - после» предполагают, что любое улучшение, наблюдаемое после лечения, обусловлено именно лечением. Это предположение может оказаться ложным, что делает схему «до - после» весьма уязвимым [4].

Проведение процедур сравнения III и IV позволяет выявить изменения в показателях пациентов экспериментальной и контрольной групп за время наблюдения. Таким образом, все четыре процедуры сравнения, указанные на рис. 1, широко используются в научных медицинских исследованиях. Расчетные методы проведения сравнений показателей (между группами и в одной группе в различные моменты времени) рассмотрены ниже (см. также [3, гл. 5]).

Разделение выборки на контрольную и экспериментальную части - не догма. В некоторых ситуациях применяют скользящий контроль, позволяющий увеличить объем данных (в экспериментальной группе) в каждом конкретном расчете путем уменьшения контрольной группы по минимуму - до одного наблюдения [5].

Отдельные методы сбора и анализа данных (как медицинских, так и иной природы) объединяются в статистические технологии. В настоящей статье рассмотрены лишь отдельные элементы статистических технологий научных медицинских исследований. Тем не менее они позволяют решать подавляющую часть задач, возникающих при анализе данных в медицинских научных исследованиях. Понятие «статистические технологии» подробно обсуждается в прикладной статистике [6, 7].

Технология распределения пациентов по группам получила название *рандомизации*. При хорошо выполненной рандомизации в каждой из групп доля пациентов с тем или иным свойством будет примерно одинакова: одинаков процент мужчин и женщин, одинаков процент лиц того или иного возраста, одинаково соотношение пациентов с различными степенями тяжести исследуемого заболевания. Способы проведения такого процесса описаны в специальной литературе, однако для полноценной рандомизации число пациентов должно быть достаточно большим - измеряться по меньшей мере тысячами.

11. Слепой метод

Базовая схема проведения научного медицинского эксперимента основана на сравнениях показателей пациентов экспериментальной и контрольной групп до и после воздействия. Могут сравниваться новая методика лечения, применяемая в экспериментальной группе, с традиционной методикой в контрольной группе. В ряде случаев, например,

при изучении эффективности нового препарата, пациенты в контрольной группе получают плацебо.

Плацебо (placebo) - лекарственная форма, которую пациент не может отличить от исследуемого препарата (по внешнему виду, цвету, вкусу, запаху и т.п.). Однако плацебо не оказывает специфического действия. Например, в качестве плацебо используют таблетки глюкозы или инъекции изотонического раствора хлорида натрия (физиологического раствора), или иное безразличное вмешательство, «воздействие без действия». Часто наблюдается плацебо-эффект - изменение состояния больного (отмечаемое самим больным или лечащим врачом), связанное с фактом лечения, а не с биологическим действием препарата. Например, плацебо, которое больной воспринимает как лекарство, уменьшает послеоперационную боль, тошноту и зуд более чем у 30% прооперированных больных [4]. Плацебо-эффект демонстрирует возможности воздействия психической сферы больного на течение заболевания и процесс лечения. Подобные эффекты наблюдаются не только в медицине, но и в других сферах деятельности, например, в менеджменте при управлении персоналом [8].

Необходимо исключить влияние психологического эффекта на результаты исследования. Проще всего это сделать, если пациентам двух групп давать совершенно одинаковые по внешнему виду, вкусу, запаху препараты, только в один из препаратов добавлять то вещество, влияние которого на человеческий организм исследуется, а во второй – ничего не добавлять. Тогда пациенты будут одинаково верить в то, что препарат им поможет, что не повлияет на конечный результат - оценку эффективности лечения лекарством. Препарат, который дается пациентам из контрольной группы и не содержит лекарства, принято называть *плацебо* (от лат. пустышка), а метод исследования с использованием плацебо - *слепым*.

Однако даже при использовании плацебо могут возникнуть проблемы: врачи, проводящие исследования, могут словом или жестом случайно дать понять пациенту, чем на самом деле его лечат, что обесценит использование плацебо. Чтобы это предотвратить, можно применить *двойной слепой* метод - т.е. не говорить об истинном содержимом препарата и лечащим врачам пациентов.

Оценку эффективности того или иного воздействия на организм человека можно осуществлять по изменившимся параметрам жизнедеятельности (например, по уровню артериального давления или концентрации в крови тех или иных веществ). Однако более привлекательным представляется оценка по наличию или отсутствию в жизни пациента благоприятных и неблагоприятных событий. Если у пациентов в группе вмешательства неблагоприятных событий в жизни было меньше, а благоприятных - больше, чем у пациентов из группы сравнения, то вмешательство являлось полезным, если же нет - вредным. Такие события получили название *конечных точек* (во многих исследованиях оценка эффективности медицинского вмешательства ведется именно по этим точкам).

Таким образом, плацебо используется в научных медицинских исследованиях для имитации лечения в целях устранения систематической ошибки, возникающей от плацебо-эффекта.

Если больные, включенные в клинико-статистическое исследование, знают, какое лечения получают, то их поведение может измениться, что приведет к появлению систематической ошибки в итоговых научных выводах. Чтобы этого избежать, используют *слепой метод*.

Выделяют пять условий, соблюдение которых позволяет во многом избавиться от нежелательных эффектов [4].

1. Те, кто формирует выборку и распределяет пациентов по группам вмешательства (в базовом случае - направляет их в две группы, она из

которых будет экспериментальной, а вторая - контрольной), не должны знать, какое лечение будет назначено каждому последующему больному, чтобы это не нарушало включение пациентов в исследование в порядке их поступления.

2. Пациенты не должны знать, какое именно лечение они получают. Тогда менее вероятны нарушения схемы лечения, субъективность при описании своего состояния.

3. Проводящие лечение врачи не должны знать, какое именно лечение назначено конкретному пациенту. Тогда не возникнут вызванные этим знаниями различия в ведении больных.

4. Врачи-исследователи, оценивающие исходы лечения, не должны знать, какое лечение было назначено тому или иному пациенту.

5. Специалисты по доказательной медицине, с помощью статистических методов сравнивающие исходы лечения в группах, не должны знать, какая из них экспериментальная, а какая - контрольная.

Когда информация о реально используемом лечебном воздействии отсутствует только у пациента, говорят о простом слепом методе. Если также и у лечащих врачей и исследователей, оценивающих исходы, - о двойном слепом методе. Если рассматриваемой информации нет и у лиц, проводящих статистическую обработку результатов исследования, говорят о тройном слепом методе. Наиболее часто используется простой слепой метод, а в хорошо организованных исследованиях - двойной слепой метод.

12. Применение статистических моделей и методов в научных медицинских исследованиях

Научные медицинские исследования можно разделить на три типа. К первому типу относятся немногочисленные работы, посвященные описанию отдельных интересных случаев. Второй тип - выборочные исследования, когда по результатам изучения сравнительно небольших

групп больных (десятки или сотни человек) формулируют выводы и рекомендации для более крупных совокупностей. Третий тип - обобщающие работы, построенные на основе ранее проведенных исследований. Как правило, именно такие работы наиболее полезны для практикующих врачей.

Методы выборочных исследований основаны на вероятностных моделях, описывающих получение ответов на вопросы, включенные в «карты больного». В случае ответов типа «есть» - «нет» наиболее распространенными являются две вероятностные модели - биномиальная и гипергеометрическая (см., например, [3, гл.1]). Эти модели *весьма близки* (с практической точки зрения совпадают), когда объем генеральной совокупности по крайней мере в 10 раз превышает объем выборки.

Биномиальная модель легче для теоретического изучения, поэтому будем именно её рассматривать в дальнейшем. Однако при проведении реального научного медицинского исследования лучше формировать выборку, исходя из гипергеометрической модели. Например, при проведении эпидемиологического исследования в целях изучения качества жизни это делают, выбирая опрашиваемых (участников исследования) из списка избирателей с помощью датчиков псевдослучайных чисел на ЭВМ или с помощью таблиц псевдослучайных чисел [9]. Алгоритмы формирования выборки встраивают во все современные программные продукты, предназначенные для поддержки проведения подобных исследований.

Объем выборки определяется финансовыми возможностями организаторов научного медицинского исследования, числом больных с соответствующими нозологическими формами, находящимися в определенном медицинском учреждении во время исследования и т.д. В отдельных случаях, когда требуется оценить определенный параметр с заданной точностью или выбрать между двумя полностью описанными

гипотезами, необходимый объем выборки можно рассчитать на основе теории математической статистики.

13. Формирование и динамика выборки

Рассмотрим конкретное научное медицинское исследование. В качестве популяции, из которой проводится выборка, рассматриваются больные с сахарным диабетом, наблюдающиеся в одной клинике. Предполагается, что клиника является типовой, и результаты могут быть перенесены на всех больных с рассматриваемой нозологической формой заболевания. Критерии отбора (включения в выборку):

1. Возраст старше 40 лет.
2. Сахарный диабет диагностирован в возрасте старше 30 лет.
3. Требуется лекарственная коррекция гипергликемии.
4. Планируемая продолжительность наблюдения - более двух лет.
5. Выполнение определенных ограничений (здесь не описываются), связанных с другими заболеваниями, инвалидностью и т.п.

Некоторая часть больных, соответствующих критериям отбора, откажется сотрудничать с исследователем. Одни пациенты не захотят лечиться предлагаемым способом. Другие станут возражать против определения выбора лечения случайным образом (при применении двойного слепого метода), а не их лечащим врачом. Пациенты-отказники обычно отличаются от давших согласие по социально-экономическому статусу, тяжести заболевания и другим признакам.

За время исследования (в данном случае за 2 года) состав выборки изменится. Отпадут пациенты, не соблюдавшие график посещения врача-исследователя, полностью или частично не проходившие регулярных обследований. Останутся лишь наиболее дисциплинированные и ответственные пациенты. Но и среди них будут выбывшие из исследования (вследствие смерти, переезда, развития другого заболевания

и т.п.). Таким образом, реальная выборка будет отличаться от намеченной. Если критериям отбора удовлетворяет достаточно много пациентов, то простейшее решение заключается в увеличении исходного объема выборки, с тем, чтобы к концу исследования обеспечить отслеживание динамики заболевания для заданного числа пациентов.

14. Доверительное оценивание доли (вероятности)

Рассмотрим проблему точности переноса выводов, сделанных для выборки, на генеральную совокупность для простейшего случая – наличия одного вопроса в карте больного с двумя возможными ответами («есть» и «нет»).

Ситуация описывается одним параметром p - долей пациентов с ответом «есть» в генеральной совокупности. Необходимо оценить этот параметр. В [3, гл.1] развита соответствующая теория. Предварительно требуется задать характеристику надежности переноса выводов с выборки на генеральную совокупность - доверительную вероятность γ , близкую к 1. Наиболее распространенным (в научных медицинских исследованиях) значением доверительной вероятности является $\gamma = 0,95$. Иногда употребляют термин «95% доверительный интервал». Тогда согласно [3, гл.1] используют множитель $U(\gamma) = 1,96$, задающий ширину доверительного интервала.

Пример 1. Пусть в выборке объемом 500 человек ($n = 500$) для 300 пациентов ($m = 300$) в карте больного ответ «есть». Тогда выборочная доля пациентов с таким ответом $p^* = m/n = 0,60$. Определить доверительный вариант для p при доверительной вероятности $\gamma = 0,95$.

Согласно формулам [3, разд. 1.3]

$$p_{\text{нижн}} = 0,40 - 1,96 \frac{\sqrt{0,6 \times 0,4}}{\sqrt{500}} = 0,60 - 0,043 = 0,557, \quad p_{\text{верх}} = 0,60 + 0,043 = 0,643.$$

Таким образом, хотя в достаточно большой выборке в карте больного ответ «есть» имеется для 60% пациентов, можно утверждать, что в генеральной совокупности таких ответов 55,7...64,3% (разность крайних значений составляет 8,6%).

Пример 2. Из 64 пациентов диагноз «шизофрения» поставлен 30 пациентам. Какие выводы можно сделать о распространенности шизофрении среди психических заболеваний?

В рассматриваемом случае $p^* = 30/64 = 0,47$. Доверительный интервал для вероятности p при $\gamma = 0,95$:

$$p_{\text{нижн}} = 0,469 - 1,96 \frac{\sqrt{0,469 \times 0,531}}{\sqrt{64}} = 0,469 - 0,122 = 0,347,$$
$$p_{\text{верх}} = 0,469 + 0,122 = 0,591.$$

Следовательно, распространенность шизофрении составляет 34,7...59,1% (35...60%). Более точную оценку можно найти на основе исследования более обширной выборки.

15. Проверка однородности вероятностей

Одна из базовых проблем, решаемых статистическими методами, - проверка однородности вероятностей. Она часто обсуждается в литературе, а методы проверки однородности применяются при решении многих практических задач, встающих перед врачом-исследователем. Например, как сравнить по тяжести заболевания две группы - мужчин и женщин, молодых и пожилых, и т.п.?

Если две группы не отличаются по ответам, значит, их можно объединить в одну и осуществлять одни и те же лечебные воздействия. Если две группы различаются, то и лечить их надо по-разному.

Постановка задачи в вероятностно-статистических терминах такова. Рассматривается вопрос «карты больного» с двумя возможными ответами, например, «есть» и «нет». В первой группе из n_1 пациентов для m_1 человек

ответ «есть», а во второй группе из n_2 больных для m_2 лиц ответ «есть». В вероятностной модели предполагается, что m_1 и m_2 - биномиальные случайные величины $B(n_1, p_1)$ и $B(n_2, p_2)$ соответственно.

Однородность двух групп означает, что соответствующие им вероятности равны, неоднородность - что эти вероятности отличаются. Необходимо проверить гипотезу однородности $H_0: p_1 = p_2$ при альтернативной гипотезе о наличии эффекта $H_1: p_1 \neq p_2$. Оценкой вероятности p_1 является частота $p_1^* = m_1/n_1$, а оценкой вероятности p_2 является частота $p_2^* = m_2/n_2$. Даже при совпадении вероятностей p_1 и p_2 частоты, как правило, различаются.

Правило принятия решения при проверке статистической гипотезы однородности двух выборок приведено в [3, гл. 1]. В клинко-статистических и медико-биологических наиболее распространен 5% уровень значимости, т.е. $\alpha = 0,05$.

Пример 3. Пусть из 500 мужчин определенное нарушение выявлено у 300, а из 600 женщин это нарушение зафиксировано у 300. Есть ли разница между мужчинами и женщинами по доле лиц, имеющих рассматриваемое нарушение?

В рассматриваемом примере нужные для расчетов величины таковы: $n_1 = 500$, $p_1^* = 300/500 = 0,6$; $n_2 = 600$, $p_2^* = 300/600 = 0,5$. Вычислим статистику

$$Q = \frac{0,6 - 0,5}{\sqrt{\frac{0,6 \cdot 0,4}{500} + \frac{0,5 \cdot 0,5}{600}}} = 3,34.$$

Поскольку $|Q| = 3,34 > 1,96$, то необходимо отклонить нулевую гипотезу и принять альтернативную. Таким образом, мужчины и женщины отличаются по доле лиц, имеющих рассматриваемое нарушение.

Итак, если частоты (выборочные доли) в группах достаточно сильно различаются, то делается вывод о различии долей в генеральных совокупностях. А если сравнительно мало различаются, то утверждают,

что различие не обнаружено. А где граница? Допустимое расхождение $\Delta = \Delta(\alpha)$ между частотами нетрудно получить расчетным путем (см. [3, гл. 1]). Для данных примера 3 $\Delta = \Delta(\alpha) = 1,96 \times 0,0299 = 0,059$, или 5,9%, для уровня значимости 0,05.

Пример 4. Фраксипарин получали 75 пациентов, из которых 53 мужчины (70,7%), фраксифорте - 75 пациентов, из которых 45 мужчин (60%). Различаются ли подгруппы по половому составу?

Данные, требуемые для расчетов: $n_1 = 75$, $p_1^* = 53/75 = 0,707$; $n_2 = 75$, $p_2^* = 45/75 = 0,600$. Вычислим статистику критерия проверки однородности (обнаружения эффекта)

$$Q = \frac{0,707 - 0,600}{\sqrt{\frac{0,707 \cdot 0,293}{75} + \frac{0,6 \cdot 0,4}{75}}} = 1,386.$$

Поскольку $|Q| = 1,386 < 1,96$, необходимо принять нулевую гипотезу и отклонить альтернативную. Различия не обнаружено, следовательно, подгруппы не отличаются по половому составу.

16. Случай малых вероятностей

Рассмотренные выше методы анализа медицинских данных (доверительного оценивания доли (вероятности) и проверки однородности вероятностей) основаны на предположении о возможности использования биномиальной модели выборки. Кроме этого, необходимо, чтобы объем выборки n был достаточно большим - несколько десятков или сотен.

Наконец, кроме биномиальной модели и большого объема выборки, необходимо выполнение третьего условия: число m ответов «есть» (в схеме с двумя ответами «есть» - «нет») также было достаточно большим, сравнимым с объемом выборки. Последнее эквивалентно предположению, что np - произведение числа опытов на вероятность рассматриваемого ответа при отдельном опыте - также достаточно велико (np - среднее

значение (математическое ожидание) числа ответов m , т.е. $M(m) = np$. Третье условие означает, что вероятность p не является слишком малой (не меньше 5-10%). Не должна быть слишком малой и дополнительная вероятность $1 - p$. Следовательно, вероятность ответа «есть» должна лежать между 5 и 95% (а лучше - между 10 и 90%).

Однако последнее условие выполнено отнюдь не всегда. Например, в области с миллионным населением наблюдаются сотни или тысячи заболеваний. Хотя их число достаточно велико, но оно составляет лишь доли процента от численности населения. Для статистического моделирования в таких случаях используют не биномиальное распределение, а распределение Пуассона.

Случайная величина Y имеет распределение Пуассона, если

$$P(Y = y) = \frac{\lambda^y e^{-\lambda}}{y!}, \quad y = 0, 1, 2, \dots,$$

где λ – параметр распределения Пуассона (интенсивность заболеваемости или иной характеристики), $P(Y=y)=0$ для всех прочих чисел y .

Для распределения Пуассона математическое ожидание и дисперсия совпадают с интенсивностью:

$$M(Y) = \lambda; D(Y) = \lambda. \quad (1)$$

Это распределение названо в честь французского математика С.Д. Пуассона (1781 - 1840), впервые получившего его в 1837 г. Распределение Пуассона - предельный случай биномиального распределения, когда вероятность p осуществления события мала, а число испытаний n велико, причем $np = \lambda$. Таким образом, справедливо предельное соотношение

$$\lim_{n \rightarrow \infty, np \rightarrow \lambda} P(Y = y | p, n) = \frac{\lambda^y e^{-\lambda}}{y!}, \quad y = 0, 1, 2, \dots$$

Поэтому распределение Пуассона (в старой терминологии «закон распределения») часто называют также «законом редких событий».

Состоятельной оценкой неизвестного параметра λ распределения

Пуассона является наблюдаемое значение случайной величины Y . Доверительное оценивание может быть основано на асимптотической нормальности распределения Пуассона и соотношении (1). Для $\gamma = 0,95$ нижняя доверительная граница имеет вид

$$\lambda_{\text{нижн}} = Y - 1,96 \sqrt{Y},$$

а верхняя доверительная граница

$$\lambda_{\text{верх}} = Y + 1,96 \sqrt{Y}.$$

Эти границы являются приближенными (асимптотическими), поскольку получены в предположении, что интенсивность λ достаточно велика (несколько десятков или сотен). Более точные расчеты можно провести по правилам, приведенным в специальной литературе [10].

Пример 5. Пусть в муниципальном образовании численностью 125 тыс. чел. наблюдается 438 случаев заболевания. Необходимо оценить заболеваемость (на 10 000 жителей).

Заболеваемость составляет

$$\frac{438}{125000} \times 10\,000 = 35,04$$

случаев на 10 000 жителей. Чтобы рассчитать доверительные границы, найдем границы для общей заболеваемости (для всего муниципального образования):

$$\lambda_{\text{нижн}} = 438 - 1,96 \sqrt{438} = 438 - 41 = 397, \lambda_{\text{верх}} = 438 + 41 = 479.$$

Нижняя доверительная граница для заболеваемости на 10 000 жителей составляет

$$\lambda_{\text{нижн}} = \frac{397}{125000} \times 10000 = 31,76,$$

а верхняя доверительная граница для заболеваемости на 10000 жителей:

$$\lambda_{\text{верх}} = \frac{479}{125000} \times 10000 = 38,32.$$

Есть ли отличия между двумя муниципальными образованиями, близкими по численности, или заболеваемости на одной и той же

территории, но в соседние годы? Следует рассматривать случайные величины Y_1 и Y_2 , имеющие распределения Пуассона с параметрами λ_1 и λ_2 соответственно, и проверять статистическую гипотезу о равенстве параметров этих распределений, т.е. гипотезу однородности $H_0: \lambda_1 = \lambda_2$, при альтернативной гипотезе о наличии эффекта $H_1: \lambda_1 \neq \lambda_2$. (Иногда представляют интерес и иные альтернативные гипотезы, прежде всего односторонние альтернативные гипотезы $H_1': \lambda_1 > \lambda_2$ и $H_1'': \lambda_1 < \lambda_2$.)

В прикладной математической статистике разработаны правила проверки гипотезы однородности для распределений Пуассона [10, п. 5.5]. При достаточно больших (не менее нескольких десятков) наблюдаемых значениях случайных величин Y_1 и Y_2 достаточно воспользоваться статистическим критерием, основанным на асимптотической нормальности случайных величин с распределениями Пуассона. В этом случае правило принятия решения при проверке однородности двух выборок формулируется следующим образом.

1. Вычислить статистику

$$R = \frac{|Y_1 - Y_2|}{\sqrt{Y_1 + Y_2}}.$$

2. Сравнить значение статистики R с граничным значением K . Если $R \leq K$, то принимается гипотеза однородности H_0 . Если же $R > K$, то однородность отсутствует и принимается альтернативная гипотеза H_1 о наличии эффекта.

Граничное значение K определяется выбором уровня значимости α статистического критерия проверки однородности. В клинко-статистических и медико-биологических исследованиях обычно $\alpha = 0,05$, а $K = 1,96$.

Примечание. Если численности сравниваемых популяций сильно отличаются, то необходимо применять более сложные способы расчета, описанные в статистической литературе [10, п. 5.5]. При этом исходная

гипотеза однородности состоит в равенстве заболеваемости на 10000 жителей. В нулевой гипотезе утверждается, что отношение параметров двух распределений Пуассона равно известной величине (отношению численностей популяций).

Пример 6. Пусть в 2009 г. на некоторой территории было 235 случаев заболевания, а в 2010 г. - 312 случаев. Можно ли объяснить это различие случайными причинами?

Примем, что за год численность населения на этой территории изменилась незначительно. Поскольку численность измеряется миллионами человек, то число заболеваний в определенном году моделируется распределением Пуассона. Рассчитаем значение статистики критерия проверки однородности двух распределений Пуассона:

$$R = \frac{|235 - 312|}{\sqrt{235 + 312}} = 3,29.$$

При $\alpha = 0,05$ $K = 1,96$ $R = 3,29 > K = 1,96$, поэтому гипотезу однородности отклоняем. При $\alpha = 0,01$ $K = 2,58$ $R = 3,29 > K = 2,58$. Можно сделать вывод о значительном увеличении заболеваемости.

Пример 7. Пусть выполнены все предположения примера 6, кроме одного - заболеваемость в 2010 г. составила 270 случаев. Изменяются ли сформулированные выводы?

Рассчитаем значение статистики критерия

$$R = \frac{|235 - 270|}{\sqrt{235 + 270}} = 1,56.$$

При $\alpha = 0,05$ $K = 1,96$ $R = 1,56 < K = 1,96$, следовательно, гипотезу однородности принимаем. Выводы примера 6 не подтверждаются. Различие между годами можно объяснить случайными причинами.

Расскажем о прикладном исследовании, в котором использовался метод проверки гипотезы о равенстве параметров двух распределений Пуассона.

Описторхоз - тяжелое глистное заболевание, склонное к хроническому течению, вызывающее нарушения различных функций организма [11, 12]. Один из путей заражения - употребление в пищу сырой рыбы, например, в замороженном (строганина) или вяленом виде, поэтому описторхоз распространен в северных районах, но встречается практически во всех субъектах РФ. Он приводит к серьезным осложнениям вплоть до развития первичного рака печени. Например, в одной из северных областей, где широко распространено заболевание описторхозом, количество заболеваний раком печени в 7 раз превышает средний показатель по стране. За год в России регистрируется около 100 тыс. лиц, заболевших описторхозом.

Анализировались статистические данные о заболеваемости описторхозом по 61 административной территории. Одна из основных задач исследования - отличить территории с наличием местных очагов заражения от территорий, где имеются лишь «завозные» случаи описторхоза. (Если пациент заболел, находясь на одной территории, а лечиться стал после переезда на другую, то наблюдаем «завозной» случай.) Другая задача - прогноз (с помощью доверительных границ) числа заболеваний на будущий год. Третья - типология (классификация) территорий по степени зараженности описторхозом.

Для решения поставленных задач использовались статистические модели и методы доказательной медицины, основанные на доверительном оценивании и проверке гипотез относительно параметров пуассоновских распределений. Число заболеваний на конкретной административной территории моделировалось распределением Пуассона. Изменение числа заболеваний с течением времени может быть связано с появлением очага заражения на рассматриваемой территории, но может объясняться и случайными отклонениями. Первый случай соответствует изменению параметра распределения Пуассона, а второй - его постоянству. Поэтому

обосновать вывод о появлении очага заражения можно, отклонив статистическую гипотезу о равенстве параметров распределений Пуассона, описывающих число заболевших в соседние годы.

Полученные результаты представляли достаточно большой интерес для практической медицины. В частности, был сделан вывод о наличии местной заболеваемости в двух конкретных областях, после чего были обнаружены очаги заболевания описторхозом в этих областях.

Итак, сравнение групп проводят с помощью статистических методов проверки однородности. В научных медицинских исследованиях до сих пор часто применяют критерий Стьюдента для проверки однородности двух выборок. Однако он опирается на два предположения - о нормальном распределении результатов наблюдений и о равенстве дисперсий для двух групп. Ни то, ни другое предположение обычно не выполнено в научных медицинских исследованиях. Поэтому современная прикладная статистика рекомендует вместо критерия Стьюдента применять иные критерии [13, 14], а именно, критерий Крамера-Уэлча [15], непараметрические критерии Вилкоксона [16, 17], Смирнова, Лемана-Розенблатта [18].

Литература

1. Голдстейн М., Голдстейн И. Как мы познаем. Исследование процесса научного познания. М.: Знание, 1984. 256 с.
2. Новиков Д.А., Новочадов В.В. Статистические методы в медико-биологическом эксперименте (типовые случаи). Волгоград: Изд-во ВолГМУ, 2005. 87 с.
3. Орлов А.И. Организационно-экономическое моделирование : учебник : в 3 ч. Ч.3. Статистические методы анализа данных. - М.: Изд-во МГТУ им. Н.Э. Баумана, 2012. - 624 с.
4. Флетчер Р., Флетчер С., Вагнер Э. Клиническая эпидемиология: основы доказательной медицины. М.: Медиа Сфера, 1998. 352 с.
5. Орлов А.И. Надо ли разбивать выборку? // Заводская лаборатория. Диагностика материалов. 1997. Т.63. №1. С. 54-54.
6. Орлов А.И. Прикладная статистика. М.: Экзамен, 2006. 671 с.
7. Орлов А.И. О высоких статистических технологиях // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. 2015. № 105. С. 14-38.
8. Орлов А.И., Федосеев В.Н. Менеджмент в техносфере. Учебное пособие для вузов. М.: Академия, 2003. 384 с.

9. Орлов А.И. Взаимосвязь предельных теорем и метода Монте-Карло // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. 2015. № 114. С. 27–41.
10. Большев Л.Н., Смирнов Н.В. Таблицы математической статистики. М.: Наука, 1983. 416 с.
11. Беэр С.А.. Биология возбудителя описторхоза. М.: Товарищество научных изданий КМК, 2005. 336 с.
12. Бочарова. Т.А. Возбудитель описторхоза и другие мышечные паразиты карповых рыб бассейна Нижней Томи. Томск: Изд-во Томского государственного университета, 2007. 66 с.
13. Орлов А.И. О применении статистических методов в медико-биологических исследованиях // Вестник Академии медицинских наук СССР. 1987. №2. С.88-94
14. Орлов А.И. О проверке однородности двух независимых выборок // Заводская лаборатория. Диагностика материалов. 2003. Т.69. №1. С.55-60.
15. Орлов А.И. Проверка статистической гипотезы однородности математических ожиданий двух независимых выборок: критерий Крамера-Уэлча вместо критерия Стьюдента // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. 2015. № 110. С. 197–218.
16. Орлов А.И. Какие гипотезы можно проверять с помощью двухвыборочного критерия Вилкоксона? // Заводская лаборатория. Диагностика материалов. 1999. Т.65. №1. С.51-55.
17. Орлов А.И. Двухвыборочный критерий Вилкоксона – анализ двух мифов // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. 2014. № 104. С. 91 – 111.
18. Орлов А.И. Состоятельные критерии проверки абсолютной однородности независимых выборок // Заводская лаборатория. Диагностика материалов. 2012. Т.78. №11. С.66-70.

References

1. Goldstejn M., Goldstejn I. Kak my poznaem. Issledovanie processa nauchnogo poznaniya. M.: Znanie, 1984. 256 s.
2. Novikov D.A., Novochadov V.V. Statisticheskie metody v mediko-biologicheskom jeksperimente (tipovye sluchai). Volgograd: Izd-vo VolGMU, 2005. 87 s.
3. Orlov A.I. Organizacionno-jekonomicheskoe modelirovanie : uchebnik : v 3 ch. Ch.3. Statisticheskie metody analiza dannyh. - M.: Izd-vo MGTU im. N.Je. Baumana, 2012. - 624 s.
4. Fletcher R., Fletcher S., Vagner Je. Klinicheskaja jepidemiologija: osnovy dokazatel'noj mediciny. M.: Media Sfera, 1998. 352 s.
5. Orlov A.I. Nado li razbivat' vyborku? // Zavodskaja laboratorija. Diagnostika materialov. 1997. Т.63. №1. S. 54-54.
6. Orlov A.I. Prikladnaja statistika. M.: Jekzamen, 2006. 671 s.
7. Orlov A.I. O vysokih statisticheskikh tehnologijah // Politematicheskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta. 2015. № 105. S. 14–38.
8. Orlov A.I., Fedoseev V.N. Menedzhment v tehnosfere. Uchebnoe posobie dlja vuzov. M.: Akademija, 2003. 384 s.
9. Orlov A.I. Vzaimosvjaz' predel'nyh teorem i metoda Monte-Karlo //

Politematicheskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta. 2015. № 114. S. 27–41.

10. Bol'shev L.N., Smirnov N.V. Tablicy matematicheskoy statistiki. M.: Nauka, 1983. 416 s.
11. Bejer S.A.. Biologija vozбудitelja opistorhoza. M.: Tovarishhestvo nauchnyh izdaniy KMK, 2005. 336 s.
12. Bocharova. T.A. Vozбудitel' opistorhoza i drugie myshechnye parazity karpovyh ryb bassejna Nizhnej Tomi. Tomsk: Izd-vo Tomskogo gosudarstvennogo universiteta, 2007. 66 s.
13. Orlov A.I. O primenenii statisticheskikh metodov v mediko-biologicheskikh issledovanijah // Vestnik Akademii medicinskih nauk SSSR. 1987. №2. S.88-94
14. Orlov A.I. O proverke odnorodnosti dvuh nezavisimyh vyborok // Zavodskaja laboratorija. Diagnostika materialov. 2003. T.69. №1. S.55-60.
15. Orlov A.I. Proverka statisticheskoy gipotezy odnorodnosti matematicheskikh ozhidaniy dvuh nezavisimyh vyborok: kriterij Kramera-Ujelcha vmesto kriterija St'judenta // Politematicheskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta. 2015. № 110. S. 197–218.
16. Orlov A.I. Kakie gipotezy mozhno proverjat' s pomoshh'ju dvuhvyborochnogo kriterija Vilkoksona? // Zavodskaja laboratorija. Diagnostika materialov. 1999. T.65. №1. S.51-55.
17. Orlov A.I. Dvuhvyborochnyj kriterij Vilkoksona – analiz dvuh mifov // Politematicheskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta. 2014. № 104. S. 91 – 111.
18. Orlov A.I. Sostojatel'nye kriterii proverki absoljutnoj odnorodnosti nezavisimyh vyborok // Zavodskaja laboratorija. Diagnostika materialov. 2012. T.78. №11. S.66-70.