

Л.А.Атраментова

ДИЗАЙН И СТАТИСТИКА
биологического исследования

Издательство «НТМТ»

Харьков

2014

УДК 57.087.1

ББК 28в6

A92

ISBN 978-617-578-198-2

Атраментова Л.А.

A92 Дизайн и статистика биологического исследования. Х. Издательство «НТМТ», 2014. – 255 с.

Книга адресована начинающим исследователям: студентам-биологам классических университетов, учащимся медицинских, сельскохозяйственных, фармацевтических, педагогических высших учебных заведений. Пособие может быть полезным также для аспирантов и научных работников в смежных с биологией отраслях науки. В книге излагаются принципы научно-исследовательской работы, описываются наиболее распространённые схемы проведения экспериментов, приёмы анализа фактического материала, приведены основы элементарной статистики, а также методы, нашедшие применение в прикладной науке, способы представления научных результатов. В книге рассмотрены наиболее распространённые ошибки, которые встречаются в научных отчётах и публикациях, даются рекомендации по их устранению.

ISBN 978-617-578-198-2

УДК 57.087.1

ББК 28в6

A92

©Атраментова ЛА, 2014


Издательство «НТМТ», 2014.

СОДЕРЖАНИЕ

Вступление	5	
Благодарности	9	
ПРИНЦИПЫ НАУЧНОГО ИССЛЕДОВАНИЯ		10
Фундаментальная и прикладная наука	10	
Наблюдение и эксперимент	11	
Ретроспективное и проспективное исследование		11
Сплошное и выборочное исследование		12
Поисковые и пилотные исследования		13
Научная проблема, цель и задачи исследования		14
Объект исследования	15	
Исследуемые совокупности	16	
ПЛАНИРОВАНИЕ, ОРГАНИЗАЦИЯ, ПРОВЕДЕНИЕ ИССЛЕДОВАНИЯ	20	
Виды научных задач	21	
Схемы исследования	24	
Субъектные исследования	29	
Уравновешивающее исследование		35
Мощность исследования	41	
Расчёт необходимого числа наблюдений		46
Условия исследования	60	
Научная документация	69	
ПОДГОТОВКА ДАННЫХ К СТАТИСТИЧЕСКОМУ АНАЛИЗУ		73
Обработка первичного материала	73	
Анализ распределения	74	
Приспособление данных к виду статистики		77

СТАТИСТИЧЕСКИЙ АНАЛИЗ ДАННЫХ	89
Статистические приёмы	89
Виды статистики	94
Статистические гипотезы и их проверка	97
РАБОТА С РАЗНЫМИ ВИДАМИ ДАННЫХ	120
Разнообразие признаков и дат	120
Статистические шкалы	123
Анализ категориальных данных (абсолютные числа)	126
Анализ долей	146
Анализ количественных данных	152
Свободно распределяющиеся данные	168
Выживаемость	175
Многомерный анализ	180
ПРЕДСТАВЛЕНИЕ РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЯ	193
Итоги статистического анализа	193
Научный отчёт	195
Ошибки в представлении результатов исследования	211
РЕКОМЕНДОВАННАЯ ЛИТЕРАТУРА	228
ИСПОЛЬЗОВАННАЯ ЛИТЕРАТУРА	228
ПРИЛОЖЕНИЯ	231

ВСТУПЛЕНИЕ



Написать эту книгу меня побудили две причины. Первая обусловлена профессиональной обязанностью – преподавание курса «Методология и организация научных исследований» студентам-биологам классического университета. Другим, пожалуй, более веским мотивом стали многочисленные обращения соискателей учёных степеней по биологическим и смежным специальностям примерно с такой речью: «Мы провели большое исследование, опубликовали достаточное количество статей, теперь можно подавать диссертацию. Помогите сделать правильную статистику». Объясняя это, претендент на учёную степень обычно держит в руках книгу по статистике, но скоро обнаруживается, что, как говорят в одном весёлом городе, книжная статистика отдельно, а статистика диссертанта отдельно. У меня всегда возникал риторический вопрос, который, я как-то ни разу не решилась озвучить: «Как же вам удалось опубликовать статьи без правильной статистики?».

Просмотр таблиц, заготовленных для диссертации, обычно приводил к заключению, что соискатель учёной степени имеет смутное представление о профессиональной научной работе. Большинство вообще убеждены, что для успешной научной работы нужно овладеть теми или иными специальными методиками, с их помощью собрать фактический материал, а уже напоследок, как они говорят, «статистически обработать результаты». Согласившись принять участие в доведении научного труда до требуемой кондиции, пытаюсь получить нужную для этого информацию. Обнаруживается, что любитель учёных степеней ни на один вопрос, касающийся организации и проведения исследования, ответить не может, так как просто не понимает, о чём идёт речь. Получается, что он работал не как исследователь (научный работник), а как исполнитель методик - лаборант (не в смысле должности, а в смысле понимания происходящего). Не ответив на вопросы о таких моментах

собственной работы, как дизайн исследования, слепой метод, мощность критерия, выпад, тип проверяемой статистической гипотезы, расстроенный соискатель, в отчаянии восклицает: «Почему нам раньше никто об этом не говорил!». Эта книга для тех, кто хочет знать то, о чём «раньше никто не говорил».

В настоящее время уже никто не оспаривает, что статистика – необходимая составляющая биологии и смежных с ней наук. Статистический анализ даёт возможность выявить связь, обнаружить влияние. Статистика – это не только вычисления. Статистика помогает учёным разрабатывать схемы исследования так, чтобы при минимуме затрат получить максимум информации, принимать решения и делать прогнозы. Научный работник, знающий статистику, может оценить результаты исследований своих коллег, сделать заключение относительно того, насколько обоснованы выводы других авторов. Не обладая статистической грамотностью, невозможно понять научную статью и оценить её уровень. Читатель, осведомлённый в статистике, критически относится к фактам, предлагаемым средствами массовой информации, и не принимает на веру всё, что публикуется.

К сожалению, не только в массовом сознании, но и среди некоторой части научных работников, отношение к статистике несколько скептическое, что отражают известные поговорки, повторение которых считаю дурным тоном, потому не привожу их здесь, дабы не тиражировать пошлость. Объяснить такое отношение к статистике можно тем, что её нередко используют в неблагоприятных целях. Другая причина – слабое знание самого предмета.

Отношения между биологией и статистикой складывались непросто. И этому есть основания. Одно из них – субъективное, заключается в том, что на биологические факультеты, в медицинские, зооветеринарные, сельскохозяйственные, фармацевтические вузы по преимуществу поступают молодые люди, не имеющие склонности к точным наукам. В их представлении биология, психология, медицина – науки описательные. Раньше они, действительно, таковыми и были, но в настоящее время это уже науки точные. Учащихся медицинских, зооветеринарных, агрономических, фармацевтических

вузов готовят к практической деятельности: врачей и ветеринаров – ставить диагнозы и лечить, зоотехников и агрономов - обеспечивать сельскохозяйственную продуктивность, фармацевтов - разрабатывать лекарства и т.д. Выполнять научные исследования специально никто не учит – нет такой дисциплины. Выпускники перечисленных вузов смутно представляют, чем отличается научная работа от их повседневных занятий. Потому нередко их статьи выглядят не как научные труды, а как отчёты о работе за определённый период. Есть и другая причина сложных отношений между биологией и статистикой. В середине прошлого века в СССР статистику вместе с генетикой власти объявили лженаукой и изгнали из университетов. Отечественные учёные десятки лет были изолированы от прогресса в этой области, а судьба многих была просто трагичной. Когда режим смягчился, и с репрессированных наук были сняты позорные клейма, обнаружилась пропасть, которую нечем было заполнить. Прерванные традиции, как известно, начальственными директивами восстановить не удаётся.

Массовая компьютеризация породила иллюзию быстрого решения проблемы. Научные работники установили на компьютерах статистические программы пиратского происхождения. Процесс пошёл. Книжный рынок отреагировал. Спасибо. Но это не решило проблемы. Большинство современных пособий по статистике для биологов-врачей-фармацевтов в значительной своей части являются набором инструкций по использованию компьютерных программ. И в этом их несомненная польза. Между тем, теоретические основы, написанные профессиональным языком математиков, как правило, плохо понимают, а чаще и вообще не понимают биологи, потому и не могут соотнести свою научную деятельность с понятийным аппаратом статистики. У биолога (врача, психолога и др.) есть набор данных, к которым, как он слышал, надо применить статистику, но какую именно статистику и что конкретно надо делать, он не знает. Что-то слышал о законе нормального распределения и читает: «Главная особенность нормального закона состоит в том, что он является предельным законом, к которому стремятся при

выполнении некоторых законов все другие законы распределения». Что должен делать биолог, прочитав такое? А как ему определиться с выбором критерия при таком пояснении: «в случае, когда H_1 сформулирована в виде $\theta \neq \theta_0$, используется двусторонний критерий, если же мы формулируем H_1 в виде $\theta > \theta_1$ (или $\theta < \theta_0$), то в этом случае используется односторонний критерий»? Сделав несколько безуспешных попыток разобраться в статистике, биолог в отчаянии забрасывает пособие, передаёт собранные фактические данные компьютерному гению, а получив результаты расчётов, не знает, что с ними делать, как, впрочем, и сам гений тоже.

Часто биологам рекомендуют осваивать азы статистического анализа сразу на компьютере. Эффективность такого методического приёма считаю сомнительной, будучи согласной с М.Е.Салтыковым-Щедриным, что «вводить просвещение надо постепенно, не допуская кровопролития». Применение статистических пакетов позволяет вычислять статистики мгновенно, но мешает понять, что происходит. Если исследователь надеется, что компьютер решит научную проблему, то его ждёт разочарование. Задавать задачи компьютеру должен учёный. Он сам, а не компьютер решает, какие показатели рассчитать и какой критерий выбрать. Базовую одномерную статистику биолог должен освоить вручную, с калькулятором. Её рецепты можно найти в пособиях по статистике для биологов, приведённых в «Списке рекомендованной литературы». До тех пор, пока исследователь сам не «переварит» данные, он не будет уверенно себя чувствовать при виде показателей, выдаваемых компьютером. Овладение вручную одномерной статистикой – это своего рода разминка, без которой, как известно, опасно приступать к серьёзным занятиям – можно надорваться.

В данном пособии я постаралась ответить на большинство вопросов, которые возникают у диссертантов при написании научного труда: как запланировать исследование, разработать дизайн, рассчитать нужное количество объектов, подобрать нужный метод статистического анализа, сделать выводы, написать научный отчёт, представить данные, оценить

научную статью, выявить фальсификацию. Хочу знать, что из этого
получилось, готова выслушать все замечания: 099-254-50-31,
atramentova@yandex.ru.