

2. МЕТОДЫ РАНДОМИЗАЦИИ

При постановке экспериментов, связанных с отбором групп животных (породоиспытание, опыты по скрещиванию, кормлению, зоогигиене, клинические испытания и т.п.), как правило, возникают вольные или невольные ошибки.

Предположим, что для сравнения эффективности использования двух молочных пород необходимо отобрать по 20 телок. Исследователь собрал данные по 100 телкам каждой породы с учетом возраста, живого веса, состояния здоровья и т.п. Как правильно выбрать из них 20 голов?

Если исследователь заранее убежден, что порода А лучше породы В, то при отборе животных он внесет (часто даже незаметно) преднамеренную систематическую ошибку - телки породы А в какой-то степени будут лучше в генетическом и других отношениях, чем телки породы В.

Если исследователь не внесет преднамеренную ошибку, то может допустить невольную ошибку - телки каждой породы будут представлять выборку лучших по фенотипу животных, но чаще не аналогов по племенной ценности. Это также исказит окончательные результаты опыта и может привести исследователя к неправильным выводам.

Другой пример, при формировании опытных групп непосредственно в животноводческих помещениях исследователь преднамеренно или непреднамеренно может отобрать более крупных или же более спокойных животных, находящихся в более (или менее) благоприятных условиях (освещения, влажности, температуры) одного или нескольких помещений.

Устранение такого рода ошибок возможно только путем *рандомизации* – случайного выбора животных в группы, случайной последовательности проведения опытов, измерений, оценок и т.п. В данном контексте термин «рандомизация» относится не к выборке, а к способу ее генерирования (формирования). Когда говорят, что группа животных численностью n является случайной выборкой из стада или популяции, то под этим подразумевают, что все возможные выборки такой численности извлекаются с равными вероятностями. Если говорят, что *обработка* (например, лечение

или дача кормовой добавки) назначается животным случайно, то под этим подразумевают, что вероятность назначения каждого вида обработки одинакова для всех животных.

2.1. Таблица случайных чисел

Простым и надежным методом достижения рандомизации является использование таблицы случайных чисел (табл. 1).

1. Случайная последовательность чисел, равномерно распределенных в интервале 0÷99999

| | | | | |
|--------------|--------------|-------|--------------|-------|
| 57705 | 13094 | 60835 | 36014 | 35950 |
| 71618 | 35193 | 42323 | 38612 | 08235 |
| 73710 | 64560 | 25732 | 93857 | 73606 |
| 70131 | 64559 | 98364 | 33749 | 66609 |
| 16961 | 68008 | 63407 | 08921 | 31842 |
| 53324 | 39848 | 72028 | 07721 | 22807 |
| 43165 | 33851 | 25496 | 58577 | 41476 |
| 26275 | 80586 | 83761 | 39303 | 74473 |
| 05926 | 69939 | 58568 | 19302 | 78489 |
| 66280 | 98351 | 27409 | 17068 | 14142 |
| 35483 | 12949 | 64789 | 59201 | 75975 |
| 09363 | 14644 | 78992 | 18688 | 55604 |
| 30304 | 32673 | 67388 | 73449 | 80702 |
| 55186 | 66887 | 75316 | 41734 | 11027 |
| 64003 | 43042 | 73673 | 17033 | 34559 |
| 20514 | 49110 | 21681 | 18664 | 73345 |
| 00188 | 18170 | 32763 | 94722 | 02783 |
| 55709 | 18187 | 50983 | 55024 | 54095 |
| 86977 | 02464 | 98359 | 85143 | 29373 |
| 31303 | 55739 | 38440 | 28564 | 96006 |
| 11578 | 52992 | 78142 | 76886 | 69351 |
| 39045 | 86513 | 25730 | 97570 | 07995 |
| 93011 | 10480 | 30454 | 26292 | 00900 |
| 42844 | 59437 | 19106 | 07120 | 29396 |
| 52906 | 13647 | 58222 | 11851 | 17727 |

Пример 2.1. Пусть имеется по 100 телок двух пород в одинаковой степени пригодных для опыта по породоиспытанию. Требуется рандомизированно отобрать по 20 телок каждой породы.

Для этого необходимо пронумеровать порядковыми номерами 100 телок первой и второй пород. Затем в табл. 1 произвольно выбрать точку начала отсчета. Например, седьмую строку первой колонки – число **43165**. От этого числа следует выписать подряд (или через одно, два, три и т.д.) две последние цифры (или первые, или

любые две) каждого числа: 65, 75, 26, 80, 83, 63, 4, 86, 3, 14, 88, 9, 77 (следующее число пропускается, так как № 3 был уже отобран), 78, 45, 11, 44, 6, 94, 93. Всего 20 чисел. После произвольного изменения точки отсчета, аналогичные операции повторяют для второй породы: из табл. 1 выписывают 20 номеров. Животные, соответствующие этим номерам, должны быть включены в опыт.

Пример 2.2. Допустим, что необходимо отобрать опытное поголовье непосредственно в свиноматке, причем из каждого станка одинаковое число животных (предположим - 4). В каждом станке не более 40 голов. Как следует поступить в этом случае?

По аналогии с примером 2.1, всем животным в каждом станке присваивают порядковые номера от 1 до 40. Затем из табл. 1 выписывают номера животных для первого станка. Пусть точка отсчета будет число **68008** в 5 строке 2 столбца. При выборе учитывают только первые две цифры каждого числа. Начиная с числа 68, просматривают последовательно все числа и отбирают те, которые не превышают 40. Для данного примера это будут числа 39, 33, (числа 80, 69, 98 пропускают, так как они больше 40), 12, 14. Таким образом, из первого станка следует взять для опыта животных под номерами 39, 33, 12, 14.

Пусть для второго станка точка отсчета будет число **58577** в 7 строке 4-го столбца. По аналогии с отбором номеров для первого станка, выписывают числа 39, 19, 17, 18. Аналогично по другим станкам.

2.2. Рандомизация в клинических испытаниях

Предположим, что необходимо провести клинические испытания лекарственного препарата, чтобы установить его эффективность. Для этого, например, 50 больным животным назначают лекарство, а другим 50 - нейтральный препарат («пустышку»). Предположим также, что животные поступают на испытания не одновременно, а группами, в течение некоторого времени.

Существует два метода рандомизации. В первом методе требуется выбрать 50 различных чисел между 1 и 100. Активное лекарство должно быть назначено тем из 100 больных животных, чьи номера попали в этот набор. Остальные 50 животных будут получать нейтральный препарат.

Этот метод имеет два недостатка.

Во-первых, если придется преждевременно завершить исследование, то общее число животных, принимавших активный препарат, с большой вероятностью не будет равно числу животных, принимавших нейтральный препарат. Между тем статистические методы сравнения теряют чувствительность, если размеры выборок различаются.

Во-вторых, если клиническое состояние животных, включающихся в испытание в один момент времени, отличается от состояния животных, включающихся в другой момент, или меняются правила приема препаратов, то, несмотря на рандомизацию, две группы, возможно, будут отличаться по типу животных или по правилам приема лекарств.

Второй метод рандомизации лишен недостатков, присущих первому. С помощью этого метода проводят независимую последовательную рандомизацию животных, поступающих в течение коротких промежутков времени, по группам лечения.

Предположим, что ежемесячно в испытаниях начинают участвовать десять больных животных. Разумно *случайно* назначать пяти животным лечение одного вида, а остальным пяти животным - другого, повторяя случайное назначение каждый месяц, по мере поступления новых партий больных животных.

Для каждой следующей группы больных животных следует получать новый набор случайных чисел, чтобы избежать смещений, которые могут появиться вследствие скрытой периодичности типа больных животных или ввиду того, что ветфельдшеру вскоре будет ясен вид лекарства (он не должен быть известен ветфельдшерам, контактирующим с животными).

Частный случай этого метода - испытания на парах животных, когда одно из двух животных получает активный, а другой - нейтральный препарат. В этом случае рандомизацию проводить очень просто.

Сначала каким-либо образом, например, по алфавитному порядку кличек, выделяют одного из двух больных животных как первого. Этот выбор надо сделать до проведения рандомизации. Затем, начиная с любого удобного места, просматривают однозначные числа в табл. 1. Если цифра нечетная - 1, 3, 5, 7 или 9, то первое больное животное принимает активный, а второе -

нейтральный препарат. Если цифра четная - 0, 2, 4, 6 или 8, активное лекарство назначают второму больному животному.

Описанные методы рандомизации приводят к назначению каждому животному одного из двух видов лечения с шансами 50 на 50. Все эти методы за исключением тех из них, в которых пары животных подбирают по признакам, взаимодействующим с изучаемым фактором, сопряжены с риском дисбаланса между группами больных животных (в которых проводят различное лечение) в распределении возраста, пола, начальной тяжести заболевания или других прогностически важных факторов. Кроме метода подбора пар по прогностическим факторам, можно использовать другое решение: провести *стратификацию*, т.е. разделить животных на определенные слои, группы (например, выделить самцов в возрасте от 6 до 12 мес., самок 6-12 мес., самцов 13-18 мес. и т.д.), а затем применить независимо и раздельно внутри каждой группы один из методов рандомизации.

2.3. Схема «несимметричной монеты»

Рандомизация с расслоением уменьшает, но не устраняет полностью риск дисбаланса, особенно если испытания проводят на животных, поступающих в разное время, и окончательное число животных в каждой группе неизвестно до конца набора животных, участвующих в испытаниях. Средством дальнейшего уменьшения возможности дисбаланса является концепция «несимметричной монеты».

Предположим, что поступающее животное относится к группе, в которой большему числу больных животных назначено лечение одного вида, а меньшему числу - лечение другого вида. Тогда, согласно схеме несимметричной монеты, новому больному животному с некоторой вероятностью $p > 0,5$ назначают лечение, которое получила на текущий момент меньшая часть больных животных. Вероятность назначения ему лечения, которое получила большая часть больных животных, будет равна $1 - p < 0,5$. При равном числе больных животных, получивших лечение того и другого вида, новому животному вид лечения назначают с вероятностью 0,5.

Продемонстрируем схему несимметричной монеты со значением $p=2/3$. Пусть животное относится к возрастной группе, в которой два вида лечения были назначены различному числу животных. Из табл. 1 извлечем одно из целых чисел от 1 до 9. Если оно делится на 3, (3, 6 или 9), то животному назначают лечение, полученное на этот момент бóльшим числом животных рассматриваемой группы, а если нет (1, 2, 4, 5, 7 или 8), то назначают лечение другого вида.

Вероятность назначения определенного лечения, отличная от $1/2$, требуется также в так называемых адаптивных клинических испытаниях. Здесь проводят лечение так, чтобы к концу испытаний бóльшее число животных прошло курс лечения, считающегося лучшим, а меньшее число - курс лечения, которое считается менее эффективным. Пусть задана вероятность $p > 0,5$ того, что животному назначают лечение, которое к этому моменту считается лучше другого (если ни одному из видов лечения не отдается предпочтение, то полагают $p=0,5$). В табл. 1 просматривают двузначные числа: если выбранное целое число лежит в интервале от единицы до $100 \times p$, то поступившему животному назначают лечение лучшего вида, если же число больше $100 \times p$ (нуль принимается за 100), то назначают лечение второго вида. Например, при $p=0,6$ любое число от 01 до 60 будет соответствовать назначению первого вида, от 61 до 00 второго.

В заключение следует отметить, что использование принципа рандомизации при формировании контрольной и опытных групп обеспечивает несмещенную (правильную) оценку статистических параметров и наименьшую ошибку эксперимента. Рандомизация является основой конкурсных испытаний животных и птицы различных линий и пород. В исследованиях по животноводству этот принцип должен соблюдаться тем строже, чем больше затраты времени и средств на эксперимент, и чем важнее выводы, вытекающие из него.

\hat{g}