А. А. Любищев
Дисперсионный анализ в биологии
Издательство Московского университета

УДК 578.087.1

Любищев А. А. Дисперсионный анализ в биологии. — М.: Изд-во Моск. ун-та, 1986.—200 с.

В монографии популярно изложены основы дисперсионного анализа Р. А. Фишера и логико-математические основы теории планирования и обработки результатов биологического эксперимента. Автор описывает конкретные математические методы обработки результатов полевого опыта и экспериментальных данных.
Для специалистов в области биологии, медицины и сельского хозяйства.

Ответственный редактор: кандидат биологических наук Б. С. *Шорников*

Рецензенты: доктор физико-математических наук *Ю. М. Свирижев*
доктор физико-математических наук *И. А. Ибрагимов*

Печатается по постановлению Редакционно-издательского совета Московского университета

МОНОГРАФИЯ
Александр Александрович Любищев
ДИСПЕРСИОННЫЙ АНАЛИЗ В БИОЛОГИИ

Зав. редакцией *Н. М. Глазкова.* Редактор *Н. Г. Комлева.* Переплет художника *Ю. И. Артюхова.* Художественный редактор *М. Ф. Евстафиева.* Технический редактор *Г. Д. Колоскова.* Корректоры *В. П. Кададинская, Т. С. Милякова*

ИБ № 2312

Сдано в набор 3.12.85. Подписано к печати 11.09.86. Л-67400 Формат 60Х90/16 Бумага тип. № 1 Гарнитура литературная. Высокая печать Усл. печ. л. 12,5 Уч.-изд. л 13,86 Тираж 3900 экз. Заказ № 259 Цена 1 р. 60 к. Изд. № 3618

Ордена «Знак Почета» издательство Московского университета 103009, Москва ул. Герцена, 5/7. Типография ордена «Знак Почета» изд-ва МГУ. 119899, Москва, Ленинские горы

АЛЕКСАНДР АЛЕКСАНДРОВИЧ ЛЮБИЩЕВ (1890–1972) И ЕГО БИОМЕТРИЧЕСКИЙ ТРУД

Руководство по применению дисперсионного анализа в биологии написано А. А. Любищевым в 1938–1940 гг., т. е. в возрасте около пятидесяти лет, и как бы завершает 25-летний период его научной, научно-педагогической и научно-организационной деятельности в области сравнительной и эволюционной морфологии, теоретической и сельскохозяйственной энтомологии, агробиологии и теории планирования, организации и анализа полевого агробиологического эксперимента, биометрии и биоматематики.

Александр Александрович Любищев родился 5 апреля 1890 г. в Петербурге в семье лесопромышленника. В 1906 г. по окончании реального училища поступил на естественно-историческое отделение физико-математического факультета Петербургского университета, который окончил в 1911 г.

Еще в студенческие годы при прохождении учебной зоологической практики на Неаполитанской биологической станции (1909) и на морской биологической станции Виллафранко (1910), а затем, по окончании университета, на Мурманской биостанции (лето 1912 г.) у А. А. Любищева возникла идея использовать математические (геометрические) методы в решении морфологических задач описания разнообразия органических форм морских червей и раковин аммонитов. В значительной степени она была навеяна чтением книги В. Скиапарелли «О сравнительном изучении естественной и чистой формы».

В последующей научной работе на Петербургских Бестужевских женских курсах (1914–1915) и во время службы в армии (1915–1918) А. А. Любищев не оставлял этих математических опытов и размышлений. В 1917–1918 гг. появились в его дневниках первые наброски сочинения «Линия Платона — Пифагора в истории культуры», частично реализованные в 1961 г. и определившие всю. последующую судьбу А. А. Любищева как биоматематика и биометрика. Именно древнегреческой школы пифагорейцев (идея «чистой» математики) и школе Платона (идея «чистого» разума) принадлежит первая системно-целостная попытка логико-математического истолкования системной целостности гармонии Природы и Вселенной: идея Порядка — Гармонии — Симметрии природы и мироздания, где Идея есть Число — Гармония — Образ (фигура).

В эти и последующие годы в творчестве. А. А. Любищева стал складываться тот естественнонаучный стиль, который позднее был назван биоэстетическим стилем Любищева [5, 5а, 11]. Развитие биоэстетического принципа нашло свою реализацию в его последующих естественнонаучных работах.

В 1918 г. начинается педагогическая деятельность А. А. Любищева. Он принимает предложение профессора А. Г. Гурвича и становится ассистентом одной из кафедр Крымского (Таврического) университета. Чтение лекций, участие в творческих дискуссиях по актуальным проблемам естествознания с выдающимися профессорами Таврического университета Н. М. Крыловым, О. В. Струве, В. А. Обручевым, В. И. Вернадским, В. И. Палладиным, Г. Ф. Морозовым, П. П. Сушкиным и др.; а затем, после перехода в Пермский университет на должность доцента (1921–1927), — с профессорами В. Н. Беклемешевым, А. А. Заварзиным, Б. Ф. Вериго, П. Г, Светловым, Ю. А. Орловым и многими другими выдающимися деятелями науки и культуры сформировали наиболее яркие критико-публицистические черты в творчестве А. А. Любищева [5].

Это нашло свое отражение и развитие в шести наиболее выдающихся естественнонаучных работах А. А. Любищева по теории эволюционизма, генетике и систематике организмов (1923–1931), сыгравших заметную роль в развитии отечественной биологии. Так, в работе «Критерии изменчивости организмов» (1923) впервые был использован дисперсионный критерий оценки индивидуальной изменчивости организма. В работе «О форме естественной системы организмов» он впервые предпринял попытку построения «периодической системы организмов», в которой эволюция представляла собой исторический процесс реализации разнообразия органических форм живого.

Особую гордость А. А. Любищева составляла его работа «О природе наследственных факторов» (1925), где впервые в мировой эмбриогенетической практике была дана геометрическая концепция пространственно-временной, следовой (память), локализации гена. Эта работа названа П. Г. Светловым [5] «пророческой» книгой.

Три другие работы А. А. Любищева — «Понятие эволюции и кризис эволюционизма» (1925), «Понятие номогенеза» (1928), «Логические основания современных направлений биологии» (1931) — отличала глубоко продуманная аргументированность и критичность мышления, оригинальность, самобытность и целостность анализа сложившейся научной ситуации, умение как предвидеть и оценить существующие тенденции развития науки, так и дать прогностическую оценку развития этих тенденций в будущем.

В одной из своих рецензий Б. С. Соколов [11] назвал такой стиль «научным критицизмом». Нам в соответствии с общепринятыми литературоведческими терминами [12] хотелось бы считать этот стиль критико-публицистическим. Ибо критика (или критичность) есть «искусство судить, оценивать и правдоподобно истолковывать достоинства и недостатки анализируемого произведения с позиций нравственных ценностей современной общественной жизни» [12]. Именно это умение критически мыслить — разобрать, сравнить и оценить исходные интеллектуальные предпосылки изучаемых научных концепций, гипотез и теорий; логико-эвристически восстановить утраченные звенья научной логики такой теории или концепции; дать критический анализ интерпретации полученных экспериментальных закономерностей — было присуще творчеству Любищева. Такой подход сам А. А. Любищев [5, 5а] называл «отделом теоретического контроля», где самым важным этапом является Этический кодекс Науки [5, 11], блистательным рыцарем которой он и был.

В 1927 г. А. А. Любищев был избран профессором кафедры зоологии Самарского сельскохозяйственного института, преобразованного затем в филиал ВИЗРа (Всесоюзного института защиты растений ВАСХНИЛ). В этот период (1927–1938) в творческом общении с выдающимися агробиологами С. М. Талайковым и К. Ю. Чеховичем, с экономистом Н. Н. Богдановым-Катковым завершилось агробиологическое образование А. А. Любищева в области теории планирования, организации и экономики сельскохозяйственного эксперимента и анализа данных.

Большую роль в этом сыграли творческие и дружеские контакты Любищева с американским биометриком, учеником Р. Фишера, Ч. Блиссом [20], которые начались в стенах ВИЗРа и продолжались в течение 35 лет. В 1950 г. Блисс рекомендовал А. А. Любищева в действительные члены Американского биометрического общества. По каналам журнала «Байометрикс», издаваемого биометрическим обществом (США) шло творческое общение Любищева с другими выдающимися биометриками: Р. А. Фишером [21], У. Г. Кохраном [22], Дж. Снедекором [23] и многими другими учеными США и Англии. Здесь же, на страницах журнала «Байометрикс», были опубликованы и некоторые статьи самого Любищева, в их числе «Об использовании дискриминантных функций в таксономии» (1962).

Руководство по применению дисперсионного анализа в биологии как бы подводит итог 25-летней довоенной деятельности А. А. Любищева. В этой работе проявились лучшие черты его творчества: критичность, биоэстетика, системность и правдоподобно-эвристические рассуждения, используемые для решения биометрических (статистических) задач, которые нашли адекватное отражение в виде числовых примеров, биометрических задач и вероятностно-статистических решений различных агробиологических проблем с использованием теории планирования и анализа полевого эксперимента.

Логико-семантические особенности «Руководства»

В своем предисловии к «Руководству...» [5] А. А. Любищев пишет, что основная цель создания этой книги — «сделать общедоступной для биологической практики методику по организации агробиологических исследований и обработке полученных результатов исследования методом дисперсионного анализа Р. Фишера», так как работы Р. Фишера [15, 21] не являются ни учебниками, ни руководствами; это оригинальное теоретическое введение в математическую теорию планирования и анализа эксперимента [5, 18].

Далее, добавляет А. А. Любищев, было бы ошибочно думать, «что овладение только самой техникой вычислений, без необходимого уяснения логической стороны метода Р. Фишера, будет достаточным. Только внимательное прочтение «Руководства...», проработка и решение приведенных числовых примеров дают реальную возможность творчески усвоить методологию дисперсионного метода Р. Фишера».

Иными словами, именно логико-семантической (смысловой) стороне методологии дисперсионного анализа Р. Фишера, разбору логико-семантических тонкостей дисперсионного анализа, адаптированного к задачам агробиологического (полевого) исследования, и посвящены многочисленные агробиологические примеры, используемые Любищевым в своем руководстве. Такому логико-семантическому разбору методологии Р. Фишера посвящено 26,4% объема рукописи.

Но наряду с методологией собственно дисперсионного анализа А. А. Любищев использует другие статистические методы, основанные на идее «дисперсии» («рассеяния» индивидуальных дат вокруг центра распределения), такие как факторный, итерационный и дискриминантный анализ [7, 7а], на долю которых приходится до 50% объема рукописи. Кроме того, он широко использует и недисперсионные статистические методы: корреляционный, регрессионный и комбинаторный анализ, а также графические методы анализа и оценки качественных результатов группировки данных. На долю графических и статистических методов приходится 23.6% объема «Руководства..»

Таким образом, перед нами не только руководство по методологии дисперсионного анализа Р. Фишера [5, 21], но и богато иллюстрированное многочисленными примерами собственное логико-семантическое элементарное введение в теорию планирования и анализа агробиологического (полевого, технологического [7, 7а]) Эксперимента. Об этом говорят и три принципа планирования эксперимента, сформулированные А. А. Любищевым, три обязательных требования, которым должен удовлетворять любой агробиологический эксперимент.

1. Планирование эксперимента должно быть *целеустремленным,* т. е. иметь перед собой определенную, подлежащую числовому решению биологическую задачу. Планы математического эксперимента должны естественным образом вытекать из задач биологических.

2. Оно должно быть *эффективным,* т. е. извлекать максимум .информативности выводов из наличествующего материала, отсеивать все случайные выводы, так чтобы полученные (обобщенные) выводы были надежны и обладали принудительной силой (вероятностной) меры надежности.

3. Исследование должно быть *экономным,* т. е. оно должно быть правильно (научно) организовано и осуществлено с минимальными затратами труда и средств.

Таким образом, в этих принципах Любищев изложил не математическую «теорию» планирования эксперимента Р. А. Фишера, а собственную системно-целостную, минимаксную и экономическую теорию биологического полевого эксперимента, что опередило труды зарубежных ученых [4, 14, 17] как минимум на пятнадцать лет. Именно в этом и состоит непреходящая логико-семантическая ценность труда Любищева.

Биометрическая структура «Руководства...»

С момента появления труда Фишера «Значимость эксперимента», 1935 [21] прошло полвека. За это время сложилась фишеровская школа математической теории планирования эксперимент та [15, 21]. Она опирается на математическую теорию эксперимента: теорию математической статистики и теорию вероятностей [7, 7а]. Она имеет свою методологию: теорию дисперсионного анализа Р. Фишера, включающую концепцию рандомизации (случайности) пространства признаков; дисперсионную теорию плана эксперимента; концепцию преобразования, редукции и оптимизации пространства переменных; теорию репрезентативности (представительности) выборочного эксперимента и статистической надежности полученных статистических выводов (обобщения) экспериментальных данных. Такова биометрическая «архитектура» (структура) математической теории планирования эксперимента Фишера (ТПЭ).

Эти особенности «архитектуры» фишеровской ТПЭ и нашли свое отражение в блок-схеме таблицы. Она подразделена на 4 структурных блока, включающих 28 основных логических параметров и характеризующих основные теоретические и методологические особенности ТПЭ.

Блок I — теория планирования эксперимента. В таблице представлена 6 параметрами.

Теория — это система руководящих идей (гипотез, концепций, законов) в познании наиболее общих (объективных) закономерностей природы и общества, дающая возможность абстрактного (математического) подхода к решению фундаментальных проблем и обнаружению функциональных связей между процессами и явлениями живой природы и общества [7, 9, 13].

Таблица

Сравнительный анализ труда А. А. Любищева, Р. Фишера и других современных работ по теории планирования эксперимента

Год	1940	1935	1970	1961	1976	1971	1972	1972	1983	1964	1970	1982
Автор	[5]	[15]	[14]	[10]	[4]	[7]	[6]	[17]	[1]	[16]	[8]	[19]
I. Основные параметры (математической) теории планирования эксперимента												
1. Логика	+					+						+
2. Эвристика (правдоподобие)	+		+			+					+	+
3. Системность	+					+						+
4. Экономика	+		+	+				+				
5. Статистика	+	+	+	+	+	+	+	+	+	+	+	+
6. Вероятность	+	+	+	+	+	+	+	+	+	+	+	+
II. Методология и основные процедуры дисперсионного анализа Р. Фишера [15]												
7. Теория выборки (репрезентативность)	+	+	+	+	+	+	+	+	+	+	+	+
8. Рандомизация	+	+	+	+	+	+	-	-	+	-	+	-
9. Планы-блоки	+	+	+	+	-	+	-	-	+	-	+	-
10. Преобразование переменных	+	+	+	+	+	+	-	-	+	-	+	-
11. Редукция набора переменных	+	+	+	+	+	+	-	+	-	-	+	+
12. Оптимизация пространства переменных	+	+	+	-	+	+	-	-	-	-	+	-
13. Адекватность статистического вывода	+	+	+	-	-	+	-	+	+	-	+	+
III. Основные методы статистического анализа набора данных												
14. Дисперсионный	+	+	+	+	+	+	+	+	+	+	+	+
15. Факторный	+	+	+	+	-	+	-	+	+	+	+	+
16. Последовательный (итерационный)	+	+	+	-	+	+	+	+	+	-	-	+
17. Ковариационный	+	+	+	+	+	+	-	-	+	-	+	-
18. Дискриминантный	+	+	+	+	-	+	-	-	-	-	+	+
19. Регрессионный	+	+	-	+	+	+	+	-	+	-	+	+
20. Корреляционный	+	+	-	+	-	+	+	-	+	+	+	+
21. Комбинаторный	+	+	+	+	+	+	-	-	-	+	+	+
22. Частотный	+	+	+	+	+	+	-	-	+	+	+	+
23. Графические	+	+	+	+	-	-	+	+	-	+	+	-
IV. Методы статистической оценки и обобщения данных. Принятие решения												
24. Интерполяция	+	+	+	-	-	-	-	+	-	-	+	+

25. Экстраполяция	+	+	-	-	-	-	-	+	-	-	+	+
26. Интерпретация	+	+	+	-	-	+	-	+	-	+	+	+
27. Моделирование	+	-	-	-	-	+	-	+	-	-	+	+
28. Прогнозирование	+	-	+	-	-	+	-	+	+	-	+	+
Итого (+) Полнота (доля в %)	28 100	22 78.6	22 78.6	17 60.7	13 46.4	24 85.7	8 28.6	15 53.6	15 53.6	10 35.7	24 85.7	21 75.0
Идентификация научной школы	Л	Ф	Ф	Ф	Др	Л	ОШ	Др	ОШ	ОШ	Ф	Л

Примечание. Сокращения в идентификации научных школ: Л – Любищев, Ф - Фишер, Др — другие статистические школы, ОШ — отечественные биометрические школы.

Современная ТПЭ, изучающая «плохо» организованные (биологические) системы [7, 7а], использует математический язык метаописания эксперимента, опирающийся на теорию вероятностей и математическую статистику, что является фундаментом фишеровской ТПЭ. Создавая свою теорию агробиологического эксперимента, Любищев дополнил фишеровскую математическую методологию ТПЭ системными, логико-эвристическими (правдоподобными [9]) и эконометрическими принципами организации и планирования агробиологического эксперимента [18].

Сравнивая шесть параметров блока I, характеризующих теоретические основания в ТПЭ, мы видим резкое различие творческих стилей обоих биометриков: формализованные, имитационно-математические основания дисперсионной теории Фишера [15, 21] и логико-семантические, эвристические основания теории планирования и организации агробиологического эксперимента Любищева. Фишеровская методология ТПЭ имитирует замену биологических закономерностей эксперимента аналоговой имитационной числовой моделью дисперсионного анализа Фишера (рандомизация выборки, редукция и оптимизация пространства переменных, основывающихся на репрезентативности выборочного исследования). Любищевский труд учитывает методологические особенности дисперсионной модели эксперимента Фишера, но дополняет их знанием логико-семантических, эвристических особенностей изучаемого биологического эксперимента, адекватно отражающих биологические закономерности анализа и обобщения биологических данных. Именно предметно-ориентированный язык «Руководства...» и делает труд Любищева доступным широкой биологической аудитории.

В *блоке II* таблицы представлены семь стандартных процедур методологии дисперсионного метода Р. А. Фишера, предложенных им в 1935 г. Она опирается на концепцию репрезентативности малой выборки, ее надежности в выборочном эксперименте [7, 7а] и концепцию рандомизации, т. е. принятия статистического решения в условиях неопределенности. Иными словами, создания такой ситуации в отборе и размещения проб и объектов агробиологического эксперимента, когда влияние факторов, мешающих эксперименту, нивелируется, т. е. делается случайным (рандомизированным), и тогда полученные случайные величины, характеризующие такое влияние (дисперсию), будут иметь нормальное статистическое распределение с дисперсией, равной 0.

На идее «учета» рассеивания (дисперсии) случайной величины и ее «отклонения» от полученной средней и построена вся методология планов дисперсионного анализа ТПЭ Фишера, т. е. схема равномерного, геометрически правильного размещения объектов на делянке, использующая ортогональные свойства латинских, греко-латинских квадратов (2х2, 3Х3 или $2^k x 2^k$), гиперкубов и кубов высших порядков [7, 7а]. Эти представления стали основой математической блок-схемы дисперсионного анализа в любой математической ТПЭ [14, 10, 4]. Именно исходя из концепции ортогональности (слагаемости) пространства переменных в таких экспериментах возможны преобразования пространства переменных: его сжатие и уменьшение до оптимального набора [7, 7а] признаков.

На основе сравнения, анализа и обобщения пространства переменных с учетом его редукции и репрезентативности делается статистический вывод о надежности эксперимента и соответствии полученных результатов исследования ранее сформулированной H_0-гипотезе, ее принятии или отклонении. Таков стандартный набор статистических процедур, используемых в методологии дисперсионного метода Фишера.

Сравнив семь параметров блока II, применяющих стандартную методологию дисперсионного анализа Р. Фишера, мы увидим, что по набору процедур дисперсионного анализа в трудах Фишера и Любищева используются тождественные методы дисперсионного анализа, предложенные Фишером. Близка к труду Фишера и работа Д. Финни [14]. В работах других статистиков наблюдается полный [6, 16, 2, 3, 9, 19] или частичный [4, 17, 1, 8] отход от стандартной процедуры Р. Фишера.

В современных трудах по теории планирования эксперимента [7, 7а] стандартная процедура дисперсионного анализа заменяется полностью или частично методологией регрессионного анализа, особенно в теории планирования технологического (промышленного) эксперимента [7, 7а, 17], а также при поиске экстремума и оптимума эксперимента (школа Налимова [7, 7а]: В. В. Налимов, Ю. П. Адлер, Е. В. Маркова, Ю. В. Грановский и их многочисленные ученики и последователи).

В *блоке III* таблицы представлены десять статистических методов анализа данных, которые можно сгруппировать в четыре различные группы методологий.

Первая группа методологий, основанных на идее оценки дисперсии (рассеивания) в теории планирования и анализа эксперимента, использует стандартную процедуру дисперсионного метода Р. Фишера и близкую к ней группу методологий факторного, последовательного, ковариационного и дискриминантного методов, разработанных Р. Фишером и модифицированных трудами Д. Финни [14], Г. Хоттелинга, Дж. Бокса, Г. Шеффе и многих других исследователей [7, 7а]. По этому параметру имеется полное сходство работ Фишера и Любищева, а также Финни [14] и Налимова [7, 7а]. Другие авторы используют эти идеи лишь частично [10, 4, 3, 6, 17, 1, 16, 8, 18, 19].

Вторая группа методологий основана на оценке степени «связности» и тесноты связи между объектами. Она представлена методами корреляционного и регрессионного (линейного и нелинейного) анализа, разработанного К. Пирсоном и его школой [7, 7а] и получившего развитие и дальнейшее широкое применение в современной теории планирования экстремального эксперимента, в том числе при поиске оптимальных условий технологического эксперимента [7, 7а, 17]. По этому параметру Фишер и Любищев сходны, они успешно применяли эти методы в своих работах по анализу эксперимента. Десять других авторов либо используют |10, 4, 7, 6, 1, 8, 19], либо не используют [14, 17, 16] эти методы в статистической практике.

Третья группа методологий, использующих идею ассоциативности (сочетательности) связи между независимыми признаками, представлена методами комбинаторного и частотного анализа. Это преимущественно статистическая оценка качественных признаков (типа «да-нет»), часто применяемых в биологическом эксперименте [7, 7а, 18]. При группировке таких признаков используются методы кластерного анализа [3] и частотно-дискриминантного анализа [8, 16, 18]. По группе этих методологий работы Фишера и Любищева не различаются. Другие статистические работы либо применяют [14, 10, 4, 7, 16, 8, 19], либо не применяют [1, 6, 17] такие качественные методы анализа данных.

Четвертая группа графических методов использует геометрические методы и методы функционального анализа в интерпретации полученных (обобщенных) данных с целью рекогносцировочной оценки предварительных выводов проведенного эксперимента. Эта оценка состоит в построении разнообразных графиков, диаграмм и графических блок-схем, иллюстрирующих пространственную и временную упорядоченность (иерархию) элементов эксперимента и функциональную связь между ними. Именно на этом этапе планирования и анализа ТПЭ находят свою реализацию эвристические и правдоподобные методы [6, 3, 18] анализа связей между признаками, их системная упорядоченность (топология) и .логико-семантическая структура [18], т. е. все то, что наиболее ярко иллюстрирует биоэстетические принципы в творчестве А. А. Любищева [5, 5а]. По степени использования изографических методов анализа данных можно выделить две группы исследователей: применяющих [5, 14, 15, 10, 6, 17, 16, 8] либо не применяющих [4, 7, 1, 19] такие методы.

В *блоке IV* таблицы представлены пять параметров, характеризующих теорию статистического вывода и принятия решения. Они включают методологию реконструирования пропущенных дат (координат) внутри полученных экспериментальных данных (интерполяция) и вне их (экстраполяция); методы логико-семантической интерпретации и обобщения полученных закономерностей эксперимента, т. е. придание математического (числового) смысла физическим параметрам эксперимента [7, 7а, 18]; методы логического, числового и машинного имитационного моделирования, а также прогнозирования полученных числовых закономерностей эксперимента, включающего оценку существующей ситуации и принятие решения с учетом трендов (тенденции) развития таких параметров в будущем [7].

Такова основная процедура принятия решения в условиях неопределенности (случайности) эксперимента [7, 7а], плохо поддающаяся алгоритмизации и машинизации в теория принятия решения [7, 7а, 9, 18, 19]. Поэтому в теории-планирования эксперимента Фишер заменил такую многофазную и последовательную процедуру однофазной концепцией, статистического вывода и альтернативного принятия решения без учета изменения анализируемой ситуации в будущем [7, 7а]. В таком вероятностно-статистическом контексте дисперсионная модель вероятностного вывода Фишера используется как дискриминационный метод (сходства — различия) оценки степени соответствия полученной фактически H_0-гипотезы ее известному теоретическому закону распределения: гауссовскому, биномиальному или любому другому распределению.

Такой же дисперсионной моделью статистического вывода Фишера пользуются восемь других авторов [5, 7, 8, 14, 15, 16, 18, 19] из 12, приведенных в таблице. Такая же дисперсионная модель используется для вероятностного прогнозирования полученных закономерностей в будущем [1, 8, 5, 7, 14, 17, 18, 19].

В отличие от методологии Фишера [15, 21], Любищев в своем «Руководстве...» не только применяет комбинаторику уже реализованных фактов (гипотез, концепций), но и учитывает логическую возможность гипотетических реализации других возможностей, не реализованных в данном эксперименте, т. е. использует понятие «потенциальной» вероятности (теория «шансов» Ог. Курно) [18].

Таким образом, по теории статистического вывода ТПЭ и его вероятностному обобщению (принятию решения) можно различать два подхода: дисперсионную вероятностную модель Фишера и комбинаторную вероятностную модель Любищева. Другие математические школы используют иные методы имитационного математического и вероятностного прогнозирования и моделирования [1, 8, 6, 16, 17], в том числе и логико-эвристического моделирования [5, 5а, 9], которое использует и А. А. Любищев в своем «Руководстве...»

Завершив системно-комплексное рассмотрение 28 параметров таблицы, мы подошли к современной трактовке теории планирования эксперимента, основополагающие идеи которой были заложены Фишером в 1935 г.

Здесь можно выделить (как минимум) три сложившиеся школы ТПЭ.

1. Фишеровская теория планирования эксперимента [15]. Включает основополагающие идеи дисперсионного анализа Фишера и его основные процедуры: концепцию рандомизации, редукции и оптимизации пространства признаков, репрезентативности выборочного исследования, надежности статистического (альтернативного) вывода.

К этой школе принадлежат сам Фишер [15, 21], Д. Финни [14], Дж. Снедекор [10, 23], У. Кокрен [4, 22] и Ч. Блисс [20J. Частично идеи дисперсионного метода Фишера нашли свое отражение и в отечественной биометрической литературе [8, 16].

2. Другие (нефишеровская школа) направления в ТПЭ используют идеи оптимальности пространства признаков и методологию корреляционного [13, 13а] и регрессионного анализа [4, 10, 6, 7, 1, 8, 19], а также принципы кибернетики [7, 18], динамического и линейного программирования [7, 7а], теории игр [7, 7а, 18, 19], и многие другие направления современной математики, информатики и семантики [7, 19].

3. Любищевская школа теории планирования агробиологического эксперимента основывается на фишеровской методологии рандомизации и редукции пространства переменных [15, 21], дополненной эвристическими, логико-комбинаторными и логико-семантическими принципами интерпретации полученных результатов эксперимента, основывающимися на идее «гетерогенности» полевого эксперимента [5].

Наряду с биологическими концепциями теории полевого опыта А. А. Любищев широко использует в «Руководстве...» и экономико-статистические (эргономические) критерии оценки эффективности борьбы с сельскохозяйственными вредителями [5, 5а], нашедшие свое развитие в трудах его учеников [5].

Труд Любищева — это редкостный синтез эвристических и логико-математических методов в теории биологического эксперимента, что представляет редчайшее исключение в общем потоке биометрических работ [8, 16], где основой является только числовой пример (рецептура) без раскрытия семантической значимости метода [14, 10, 4]. Именно в логико-семантических особенностях биологического эксперимента, раскрытых в «Руководстве...» Любищева, и состоит непреходящая новизна его труда.

Б. С. Шорников

ЛИТЕРАТУРА

1. Асатурян В. И. Теория планирования эксперимента. М., 1983, 248 с.
2. Джинни К. Средние величины. М., 1970, 417 с.
2а. Джинни К. Логика в статистике. М., 1973, 126 с.
3. Дюран Б., Оделл П. Кластерный анализ. М., 1977, 127 с.
4. Кокрен Ул. Методы выборочного исследования. М., 1976, 440 с.
5. Александр Александрович Любищев (библиографический очерк). М., 1983, 143 с.
5а. Любищев А. А. Проблемы формы, систематики и эволюции организмов (сборник избранных статей 1920–1970 гг.). М., 1982, 278 с.
6. Митропольский А. К. Техника статистических вычислений. М., 1972, 576 с.
7. Налимов В. В. Теория эксперимента. М., 1971, 207 с.
7а. Налимов В. В., Голикова Т. И. Логические основания планирования эксперимента. М., 1976, 128 с.
8. ПлохинскийН.А. Биометрия. Изд. 2-е. М., 1970, 367 с.
8а. Плохинский Н. А. Алгоритмы биометрии. Изд. 2-е. М., 1980, 150 с.
9. Пойя Д. Математика и правдоподобные рассуждения. М., 1976, 462 с.
10. Снедекор Д. У. Статистические методы в применении к исследованиям в сельском хозяйстве и биологии. М., 1961, 503 с.
11. Соколов Б. С. Полвека размышлений о биологии. (Рецензия на работу А. А. Любищева [5а]). — Природа, 1983, № 6, с. 118—120.
12. Тимофеев Л. И., Венгров Н. Н. Краткий словарь литературоведческих терминов. Изд. 3-е. М., 1958, 188 с.
13. Терентьев П. В. К истории биометрии. — В кн.: Методы современной
биометрии. М, 1978, с. 5—22.1
13а. Терентьев П. В. Биометрия (ретроспективный указатель отечественной литературы 1870—1970 гг.). М., 1980, 126 с.
14. Финни Д. Введение в теорию планирования эксперимента. М., 1970, 287 с.
15. Фишер Р. А. Статистические методы для исследователей (перевод с 11-го англ. издания). М., 1958, 268 с.
16. Рокицкий П. Ф. Биологическая статистика. Минск, 1964, 326 с.
17. Шенк X. Теория инженерного эксперимента. М., 1972, 381 с.
18. Шорников Б. С. Классификация и диагностика в биологическом эксперименте. М., 1979, 140 с.
19. Шрейдер Ю. А., Шаров А. А. Системы и модели. М., 1982, 151 с.

Мемориалы:

20. In memorial Chester Ittner Bliss 1899—1979.— Biometrics, 1979, voL 135, N 4, p. 715-717.
21. In memorial Ronald Aylmer Fisher 1890—1962. — Biometrics, 1964, vol. 120, N 2, p. 237—373.
22. William Gemmell Cochren (1909—1980). A Personal Tribune. — Biometrics, 1980, vol. 36, N 4, p: 575—578.
23. In memorial George William Snedecor 1881—1974. — Biometrics, 1975, vol. 31, N 1, p. 7—9.

ПРЕДИСЛОВИЕ

Настоящая книга имеет целью сделать доступной для биологов методику Р. А. Фишера по организации исследований и по обработке полученных данных методом дисперсионного анализа (анализ вариансы, ковариансы). Этот труд отнюдь не претендует заменить классические книги Р. А. Фишера (Fisher, 1937a; 1937b), но автор надеется, что она сможет оказать помощь для более легкого усвоения этой методики и для лиц, желающих дальше углубиться в предмет, может послужить введением к более близкому знакомству с оригиналом.

Сочинения Р. Фишера (1937b) при всех своих достоинствах не отличаются легкостью изложения. Наиболее доступным является его сочинение «The Design of Experiments», но оно предполагает достаточное знакомство с книгой «Statistical Methods...» (Fishery 1937a), отличающейся гораздо более трудным изложением. Кроме того, в этих книгах путь вычислений большей частью только намечен, и потому проследить весь логический ход часто бывает нелегко.

Однако совершенно ошибочно думать, что овладение только техникой вычислений (несмотря на многие высказывания, она вовсе несложна) уже позволит правильно пользоваться данной методикой. Необходимо усвоение логической стороны метода, на что и сам Фишер обращает большое внимание. Без такого усвоения и понимания тех требований, соблюдение которых обязательно для избежания ошибок, овладение только техникой вычислений не принесет пользы. Напротив, овладение математической теорией дисперсионного анализа будет полезно лишь для тех, кто стремится совершенствовать сам метод: вполне же продуктивное применение метода и приспособление его к конкретным методам требует лишь освоения логической стороны применяемых приемов. Одно из огромных преимуществ дисперсионного анализа заключается в том, что он дает возможность проверки правильности модификаций применяемых обработок. Достаточно подготовленные лица могут ознакомиться с математической теорией дисперсионного анализа по книге В. Романовского «Математическая статистика» (1938).

Необходимость остановить особое внимание на некоторых общих понятиях привела к возможности предпослать изложению дисперсионного анализа общую часть; знакомство со многими биологическими работами, в особенностигде речь идет о полевых исследованиях, обнаруживает огромное количество методических ошибок, заключающихся в несоблюдении самых основных .методических требований. Очень часто исследования предпринимаются в огромном масштабе без соблюдения гарантий отсутствия систематической ошибки, и обнаружение явно неверных выводов приводит не к реорганизации исследования и контроля за правильностью наблюдений, а к еще большему расширению так же поставленных исследований, отчего ошибка не устраняется. Ошибочность выводов не устраняется ни соблюдением единой методики, ни неограниченно большим масштабом исследования, ни полным овладением голой техникой вычислений, а исключительно сознательным применением используемых методов. Поэтому настоящая книга, как и всякое сочинение по методике организации исследований и обработке данных, отнюдь не может считаться справочником для извлечения той или иной формулы. Только внимательное, последовательное чтение с обязательной проработкой числовых примеров даст возможность применять излагаемую методику как следует.

Кроме основных книг Р. А. Фишера огромную пользу при усвоении этого метода я получил от лекций, прочитанных американским ученым Ч. И. Блиссом в 1936—1937 гг. во Всесоюзном .институте защиты растений в Ленинграде. Некоторые приемы использованы из сочинений Дж. Снедекора (Snedecor, 1934) и Ю. Л. Поморского (1935).

Для числовых примеров в большей части мной использован материал, собранный сотрудниками Института зоологии АН УССР и Украинского НИИ плодоводства. Проработка этих материалов .послужила хорошим стимулом для уточнения и углубления некоторых технических приемов статистической обработки данных.

ВВЕДЕНИЕ

Всякое исследование должно стремиться к тому, чтобы удовлетворить следующим трем требованиям: 1) оно должно быть целеустремленным, т. е. иметь перед собой определенную подлежащую решению задачу; 2) должно быть эффективным, т. е. полученные выводы должны быть настолько надежны, чтобы обладать принудительной силой, и мера надежности должна быть известна; 3) наконец, должно быть экономным, т. е. оно должно быть осуществлено с минимальной затратой сил и средств.

Осуществление первого требования вытекает из конкретных задач науки и практики и не относится к вопросу об организации опыта. Второе требование вряд ли может вызвать какой-либо спор, что же касается третьего требования, то, несмотря на его совершенную, казалось бы, очевидность, оно часто не выполняется. Напротив, многие исследователи видят особую заслугу в том, что проделано огромное количество проб, подвергнуто исследованию огромное количество объектов, даже если таким объектом является человек (в работе Де Кюри «Стоит ли им жить?» без всякого негодования сообщается об опыте, поставленном в Америке над десятками тысяч школьников, о влиянии на них пищи, лишенной витаминов). Другие исследователи стараются заменить огромное количество объектов исследования чрезмерной точностью наблюдений, полагая, что излишняя точность никогда не повредит. Очень немногие ясно сознают, что даже при правильно организованном исследовании, достаточно гарантирующем от ошибочных выводов, число исследованных объектов и точность должны вытекать из конкретных условий исследования. Если же опыт неправильно организован, то педантическая точность и огромность материала ошибочных выводов не предотвратят. Получается, как говорит Р. Фишер, что не только начинают стрелять из пушек по воробьям, но, что еще печальнее, не попадают в воробьев.

План опыта должен естественно вытекать из биологических заданий: роль же биометрии, будучи служебной, однако является далеко не маловажной, а часто совершенно необходимой. Но при всяком исследовании никогда не следует сразу использовать самые тонкие, имеющиеся в нашем распоряжении, приемы (как при микроскопическом исследовании следует прибегать к сильным увеличениям лишь тогда, когда мы уже использовали слабые). Поэтому, чисто глазомерный подход к суждению о полученных результатах, отнюдь не являясь окончательным, требует основательной проверки, сохраняя свою силу как рекогносцировочный обзор достигнутых результатов, позволяющий сконструировать рабочие гипотезы, на которые обработка должна дать тот или иной ответ. Очень часто глазомерный результат тех или иных графиков уже может указать на целесообразность или ненужность дальнейшей обработки или позволяет выбрать наиболее удобный путь обработки материалов.

Без биологически направленной мысли биометрическое исследование может привести только к накоплению часто совершенно ненужных материалов и может оказаться совершенно бесцельным. Но, с другой стороны, без математической обработки часто даже очень изощренная биологическая мысль для решения многих актуальных вопросов не в состояние преодолеть хаос изолированных фактических данных и пробиться сквозь дебри необоснованных предположений.

Задача организации опыта заключается в таком его построении, чтобы возможность неверных выводов была устранена; задала обработки материалов, напротив, заключается: 1) в отсеве всех случайных выводов, объясняемых изменчивостью собранного материала; 2) в проверке глазомерных выводов и оценке степени их надежности; 3) в извлечении максимума выводов из имеющегося материала и 4) в контроле за правильностью организации я за экономичностью работы. Изучение изменчивости явления позволит при дальнейшем планировании оценить необходимый размах работы для решения тех или иных задач.

Чтобы сделать более наглядным развиваемые соображения, я приведу несколько конкретных данных и прежде всего коснусь чисто глазомерной оценки результатов наблюдений.

ГЛАВА 1

О МАТЕМАТИЧЕСКОЙ ОБРАБОТКЕ БИОЛОГИЧЕСКИХ ДАННЫХ

1.1. ГЛАЗОМЕРНАЯ ОЦЕНКА РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЯ

В отношении к глазомерной оценке результатов исследования широко распространены два противоположных взгляда: одни считают ее достаточной для получения надежных выводов, другие, напротив, совершенно ее отвергают как ненаучную. И то, и другое суждение неправильно: и глазомерная оценка при критической ее применении может дать ценный результат, но всегда надо иметь в виду числовой контроль за ней.

Пример ошибочности глазомерного суждения приведен в работе американского ученого Г. С. Дженнингса (Jennings, Lashley, 1903) над инфузориями парамециум. Как известно, у инфузорий происходит конъюгация, при которой оба конъюганта обменивают части своего ядерного вещества. Так как обмен происходит взаимный, то физиологически оба индивида являются гермафродитами, действующими одновременно как самец и как самка, и, естественно, возникает вопрос: нельзя ли у этого животного заметить хотя бы слабое начало половой дифференцировки? Такую начинающуюся половую дифференцировку и усмотрели Калкинс и его ученица Кулль. Сущность ее они видели в том, что судьба двух инфузорий после конъюгаций имеет тенденцию быть неодинаковой: большей частью, по их мнению, из двух конъюгантов один вскоре погибает (этим он физиологически напоминает самца), а другой остается жить долгое время (этим он физиологически напоминает самку). Этот вывод Кулль делает на основании наблюдений за судьбой 186 линий (потомство 93 пар конъюгантов), причем оказывается, что через 30 дней из этих 186 линий сохранилось 103, 83 же вымерли, причем из числа этих 83 пятьдесят шесть доставляли потомство 28 пар, а остальные 27 принадлежали к разным парам. Аналогичные цифры были получены и Калкинсом. Калкинс и Кулль на основании такого материала сделали заключение о намечающейся половой дифференцировке. Другие авторы считают, что эти цифры неубедительны. Кто прав? Вопрос решен Дженнингсом в окончательном виде. Для этого надо себе задать вопрос: сколько следует ожидать погибших пар при исходном количестве 93 пары и при гибели 83 линий, если никакой связи между судьбой конъюгантов не имеется, т. е. если гибель той или иной линии ни к какой связи с происшедшей конъюгацией не состоит? Путем простых рассуждений можно показать, что если из m номеров составить пары и вытащить n номеров, то среднее число таких пар будет равно: $k = \frac{n(n-1)}{2(m-1)}$

Отсюда, применяя также формулу вероятности, вытащить определенное число пар при n вытягивании из m номеров, можно получить суждение о том, сколь велики шансы, что то или иное отклонение от наиболее вероятного значения будет иметь место. Дженнингс приводит следующие результаты таких подсчетов: как сказано, число пар 93, т. е. 186 линий (табл. 1).

Таблица 1

Срок после конъюгации	Число вымерших линий	Число вымерших пар	Вероятное число вымерших пар	Разница наблюдения и вычисления	Количество шансов против наблюдавшейся разницы
7 дней	32	6	2	4	39:1
20 »	51	13	7	6	445:1
30 »	83	28	18	10	21016:1

Результат, как видим, совершенно противоположный тому, который казался Кулль и Калкинсу: не только нет никакого намека на половую дифференцировку (тогда наблюденное число вымерших пар должно было бы быть ниже вычисленного), но, напротив, имеются вполне ясные указания на противоположный эффект конъюгации: конъюганты делаются более схожими через конъюгацию и потому судьба их оказывается более сходной. Рассмотрение всех материалов у Калкинса, Кулль и результатов собственных исследований показало Дженнингсу, что ни в одном случае никакого намека на половую дифференцировку ни разу не наблюдалось. Таким образом, эта гипотеза, конечно, в отношении парамеций, просто должна быть отброшена как основанная на неправильной интерпретации экспериментального материала.

Но, однако, из того, что глазомерная оценка, как это имело место в данном случае, приводит подчас к грубым, даже качественным ошибкам, не следует, что ее нужно вовсе отбросить. При правильном подходе она, напротив, может предостеречь нас от ошибочных выводов. Для примера приведу результаты обработки данных по связи зараженности сортов яровой пшеницы шведской мушкой и урожайностью. Для трех станций — Тамбовской, Анучинской и Воронежской — на рис. 1 приведены линии регрессии урожая по пораженности первичных стеблей шведской мушкой: эти три станции выбраны как показавшие наиболее яркую зависимость между этими двумя величинами: при статистической обработке получаются следующие коэффициенты корреляции (табл. 2).

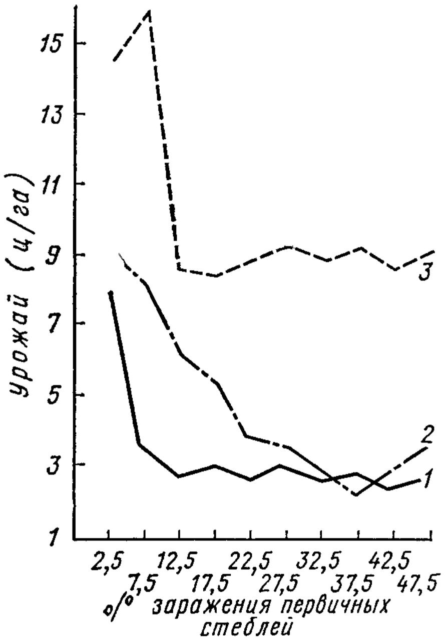

Рис. 1. Зависимость урожая пшеницы от зараженности шведской мушкой:
1 — тамбовская станция; *2* — воронежская станция; *3* — анучинская станция

Связь, таким образом, получается вполне отчетливая и на первый взгляд дает хорошее подтверждение господствовавшему одно время взгляду, что шведская мушка является главной причиной низких урожаев яровой пшеницы в Центральном районе европейской части СССР Но не говоря о том, что доводы этой гипотезы при ближайшем рассмотрении оказываются не выдерживающими .критики (Любищев, 1933б), и этот наиболее серьезный аргумент при ближайшем рассмотрении теряет свою убедительность. Общеизвестно, конечно (хотя часто забывается), что корреляция не доказывает причинной связи между явлениями и в данном случае чисто глазомерное рассмотрение линий регрессии ясно показывает, что шведская мушка не может быть не только единственной, но даже главной причиной возникновения обнаруженной связи. В самом деле, посмотрим на начерченные линии: две из них дают резкое падение, третья — сначала небольшое повышение, а затем •тоже быстрое падение. На Тамбовской станции при переходе от зараженности первого класса (в среднем принимаемого за 2,5%) к следующему (7,5%) происходит падение урожая более чем в 2 раза, а дальше колеблется около одного уровня.

Может ли быть такая картина истолкована как следствие вредного влияния шведской мушки. Для этого мы должны предположить, что увеличение на 5% зараженности первичных стеблей ведет к снижению урожая на 50%, а дальше вредное влияние прекращается. Для Воронежской станции аналогичное увеличение пораженности на 10% вызывает уменьшение урожая в 2 раза, а дальше — никакого действия. Наконец, для Анучинской станции минимум урожайности достигается при зараженности первичных стеблей в 37,5%, и урожай при этом снижается в 4 раза против максимального урожая. Ясно, что такой размер вредного влияния вредитель мог бы проявить лишь при наличии двух обстоятельств: 1) исключительно резко выраженной избирательности вредителя, т. е. если бы он нападал на растения, превышающие в несколько раз непораженные по своему потенциальному урожаю (т. е. по тому урожаю, который дало бы растение при отсутствии повреждения); 2) стопроцентной вредоносности, т. е. полной гибели всех пораженных растений. .Если бы, например, растения делились на две части: очень мощные растения, составляющие 10% от числа всех растений, но дающие 5?%. урожая, и остальные 90%, которые давали бы другие 50% урожая, и вредитель, избирательно поразивший мощные растения, .уничтожил бы их полностью, причем оставшиеся растения никакой компенсации причиненного повреждения не дали бы, то тогда поражение 10% растений могло бы повести к снижению урожая на 50%.

Таблица 2

Станция	Число наблюдений	Коэффициент корреляции	Средняя зараженность, %	Средний урожай, ц
Тамбовская	66	0,42±0,10	10,2±0,87	5,33±0,52
Анучинская	86	0,66±0,061	20,3±0,35	4,96±0,31
Воронежская	68	0,47±0,091	5,2±0,58	10,23±0,49

Но мы знаем, что подобная картина при поражении шведской мушкой яровой пшеницы никогда не наблюдается, и потому такое глазомерное рассмотрение линий регрессии заставляет нас искать какую-то другую причину низкой урожайности ряда станций и дать другое толкование.

На основании совокупности данных, изложенных в уже цитированной работе (при, установлении связи между зараженностью кустов и урожаем очень часто получается даже вполне доказанная положительная корреляция), приходится прийти к следующему толкованию. В годы с хорошими метеорологическими условиями первичные стебли уходят от повреждения шведской мушкой (а повреждения вторичных стеблей имеют гораздо меньшее значение), при весенней же засухе и отсталом развитии пшеницы первичные стебли поражаются гораздо сильнее, но и без поражения урожай оказывается очень низким: шведская мушка на яровой пшенице не является главным фактором неурожайности, а лишь добивает плохой урожай — вследствие ведущих неблагоприятных (главным образом метеорологических) условий.

Простой глазомерный подход может часто без всякой обработки показать наличие ошибок и показать необходимость изменения характера исследования. Одну из ярких иллюстраций можно заимствовать из работы Б. С. Ястремского (1937).

На рисунке 2 приведены колебания численности женского населения России по переписи 1897 г. в зависимости от возраста. Мы сразу. видим, что вместо плавной кривой, которая должна была бы быть, имеем чрезвычайно резкие зигзаги: вершины зигзагов соответствуют годам, кратным пяти и в особенности кратным десяти. Дело объясняется, конечно, тем, что среди населения широко распространена тенденция к округлению возраста, причем образование сильно снижает эту тенденцию (для грамотного населения зигзаги очень смягчены). Ясно, что если мы хотим знать точное распределение по возрастам, мы не можем непосредственно пользоваться этими данными, несмотря на то, что они основаны на многих миллионах исходных показаний. Приходится их переработать, и результат переработки, сделанный Ястремским, показан в виде пунктирной линии. Само собой разумеется, что обработка Ястремского устраняет только ошибку как следствие стремления к «округлению» возраста: если же есть тенденция к уменьшению возраста в показаниях, то для устранения этой ошибки требуются уже иные приемы. Но для нас этот пример важен как яркая иллюстрация того, что простое увеличение числа наблюдений недостаточно для избежания ошибок и что получение этих данных о возрасте должно быть не в порядке сплошной переписи, а путем правильно организованной сравнительно небольшой выборки (нескольких процентов населения), но основанной не на личных показаниях, а на документах, что даст гораздо более точную картину распределения населения по возрастам, чем сплошная перепись.

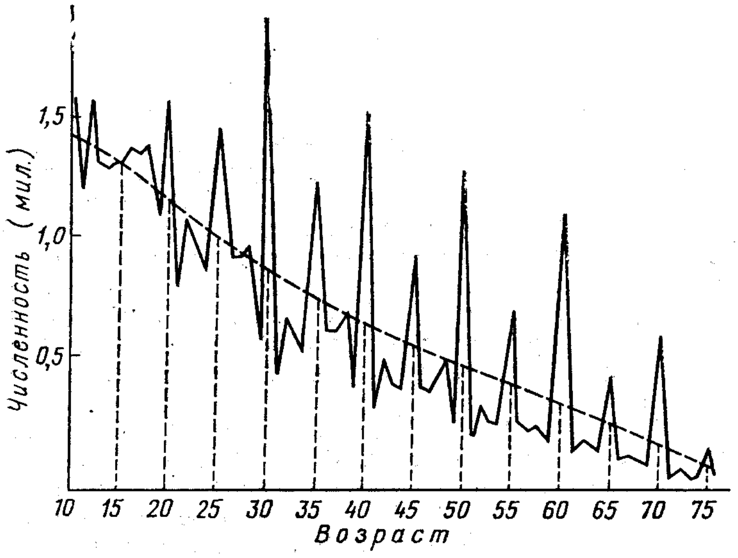

Рис. 2. Графическое изображение возрастного состава женского населения России по данным переписи 1897 г. (из работы Б. С. Ястремского, 1937) — непосредственные данные; -— выравненные данные

В области биологии нередки случаи такого же характера: производится массовое обследование по учету какого-либо вредителя, например свекловичного долгоносика. Обычно гарантии правильности видят в единстве методики обследования и в количестве проделанных проб, но если вкралась какая-либо систематическая ошибка (например, наблюдатели по неопытности и торопливости пропускают значительное количество вредителя в пробах), то увеличение числа проб может привести не к улучшению, а к ухудшению результатов, что мы часто встречаем на практике. Правильно поставленное обследование должно обязательно заключать целесообразно поставленный контроль качества работы.

1.2. О НЕОБХОДИМОСТИ ОБРАБОТКИ МАТЕРИАЛОВ

Широко распространено мнение, что при ясных и очевидных результатах и значительной разнице между различными вариантами опыта вполне достаточно привести суммарный результат по опыту и что такое сравнение дает гораздо более правильный и ценный вывод, чем последующая математическая обработка, которая, по мнению некоторых, может даже ухудшить результат. Для иллюстрации того, к каким ошибкам может приводить такой примитивный подход, разберу данные Я. В. Чугунина (1937) па эффективности пиретрума против яблоневой плодожорки (некоторые явные опечатки в его данных исправлены, но это не может повлиять на выводы).

Опыт ставился таким образом, что опрыскиванию эмульсиями пиретрума с керосином подвергались 15 деревьев (по 5 деревьев для каждой из трех концентраций пиретрума) и 3 дерева оставались контрольными — неопрыснутыми, всего 18 деревьев одного сорта и вообще одинаковых в смысле условий. Учет результатов осуществлялся при съемке урожая путем подсчета червивых и здоровых плодов. Оказалось, что среднее количество червивых плодов на. одно дерево с 911 по контролю падает до 477 у опытных деревьев, т. е. в среднем при обработке пиретрумом количество червивых плодов на дерево уменьшится на 434 плода и, следовательно, на ту же цифру возрастает количество здоровых плодов. Принимая очень умеренный средний вес плода за 50 г и беря опять-таки очень умеренно разницу в цене здоровых и червивых плодов в 60 коп. за килограмм, Чугунин приходит к выводу, что однократное опрыскивание пиретрумом дает доход в 2000 руб. на 1 га. Вот какие радужные перспективы открывает этот новый метод борьбы.

Разница в размерах червивости (почти в два раза) настолько велика,, что как будто в надежности результата сомневаться нельзя: если мы будем сравнивать средние проценты червивости, то получим опять-таки аналогичные данные? процент червивости для опытных — 22,15%, для контрольных — 49,63%, т. е. больше чем в 2 раза. Наконец, если мы применим те приемы статистической обработки, которые уже достаточно внедрились в практику, и сравним разницы, то получим разницу 49,63—22,15, или 27,48%, а средняя ошибка этой разницы будет 4,965 при обычном определении (где средняя ошибка разницы вычисляется как квадратный корень из суммы квадратов ошибок обоих величин) или 6,57, если пользоваться формулой Р. А. Фишера (1937а) с объединением сумм квадратов для обеих величин. Отношение разницы к средней ошибке получаем равное 5,535 или 4,18, что при 16 степенях свободы соответствует вероятности отсутствия существенного различия меньшей 0,001. Мы можем, следовательно, считать доказанным, что различие между опрысканными и контрольными деревьями, безусловно, не случайно.

Очень часто в таких случаях авторы выражаются так: «Эффективность опрыскивания математически доказана» — и подобные неосторожные выражения не мало послужили к тому, чтобы скомпрометировать применение биометрии в глазах опытников. На самом деле доказанной может считаться в данном случае лишь практическая невозможность (очень малая вероятность) случайного возникновения замеченных различий, но для того, чтобы быть уверенным, что это различие в червивости действительно есть следствие обработки, необходимо быть также уверенным в том, что никакое иное влияние (кроме наличия опрыскивания или неопрыскивания) не примешалось к участвующим в опыте деревьям. Предполагать наличие такого непредвиденного влияния в опыте заставляет нас прежде всего биологическая неправдоподобность результатов: пиретрум не сохраняет долго своего инсектицидного действия на воздухе, яйцекладка плодожорки чрезвычайно растянута и, следовательно, даже если бы пиретрум убивал все яички и личинки, находящиеся на яблоках в момент опрыскивания, трудно было бы ожидать 5.0% эффективности в урожае от однократного опрыскивания. Но, как увидим дальше, и чисто статистическая, несколько более углубленная обработка наводит нас на большие сомнения в правильности вывода. Приведем исходные данные Я. В. Чугунина (1937), причем отметим, что серия 1 соответствует минимальной концентрации пиретрума, серия 2 — двойной, а серия 3 — пятикратной (табл. 3).

Теперь, если мы обратим внимание на проценты зараженности по сериям, то увидим, что средние проценты пораженности возрастают от серии 1 к серии 3 и средняя пораженность по серии 3 (29,12%) почти в два раза превосходит среднюю пораженность по серии 1 (15,00). Если бы концентрация пиретрума возрастала от серии 3 к серии 1, то это было бы рассмотрено как новое доказательство эффективности пиретрума, но концентрация наибольшая как раз в серии 3 и, конечно, биологически просто невозможно себе представить, каким образом увеличение концентрации пиретрума ослабило бы его эффективность против плодожорки. От автора не ускользнула эта особенность материала. Он отмечает, что от увеличения дозировки пиретрума червивость не снижается, а даже несколько повышается. Отсюда он делает вывод, что увеличение дозировки не дает увеличения эффективности и что концентрацию пиретринов в 0,00035% надо считать предельной.

Здесь обнаруживается, к сожалению, чрезвычайно распространенная непоследовательность суждения. Снижению червивости в два раза там, где оно отвечает нашим чаяниям и ожиданиям,. придается существенное значение, а снижению почти такого же размера, но которому мы не можем дать никакого объяснения, никакого значения не придается. Последовательным ходом рассуждения будет: 1) или считать, что наш метод организации опыта безупречен, и тогда обоим сравнениям необходимо придавать одинаковое или почти одинаковое значение и делать отсюда вывод» что мы наткнулись на какое-то новое явление (повышение червивости при повышении концентрации), которому, очевидно, надо искать объяснение; 2) или признать, что такой странный вывод указывает на порочность нашего метода работы, а следовательно, подрывает, доверие и к тому выводу, который казался доказанным. Надо сказать, что в большинстве случаев при таком расхождении правильным будет именно второе заключение.

Таблица 3

Данные по эффективности пиретрума против плодожорки по Я. В. Чугунину (1937)

Количество здоровых плодов	Количество поврежденных плодов	Процент поврежденных плодов	Количество здоровых плодов	Количество поврежденных плодов	Процент поврежденных плодов
Серия 1			Серия 3		
634	137	17,7	3897	966	19,9
3341	694	17,2	3055	823	21,2
2900	367	11,2	619	359	36,7
1827	382	17,3	326	198	37,8
4287	560	11,6	968	415	30,0
Всего					
12989	2140		8865	2761	
Среднее					
2598	428	15,00	1773	552	29,12
Среднее по всем опрыснутым деревьям				476,2	22,15
Серия 2			Контроль		
3151	379	10,8			
3429	591	14,7			
2170	299	12,1	1876	1901	50,3
372	167	31,0	153	198	56,4
1064	806	43,1	870	635	42,2
Всего					
10186	2242		2899	2734	
Среднее					
2037	448	22,34	966	911	49,63

Но, может быть, разница между сериями 1 и 3 статистически не доказана. Уже простой взгляд на обе серии цифр позволяет нам утверждать, что это не так. В самом деле, сравнивая оба ряда цифр, мы видим их резкую отграниченность: трансгрессия отсутствует, самая высокая пораженность серии 1 (17,7%) ниже самой низкой пораженности серии 3 (19,9%). Случайно (а мы предполагаем, естественно, если низшая дозировка является предельной, то разница между сериями 1 и 3 носит случайный характер) Из десяти деревьев разной зараженности пять деревьев с низшей-зараженностью попадут в одну группу, а пять с высшей — в другую с вероятностью, равной

$$\frac{1\cdot2\cdot3\cdot4\cdot5;1\cdot2\cdot3\cdot4\cdot5}{1\cdot2\cdot3\cdot4\cdot5\cdot6\cdot7\cdot8\cdot9\cdot10},$$

или 1/252 (для получения этой вероятности мы подсчитываем числом всех возможных перестановок в пределах группы пяти менее пораженных деревьев, каковых будет 5 или 1-2-3-4-5, умножаем на точно такое же число всех возможных перестановок для группы; пяти более пораженных деревьев и произведение делим на общее число перестановок для всех десяти деревьев).

Если к этой вероятности прибавить вероятность прямо противоположного распределения, т. е. такого, где пораженность всех деревьев, опрыснутых низшей концентрацией, выше пораженности деревьев, опрыснутых концентрацией высшей, то мы получим вероятность случайного возникновения таких нетрансгрессивных распределений, равную 1/128, т. е. такой величины, которая ясно указывает на значимость этого различия. Таким образом, если принимать доказанность эффективности пиретрума вообще, то надо признавать и наличие противоположного эффекта: повышение пораженности с повышением концентрации пиретрума.

На самом деле вопрос решается иначе. Обратим внимание на цифры в пределах серий. Они подвержены значительным колебаниям, которые особенно велики для серии 2: максимальная пораженность в 4 раза выше минимальной, причем максимальная пораженность выше даже одного из контролей. Распределение деревьев в саду не отмечается в статье (согласно установившейся ошибочной традиции топографическое распределение считается неважным признаком), но можно думать, что оно близко к распределению деревьев на таблице, и мы видим, что деревья серий 2 и 3, близкие к контрольным, показывают пораженность значительно выше остальных. Почти с полной уверенностью можно сказать, что эффективность пиретрума в данном случае полностью или по крайней мере в значительной части объясняется неудачным выбором деревьев для опыта: контроль и несколько прилежащих к нему деревьев серий 2 и 3 выбраны в месте особо высокой естественной зараженности плодожоркой. Топографические различия зараженности были смешаны с различиями, вызванными обработкой, и получился совершенно неправильный вывод.

Но может быть задан вопрос: почему же во всех цитируемых статьях всегда бывает так, что «неудачно» выбранный участок для контроля совпадает с максимальной зараженностью? Если дело объясняется случайностью, то должны быть и противоположные примеры. Во-первых, из данного же примера мы видим, что распределение пораженностей для серий 1 и 3 оказалось противоположным ожидаемому, и это оказалось одним из доказательств мнимости вывода о высокой эффективности пиретрума, а, во-вторых, случаи, когда контроль оказывается на наименее пораженном месте, тоже нередки, но такие случаи просто не попадают» ж печать. Как раз в отношении того же пиретрума при постановке опытов Украинским НИИ плодоводства в совхозе «Зеленый Яр» в 1938 г. пораженность делянки, обработанной пиретрумом, оказалась значительно выше пораженности контроля (и при том статистически различие было вполне доказанным). Ясно, что такой результат для всякого показывает какую-то ошибку в опыте и просто отбрасывается, но почему-то огромное количество таких дефективных опытов не заставили исследователей усомниться в порочности метода, приводящего к нелепостям. Исследователи вместо этого ограничились тем, что, откидывая просто опыты, приводящие к явным нелепостям, продолжали приводить опыты, согласные с их желаниями или ожиданиями, как вполне доказывающие их выводы.

Порочность же данного метода заключалась в том, что в нем практически отсутствовала повторность, и потому топографические различия в зараженности были смешаны с различиями, вызываемыми различной обработкой. Такой ошибки мы не сделали бы, •если бы разместили деревья иначе, именно разбили бы весь наш материал (удобнее, конечно, взять не 18, а 20 деревьев, чтобы уравнять количество деревьев каждого варианта) на 5 групп по 4 дерева в каждой, причем каждая группа состояла бы из четырех, возможно близко расположенных и возможно сходных деревьев. В пределах каждой группы жребием распределили бы .деревья по намеченным вариантам. Тогда все топографические различия оказались бы связанными с группами деревьев и мы бы их не смешали с различиями между вариантами опыта.

Примеров, подобных данному, можно было бы привести неограниченное количество, частично они разобраны в моей статье о методике изучения эффективности борьбы с садовыми вредителями, напечатанной в Сборнике Украинского НИИ плодоводства (Любищев, 19406). Но для того чтобы проиллюстрировать огромную практическую важность работать правильными методами и огромные убытки, причиненные традиционный методом рассуждения, приведу далеко не полный список мероприятий, в свое время сильно рекламированных и которые даже пытались внедрить в производство, но которые все оказались целиком или в большей части основанными на методических ошибках при постановке опытов. Перечислю те, в прочности которых мне пришлось убедиться самому по собственным наблюдениям или анализу литературных и иных материалов. Вот этот крайне неполный список:

1) кремнефтористый барий в борьбе с плодожоркой;

2) трихограмма в садовом хозяйстве;

3) химический метод борьбы с садовыми слониками как радикальное средство, дающее почти 100%-ную эффективность;

4) ультрасера как полный эквивалент при опылении;

5) авиаметод (опыление серой) в борьбе с бурой ржавчиной пшеницы;

6) механические ловушки при борьбе с земляными блошками клеверным семеедом-апионом;

7) приманочные посевы для борьбы с гессенской мушкой.

Наконец, из старых методов, как будто не проводившихся научными работниками, но часто цитируемых в руководствах по борьбе с вредителями, отмечу:

8) посыпание растений пылью для предохранения от повреждений земляных блошек;

9) размещение культур (посев томатов около крестоцветных) с целью той же борьбы с земляными блошками.

Мероприятия подобного рода приводили подчас к огромной,. совершенно непроизводительной трате сил и средств и, кроме того, у многих лиц, не имевших возможности разобраться критически в вопросе, подрывали доверие и к тем средствам, которые дают действительную эффективность. Например, есть работники,. которые (очевидно, на основании справедливо критического отношения к таким инсектицидам против плодожорки, как пиретрум и кремнефтористый барий) огульно осуждают всякий химический метод борьбы против плодожорки, хотя ряд средств (парижская зелень, арсенат кальция, не говоря уже о не применяемом у нас джипсине) при правильном применении дают, несомненно, хорошую, хотя и далекую от идеала эффективность.

1.3. МЕТОДОЛОГИЧЕСКИЕ ТРУДНОСТИ СТАТИСТИЧЕСКОЙ ОБРАБОТКИ РЕЗУЛЬТАТОВ ЭКСПЕРИМЕНТА

Представление об особенной «опасности» применения статистических и вообще математических методов в биологии широко распространено и очень часто подчеркивается в руководствах по полевому методу П. Г. Лобашева (1939) и П. Н. Константинова (1939). Это нисколько не удивительно, всякое орудие тем опаснее при неразумном употреблении, чем оно совершеннее: автомобиль гораздо опаснее телеги, ружье гораздо опаснее топора, а в особенности дубины. Но вряд ли будет разумным из того, что имеет место много несчастных случаев с автомобилями и ружьями, отказаться от них и вернуться к телеге и дубине.

Все примеры опасности применения статистики и математики в биологии основаны целиком на неумелом их использовании, игнорировании при пользовании формулами тех ограничений, которые были положены при их выводе. Эти примеры полезно разобрать; вместе с тем станет многим ясно, что устранение математики от ошибок все равно не спасет, так как математика только придает точную форму верным или неверным суждениям, и если суждение было неверно, то сохранение его в неточной форме не сделает его правильным. Перечислю эти примеры, причем, конечно, приводимый перечень составляет лишь ничтожную долю в огромном разнообразии ошибок.

1.3.1. Смешение теоретических и интерполяционных формул и кривых

Под теоретической зависимостью надо понимать такую, где и форма зависимости, и все параметры уравнения выводятся на основании теоретических соображений. Если такая формула при проверке на практике на определенном отрезке подтверждается, то, естественно, мы получаем право применять формулу и за пределами проверенного нами отрезка, причем чем лучше совпадение, тем большее право мы получаем на такую экстраполяцию. Интерполяционные же формулы выбирают форму зависимости не исходя из теоретических соображений, а исключительно для удобства вычислений. Поэтому как бы ни было тесно совпадение вычисленных и наблюденных данных, делать заключение, выходящее .за пределы наблюдений, мы можем только с величайшей осторожностью и не отступая далеко за пределы наблюденного интервала. Это различие, однако, осознается не всеми, и, как это ни странно, один из наиболее ярких примеров неправильной экстраполяции я встретил не в биологической литературе. Так, известна формула геотермического градиента (т. е. возрастания температуры с углублением внутрь земли), в простейшем виде эта формула имеет вид линейной связи:

$t=a+bx,$

где x — глубина, t — температура, a и b — постоянные параметры. Но было обнаружено, что возрастание температуры идет медленнее, чем по формуле, и поэтому ввели третий параметр — с:

$t=a+bx+cx^2,$

который оказался отрицательным. В пределах исследованных глубин (так как c по абсолютному значению очень мало) с повышением глубины температура возрастает, но ясно, что при очень большом x она станет уменьшаться, а потом станет отрицательной. И вот на этом основании одному ученому пришло в голову доказывать, что, следовательно, в центре земли господствует температура чуть ли не абсолютного нуля, и подобное рассуждение по недосмотру редакции появилось даже в серьезном журнале.

В биологии особенную известность приобрел так называемый экспоненциальный закон Яниша, который некоторыми рекламировался как одно из величайших достижении биологии XX в. Между тем все теоретическое обоснование этого закона заключается в некоторых далеко не убедительных аналогиях, по существу же это чисто интерполяционная формула, менее удобная, чем давно известная гиперболическая зависимость. Однако на основании этой чисто интерполяционной зависимости Яниш считает возможным утверждать, что так как экспоненциальный закон не предусматривает порога развития, то, значит, его и не существует.

1.3.2. Игнорирование ограничений, лежащих в выводе теоретических зависимостей

Ошибка, родственная предыдущей, но более тонкая, заключается в том, что, используя ту или иную теоретическую кривую, забывают, что при выводе таких теоретических кривых практически всегда допускается некоторое упрощение действительности и что поэтому нельзя полагаться на такие кривые там, где принятое упрощение приводит к заметному отклонению от действительности. Например, общеизвестная, так называемая нормальная, гауссовская кривая распределения выведена на основе ряда предположений, из которых отметим три: 1) равная вероятность отклонения в обе стороны от средней величины; 2) возможность сколь угодно большого отклонения; 3) взаимная независимость воздействий. Если одно из требований резко не соблюдается, то и применение таблиц, основанных на гауссовской кривой, неуместно. Например, мы знаем, что фактически очень большие отклонения от средней величины просто невозможны: муха не , может быть отрицательной величины, с другой стороны, не может быть чрезвычайно большой, так как с определенным строением тела связана в известной степени и его величина. Поэтому суждение о вероятности очень больших отклонений (больше 6—7 сигм) не может основываться на таблица для гауссовской кривой: вероятность таких отклонений, вообще говоря, будет меньше, чем вычисленная по таблицам.

Другим, более распространенным и более опасным применением является использование средней ошибки и других констант, основанных на той же гауссовской кривой для асимметрических кривых, которые особенно часто встречаются при количественном учете насекомых и других организмов. Пользуясь формулами, примененными в случаях нормального симметрического распределения, делают заключение о том, что большинство вариантов должно лежать в пределах трех сигм (квадратичных отклонений), что для асимметрических распределений вовсе не обязательно. Отсюда часто вытекает и другой неправильный вывод: приблизительное суждение о вероятном размере среднего заражения по размерам максимального заражения. Избегание «опасной» статистики, конечно, от ошибок не спасает, наоборот, их усугубляет. Например, очень распространено (при оценках экономического значения вредителей и болезней культурных растений) такое рассуждение: доказанными являются случаи очень высоких процентов заражения (например, 70—80% в случае поражения мокрой головней), значит, вероятно, среднее заражение лежит где-то не ниже 10-15%. Такое заключение было бы не лишено справедливости, если бы кривая распределения зараженности была бы близка к нормальной. Но эта кривая резко асимметрична, и поэтому в том случае, где проводится действительно тщательная оценка, средняя величина оказывается значительно низшей и иногда не превышает 0,5%.

Совершенно такого же рода ошибкой будет применение формул (Любищев, 1937), основанных на законе больших чисел, к малым выборкам, на чем, например, настаивает П. Н. Константинов (1939). Но теория малых выборок имеет такое важное значение для всего дисперсионного анализа и высказанные Константиновым суждения настолько ошибочны, что об этом придется коснуться подробнее в особой главе.

Такого же рода ошибкой является применение вариационного коэффициента и показателя точности там, где они, очевидно, неприменимы. Вариационный коэффициент, вообще, имеет очень слабое теоретическое обоснование и вызывает многие возражения, но есть области, где он определенно неверен. Применение его допустимо лишь там, где есть основания (по крайней мере, эмпирические) принимать, что при одинаковой изменчивости среднее квадратическое отклонение возрастает пропорционально среднему арифметическому. Но, например, при распределении Пуассона среднее квадратическое отклонение равно корню квадратному из среднего арифметического, среднее квадратическое отклонение для вероятности одного из двух противоположных событий выражается формулой

$$\sigma = \sqrt{\frac{pq}{m}} = \sqrt{\frac{p(1-p)}{m}},$$

где p — вероятность данного события, q — вероятность события,, ему противоположного, m — число испытаний в пробе. Если для данного случая предъявлять требования (обычные в агрономической практике) одинаковой относительной точности, то мы придем к совершенно нелепому (и несмотря на это долго отстаиваемому) требованию, что чем ниже зараженность вредителями,. тем большее число проб должно быть взято.

1.3.3. Смешение вероятностей априорных и апостериорных суждений

Огромное число ошибочных применений статистических методов объясняется тем, что авторы недостаточно ясно осознали логическую природу суждений статистического характера. Этот вопрос имеет огромную важность, так как ошибки принципиально того же характера совершенно не связаны с применением статистики и вообще математики, и поэтому решительно во всех науках можно отыскать примеры, где выводы как будто убедительные на самом деле основаны на указанном смешении. Применение статистики и вообще математики только с особой наглядностью и особой точностью проявляет необходимость видения такого различия.

Под априорными суждениями мы понимаем такие, которые высказываются до опыта, и затем проверяются опытом; апостериорными же такие, которые извлекаются из опытных данных и до опыта не ставятся (подразумевая под опытом не только искусственный эксперимент, но вообще совокупность наших сведений о внешнем мире). Это более или менее соответствует теоретическим и эмпирическим суждениям, но априорные — шире теоретических. Например, в испытании сортов наше суждение о лучшей урожайности того или иного сорта может быть просто основано на предыдущем опыте, а не на какой-либо теории.

Теперь предположим, что мы произвели испытание сравнительной урожайности пяти сортов. Естественно, что пять сортов дадут неодинаковую урожайность, но ценность этого результата будет зависеть от того, ожидали ли мы до опыта именно наблюденный порядок урожайности или подходили к опыту без всякого предвзятого предположения. Если обнаруженный в опыте порядок урожайности в точности совпал с ожидаемым, то это уже очень хорошее подтверждение нашего предположения, так как случайное возникновение такого порядка можно ожидать лишь в одном случае из 120. В самом деле, пять сортов могут дать 5-4-3-2 перестановок из 120, и из них только одна соответствует нашему априорному предположению. Но если мы никаких априорных предположений не делали, то (без повторностей или других .материалов для суждения о надежности различий) опыт вообще сам по себе ничего не дает, так как какой-то порядок в урожайности сортов должен же быть.

Если опыт производится с повторностями, то он позволяет оценить вероятность и для апостериорных суждений, но следует твердо помнить, что при оценке апостериорных суждений мы должны предъявлять гораздо более высокие требования, чем при оценке априорных.

Предположим, мы сравниваем два сорта. Тогда обычно мы выясняем вопрос, есть ли между ними существенная разница или ее нет. Из этих двух возможных формулировок (есть разница или ее нет) выбираем вторую, так как такая формулировка вполне точна (разница равна нулю) и потому может служить так называемой нулевой гипотезой. Формулировка же: «есть различие» не может служить нулевой гипотезой, так как она неточна, но, например, формулировка: «разница равна одному центнеру на 1 га», как вполне точная, может служить нулевой гипотезой. Нулевую гипотезу по самой ее природе нельзя доказать, но можно опровергнуть. В самом деле, в опыте два сорта будут показывать хотя бы минимальнейшее различие, и, следовательно, мы никогда не :можем быть уверены, что нет очень малого различия. Но если различие оказывается достаточно большим, то ясно, что гипотезу отсутствия различия можно считать опровергнутой. Точно так же, если мы оцениваем нулевую гипотезу: «разница равна одному центнеру на 1 га», то мы никогда не можем быть вполне уверены, что разница точно один центнер (может быть, положим, она равна 1.00001 ц), но если обнаружится значительное отклонение от одного центнера, то поставленную гипотезу можем считать опровергнутой.

Степень надежности такого опровержения выражается вероятностью того, что обнаруженное в опыте различие может возникнуть в силу комбинации случайных причин (не связанных с сортовым различием). Чем меньше такая вероятность, тем с большей уверенностью мы можем считать нашу нулевую гипотезу опровергнутой. Поэтому, чем выше требования мы предъявляем к надежности наших выводов, тем меньшую вероятность надо брать как верхний предел вероятности случайного возникновения наблюдаемого различия. При очень широком распространенном приеме требовать тройную разницу (по отношению к средней ошибке) такая. вероятность равна приблизительно 0,003. В настоящее время принимается несколько таких уровней значимости, соответствующих. вероятностям р=0,05, 0,01 и 0,001. Эти вероятности, пользуясь терминологией Р. Фишера (1937а,в), называются доверительными вероятностями (fiducial probability), но некоторые авторы предпочитают вместо них употреблять дополнения до единиц, т. е. 0,95, 0,99 и 0,999; такие величины (по предложению Ю. Неймана) называются коэффициентами доверия (доверительные интервалы).

Таким образом, мы не должны смущаться, если в .одних таблицах стоят величины 0,05, 0,01 и 0,001, а в других при тех же значениях — 0,95, 0,99 и 0,999.

Но в некоторых таблицах (например, в таблице Н. Ф. Деревицкого (1933, 1962) в главе «Общее руководство по методике сортоиспытания») вместо 0,95 стоит 0,975, вместо 0,99 стоит 0,995 и т. д., т. е. доверительная вероятность в данном случае в два раза меньше, чем при пользовании другими таблицами, иначе говоря, те же результаты при пользовании разными таблицами как будто получают разную квалификацию. Дело объясняется тем, что таблицы Стьюдента (1942), Фишера (1938) и др. составлены из расчета оценки вероятности случайного возникновения разницы определенных размеров в обоих направлениях, т. е. и в сторону плюса, и в сторону минуса; тогда, очевидно, и вероятность такого двойного события будет в два раза большей. Таблица же с половинной доверительной вероятностью предназначена для оценки только одних положительных или одних отрицательных различий. Поэтому, если при испытании двух сортов мы испытываем гипотезу: сорт *А* не выше по урожайности сорта *В* (не ожидая, по нашим априорным соображениям, что сорт *А* может быть ниже сорта *В),* то можно пользоваться половинными доверительными вероятностями. Но так как обычно вопрос ставится не так, а проверяется гипотеза: нет ли различия в урожайности сортов *А* и *В* (не предполагая, какой сорт выше по урожайности), то необходимо пользоваться обычными доверительными вероятностями.

Но обычно испытывается не два, а много сортов и возникает вопрос, можем ли мы разницу крайних по урожайности сортов оценивать, Пользуясь таблицами интеграла вероятности (все равно для больших или малых выборок). Ответ на вопрос будет зависеть от того, является ли эта разница априорной (т. е. ожидаемой до испытания) или простым эмпирическим следствием опыта. Если, положим, в испытании имеется 10 сортов и мы будем сравнивать наиболее урожайный с наименее урожайным, то этим мы выберем самую большую разницу из 45 возможных, так как 10 сортов 10. 9 дают возможность произвести $\frac{10\cdot 9}{2}$ или 45 сравнений. Конечно, не все эти сравнения являются независимыми друг от друга, и потому предложение Р. Фишера, что в данном случае для минимальной значимости доверительную вероятность надо брать не 1/20, а 1/900, является слишком строгим, но, во всяком случае, 9 сравнений (соответственно 9 степеням свободы) являются совершенно независимыми и требование доверительной вероятности 1/180 (вместо 1/20) будет минимальным.

Непонимание разницы априорных и апостериорных предположений приводит, например, П. Н. Константинова (1939) к ошибочным выводам о недопустимости (чисто статистической) мелких делянок. Он комбинирует попарно делянки разведочного посева и, сравнивая такие пары, указывает, что «такой учет может приводить к совершенно нелепым выводам, что разница между определениями превосходит ее ошибку больше, чем в три раза, т.е. как будто мы имеем дело с двумя разными сортами». Выбирая из 840 делянок разведочного посева две пары (одну пару из малоурожайных делянок и другую —из делянок максимальной урожайности) и сравнивая их средние со средней ошибкой разности, он получает отношение разности к средней ошибке разности $\frac{69}{13.2}$, или 5,2 (т. е. значительно больше трех). Константинов еще не использовал полностью свой материал и взял не наиболее яркий контраст; если сравнить пару делянок 74 и 76 с парой 142 и 144, то получится разница средних 68 со средней ошибкой 1,41, т. е. отношение *t* равно 48/2. Доказывает ли это наличие двух сортов или непригодность малых делянок? Ни то, ни другое. Здесь две ошибки: 1) в данном случае при сравнении средних арифметических двух пар нельзя применять критерии, основанные на теории больших выборок, а надо пользоваться теорией малых выборок, где критерии значимости значительно повышены: для минимальной значимости (доверительная вероятность 0,05) требуется /, равное 4,3, для значимости 0,01—9,9, для 0,001—31,6; 2) так как в разведочном посеве 840 делянок, то число возможных комбинаций двух пар равно

$$\frac{840\cdot 839}{2}\cdot\frac{838\cdot 837}{2}$$

(так как любая комбинация двух делянок может сочетаться с любой комбинацией двух делянок из оставшегося числа), или более 12 358 000 000 комбинаций. Число независимых комбинаций будет, конечно, гораздо меньше, но при 840 исходных данных все же выражается многими тысячами (имея в виду, что мы подбираем делянки так, чтобы получить минимальное расхождение одной пары делянок), и поэтому и значения доверительных вероятностей надо уменьшить в несколько тысяч раз, чтобы получить суждение того же уровня значимости. Для суждения о том, имеются ли в материале, по отношению к которому у нас нет никаких априорных предположений, существенные различия или нет, необходимо выяснить, в какой степени средний квадрат, соответствующий определенной категории изменчивости (в случае разведочного посева это будет чисто топографическое различие поля), существенно больше среднего квадрата ошибки опыта. Об этом будет подробно изложено в специальной части данной книги (в главе о радомизированных блоках и латинском квадрате).

При суждении о различиях чисто эмпирического характера мы всегда должны стараться сообразить, как велико число проделанных испытаний, и тогда многие невероятные на первый взгляд вещи окажутся вполне вероятными. В истории рулетки в Монте-Карло был случай, когда на один цвет шарик попал 17 раз подряд: если бы поставили на то, что шарик подряд попадет на один цвет (положим красный), то вероятность осуществления нашего прогноза была бы $\frac{1}{2^{17}}$, или $\frac{1}{131072}$ - исчезающая малая вероятность. Но так как рулетка существует десятки лет и за эти десятки лет прошло сотни тысяч серий по 17 испытаний, то неудивительно, что один раз проявился и такой невероятный случай. Я не могу воздержаться, чтобы не привести несколько примеров того, к каким заблуждениям приводит это игнорирование различий априорных и апостериорных суждений, чтобы показать, что дело действительно имеет огромный и теоретический, и практический интерес.

В животноводстве были широко распространены взгляды, что существуют производители, дающие преимущественно потомство одного пола. Это основывалось на случаях, где, например, одна кобыла давала большое число потомков одного пола. Но статистическая проверка показала, что при огромном материале коневодческих заводов случаи подобных длинных «итераций» (непрерывная последовательность одного и того же события) имеют место как раз в таком количестве, как и надо ожидать при полной независимости пола потомства от производителя.

Подобное же непонимание привело к построению целых философских теорий прямо противоположного характера: так, вюрцбургский философ Марбе (исходя из редкости длинных итераций) написал целую книгу о наличии статистического выравнивания, напротив, другой философ — Штерцингер и биолог П. Каммерер развили противоположную теорию «скучивания», «закон серии», по которым краткие итерации встречаются неожиданно часто, буд-» то бы чаще, чем соответствует теоретическим расчетам (Мизес, 1930), из чего Штерцингер даже заключает о несостоятельности теории вероятностей.

Наконец, из области философии можно указать, что из факта возможности истолковать стихи известного астролога Нострадамуса как завуалированные пророчества ряда исторических событий некоторыми философами делались далеко идущие заключения совершенно мистического характера. Дело же объясняется тем, что когда мы совершенно свободны в выборе апостериорных суждений, то к туманным фразам можно подыскать бесчисленное множество толкований и выбрать из них одно, которое более или менее подходит к разбираемому нами историческому явлению. Путем такого свободного оперирования можно было бы прийти к заключению, что названия крупных городов по обеим сторонам Волги давались разными племенами (так как на правом берегу пять крупных городов мужского рода, а на левом — женского) или что названия месяцев приспособлены для любителей раков.

В биологии огромное число гипотез о филогенических связях основано на сопоставлении изолированных сходств вне всякой связи с остальными органами, поэтому и получается, что разные ученые производят (и примерно с одинаковым остроумием), например, позвоночных едва ли не от всех крупных групп беспозвоночных животных.
Соблюдение указанного различия является фундаментальным для всякой науки, но оно особенно ясно выступает там, где можно оценить количество возможных апостериорных суждений.

На разнообразных выше примерах видно, что даже в простых случаях оно чрезвычайно велико, а при сколько-нибудь сложных явлениях — практически безгранично. Это и приводит к тому, что каждый из нас, кто внимательно следит за событиями собственной жизни, без труда найдет много случаев удивительных совпадений, которые суеверных людей приводят к убеждению о наличии каких-то непонятных связей между событиями и которые на самом деле объясняются тем, что мы из бесчисленного множества сопоставлений элементарных явлений нашего жизненного опыта естественно выбираем те, которые обнаруживают совпадение.

Практическое значение правильного освоения этой разницы помимо правильной оценки результатов опыта можно видеть хотя бы из того, что распространенность азартных игр в огромной степени зависит от широко распространенного убеждения, что можно выдумать безошибочную «систему игры».

ГЛАВА 2
ТЕОРИЯ МАЛЫХ ВЫБОРОК

2.1. О ТЕОРИИ МАЛЫХ ВЫБОРОК

Теория малых выборок для полного понимания своего обоснования требует основательной математической подготовки, но логическая сторона ее чрезвычайно проста и пользоваться таблицами, основанными на этой теории, можно лишь усвоив ее логическую сущность. К сожалению, некоторые руководства (например, книги П. Г. Лобашева (1935), П. Н. Константинова (1939) о полевом методе) внесли изрядную путаницу в этот простой вопрос, а так как теория малых выборок лежит в основе всего дисперсионного анализа и создана Стьюдентом (1942) и Фишером (1937а), то целесообразно остановиться на ней несколько более подробно.

Теория малых выборок появилась лишь в начале XX в. и заполнила весьма существенный пробел математической статистики. Все существовавшие до этого таблицы интегралов вероятности основывались на том предположении, что число наблюдений, послуживших для определения эмпирических констант (средней арифметической,, среднего квадратичного отклонения, средней ошибки и т. д.), очень велико. При справедливости такого предположения (в выводе число наблюдений принималось даже бесконечно большим) справедливы и соотношения между величиной и своей средней ошибкой, т. е., например, согласно хорошо известному правилу трех сигм вероятность того, что случайно взятый индивид совокупности будет отстоять от арифметического среднего (в ту или другую сторону) не менее чем на три средних квадратичных отклонения, равна приблизительно 0,003.

Всякая математическая формула справедлива, конечно, только при соблюдении тех условий, которые были положены в ее выводе. Так как эти условия строго никогда ; не соблюдаются, то формулы применяются и при отклонении от принятых условий, но только, конечно, если эти отклонения не будут очень велики. В практической работе исследуемая выборка из общей совокупности, конечно, никогда не бывала бесконечной, а иногда состояла всего из нескольких экземпляров. Естественно, что применимость формул классической математической статистики к подобным случаям вызывала большие сомнения, и, например, в прекрасном руководстве Е. Е. Слуцкого (1912) по теории корреляции не рекомендовалось вычислять коэффициенты корреляции менее чем для 30 наблюдений, причем и это число устанавливалось на глаз. Имелись и работы, прямо доказывавшие, что применимость критериев ошибки, выработанных для выборок большого объема к малым выборкам, приводит к грубо ошибочным суждениям.

Работы Стьюдента (1942) и Фишера (1937а, b) в этом отношении составили эпоху в математической статистике, так как при соблюдении определенных ограничений (именно наличие закона нормального распределения в генеральной совокупности) они дали возможность определить вероятность отклонения определенного размера для выборок любого объема. На основе теории составлены таблицы, приводимые во всех новейших руководствах по вариационной статистике. Из этих таблиц ясно, что если для получения минимальной значимости (доверительная вероятность, равная 0,05) отношение величины к своей средней ошибке должно равняться 1,96 при бесконечном числе испытания, 2,00 при 60 степенях свободы, 2,042 при 30 степенях свободы (разница, как видим, очень небольшая, позволяющая даже число 30 признавать достаточно большим), то для пяти степеней свободы мы имеем t, равное уже 2,57, для трех— 3,18, для двух — 4,3 и для одной степени свободы (т. е. при наличии, например, только двух исследованных индивидов совокупности) t возрастает до 12,71. Мы видим, что для очень малых выборок совершенно неприложимо правило, применимое для больших выборок, что точность вывода возрастает пропорционально корню квадратному числа исследованных: выигрыш точности при переходе от одной степени свободы к двум оказывается неизмеримо большим, почему при простых сопоставлениях работать с одной повторностью совершенно не рекомендуется. При использовании коэффициента корреляции ограничение, выдвигаемое Е. Е. Слуцким (1912) (наличие не менее 30 дат), отпадает. Можно работать с любым количеством данных, но значимость вывода будет очень различна. Так, если мы имеем сто наблюдений двух признаков, то коэффициент корреляции между ними, равный 0,19, уже достигает минимального уровня значимости (доверительная вероятность 0,05 или коэффициент доверия 0,95); при 30 степенях свободы (32 наблюдения двух признаков) для получения того же коэффициента доверия требуется коэффициент корреляции, равный уже 0,35, при восьми наблюдениях (шесть степеней свободы) — 0,707, при пяти наблюдениях (три степени свободы) — 0,878, при двух степенях свободы — 0,950 и при одной — 0,997. Ясно, что при наличии трех наблюдений минимально надежный вывод о наличии корреляции можно сделать только в случае исключительно высокой корреляционной связи.

Насколько мне известно, теория малых выборок пользуется бесспорным признанием среди лиц, разбирающихся в математических основах этой теории. Она же послужила основой для развития всей системы дисперсионного анализа, чрезвычайно повышающего эффективность и экономичность наших исследований: внимательный читатель в этом убедится из содержания данной книги. Тем более удивительными являются те возражения, которые выдвигаются в силу недоразумений авторами обоих указанных руководств по полевому методу, — П. Г. Лобашевым (1939) и П. Н. Константиновым (1939). Эти авторы выдвигают следующие возражения против теории малых выборок: 1) теория малых выборок предъявляет чрезмерно строгие требования и часто приводит к тому, что вполне зарекомендовавшие себя методы оказываются опороченными как недоказанные; 2) теория малых выборок стремится путем математических операций улучшить качественно негодный материал; 3) теория малых выборок родилась из стремления построить опыт на очень мелких делянках. Разберем эти возражения по очереди.

2.1.1. Чрезмерная строгость теории малых выборок

П. Н. Константинов (1939, с. 136) указывает, что, пользуясь таблицей Стьюдента—Фишера, якобы совершенно невозможно получить достаточной достоверности вывода при трех повторениях, «между тем каждый опытник знает, что не только трех, но и двух повторений при оптимальной учетной площади, например, в 0,05—1,0 га, часто достаточно, чтобы получить совершенно достоверные результаты». Здесь целый ряд недоразумений. Таблица t у П. Н. Константинова приведена только для значений t, не превышающих 6,1, и при таком значении действительно трех повторностей недостаточно для получения вывода высокой надежности, но сам автор приводит примеры сортоиспытания, где при трех повторностях t достигает 11,0—14,3 и выводы приобретают высокую статистическую надежность. Очень часто ссылаются, например у П. Г. Лобашева (1939), на данные Ставропольской станции, где якобы обработка не выявила надежной разницы даже для вполне оправдавших себя приемов. Эти данные, по-видимому, не попали в печать, и потому о них судить, вообще говоря, трудно, но все такие выводы (ненадежность статистической разницы при несомненном наличии разницы вариантов) объясняются неумелой обработкой материала. В главе о факториальном анализе можно видеть разбор примера, взятого из руководства П. Н. Константинова (1939), где различие урожая для сроков посева только при выбраковке значительной части материала приводит к выводам удовлетворительной надежности, при правильной обработке без всякой браковки и при пользовании теорией малых выборок приводит к результатам исключительной надежности.

Такие неудачные выводы из хорошего материала объясняются двумя основными ошибками: 1) неумением выделить изменчивость, связанную с повторениями опыта (путем взятия вместо разности средней — средней разности) применением корреляционного метода или дисперсионного анализа; 2) неумением объединять материал. Например, опыт велся в течение пяти лет и за каждый год вариант А давал больший урожай, чем вариант В. При неумелой обработке сравниваются варианты за каждый год порознь, где разность сама по себе ненадежна, и если даже во всех пяти случаях разность одного знака, то такая однозначность результата (которая и служит основанием для вывода о бесспорной эффективности исследуемого приема) не учитывается при обработке. Умение же объединять данные опыта приводит к тому, что, стоя целиком на почве теории малых выборок, сложные опыты можно проводить с одной повторностью, а в некоторых случаях и без повторности вовсе. Об этом будет разъяснено в главе по факториальному анализу и в главе о повторностях.

Что же касается мнения опытников о том, что и при малой повторности можно получить высоко надежный вывод, то истинное зерно этого утверждения заключается в том, что при подробных осмотрах полей опытный работник учитывает, во-первых, огромную массу сопутствующих признаков (рельеф, почва, наличие или отсутствие сорняков и т. д.), а во-вторых, глазом можно оценить степень равномерности стояния растений на обследованных участках. Иначе говоря, в этом случае уже получается глазомерное суждение о степени надежности общей цифры, характеризующей урожай любого участка. Если же мы имеем просто эту одну цифру, то значимость ее может быть очень различной, смотря по тому, равномерен ли урожай по полю или нет, как он распределен и т. д.

Знание этих данных позволит, опять-таки применяя правильную статистическую обработку, прийти к выводам несравненно более надежным, чем простой глазомерный подход даже при отсутствии повторностей (об этом опять см. в главе о повторностях), но это ни в какой мере не подрывает значения теории малых выборок для оценки степени надежности отдельных цифр, касающихся одного признака. Можно совершенно смело утверждать, что не может быть таких случаев, где бы правильная статистическая обработка ухудшила вывод, сделанный без ее помощи, но только нужно, чтобы статистическая обработка имела в своем распоряжении весь материал, лежащий в основе глазомерного вывода, а не выжимку (часто даже неправильно произведенную) из данного материала.

2.1.2. Стремление «улучшить» путем математической обработки качественно негодный материал

В своих работах П. Н. Константинов (1939) прямо утверждает, что при малых размерах делянок и малом числе повторений вообще нельзя обрабатывать материал и что никакие математические ухищрения не позволят прийти к правильному выводу.

Это обвинение против теории малых выборок прямо противоположно первому: там теорию упрекали в чрезмерной строгости, в том, что даже очевидные выводы оказываются неубедительными, а здесь, наоборот, ей инкриминируется то, что путем математической обработки стремятся выжать нечто из качественно непригодного материала.

Цитируется выражение Гексли, что «математика — жернов, который всякую засыпку смелет, но ценность помола определяется исключительно ценностью засыпанного».

Выражение Гексли, как и всякое удачное выражение, правильно только при правильном понимании, а именно: 1) математика не один жернов, а огромный набор разных жерновов, и умелый выбор подходящего жернова может сильно улучшить качество помола; 2) ценность материала очень часто можно определить только произведя пробный помол.

Теория малых выборок и дает возможность определить, в какой мере ценным является собранный материал. Использование же многих признаков в нашем материале часто позволяет значительно улучшить наши выводы, но само собой разумеется, что если исследователь не владеет предметом своего исследования, то обработка приведет только к тому выводу, что собранный материал никуда не годен, и в том случае, когда ведется учет количества индивидуумов и в рамках опыта имеются сильные колебания численности, эта зависимость есть и она действительно может привести к тому, что существенные контрасты снизят свою надежность, а малосущественные могут приобрести мнимую значимость. В этих случаях большую пользу приносит преобразование исходных данных — приемы Бартлета (Кендалл, Стьюарт, 1973) и Блисса (Bliss, 1925, 1937), о чем будет речь в гл. 4.11.

2.1.3. Теория малых выборок и размеры делянок

Как было уже указано, П. Н. Константиновым утверждается, что теория малых выборок родилась в связи со стремлением работать на малых делянках. В этом нет никакой связи. Теория малых выборок просто учитывает то обстоятельство, что дат для определения средних немного. Если проводится опыт в каком-нибудь огромном совхозе и испытывается один прием, положим, в трехкратной повторности, то всего получается шесть делянок и *за* вычетом двух средних арифметических мы получаем четыре степени свободы для суждения о степени надежности разницы полученных средних. Для суждения о степени надежности полученной разницы (при условии, конечно, что в постановке опыта были приняты предосторожности, устраняющие возможность систематических ошибок) размер делянок совершенно безразличен (он может равняться даже клетке, т. е. 400 га), но, конечно, мы вправе ожидать, что надежность будет тем больше, чем крупнее делянки (при той же повторности). Чем это объясняется? При малых делянках, естественно, даже если предположить, что наше поле прекрасно выравнено и в общем однородно, мелкие отличия в гетерогенности поля существуют — здесь теория малых выборок решительно ни при чем.

Совершенно правильно, что не следует любой материал чисто механически обрабатывать самыми разнообразными способами, и выбор наиболее подходящих приемов обработки часто вытекает из глазомерного исследования (в форме кривых, графиков и т. д.), ,но неверно, что более точная обработка может дать худшие результаты, чем «простой агрономический анализ». Пример, приводимый П. Н. Константиновым (1939), где простое сравнение двух сортов дает отношение разности к средней ошибке, равное 4,7 (очень хорошее), а по формуле Стьюдента (основанной на обобщенной средней ошибке) 1,4 (совершенно ненадежное), заключает в себе две ошибки: 1) грубая арифметическая: П. Н. Константинов использует квадраты ошибок вместо самих ошибок, и поэтому отношение разности к средней ошибке в его примере равно 2,24 (а не 4,7), что при данном числе степеней свободы едва .достигает минимального уровня значимости; 2) для сравнения Константинов взял сорта (два из 13), показавшие минимальную изменчивость при сравнении повторностей: так как это сравнение чисто эмпирическое, то, учтя это обстоятельство, мы приходим к выводу, что надежность взятого им сравнения не достигает минимального уровня значимости.

Что же касается возражения против обобщенной средней, то оно имеет некоторую справедливость лишь в том случае, если размер дисперсии не независим от количественного значения переменных. В большинстве случаев (с урожайными данными и другими величинами) зависимость если и есть, то слабо выражена и не вызывает никаких искажений, в некоторых же случаях (там, где даты выражают в процентах определенные результаты, или на однородных участках) она будет сильнее сказываться: чем больше делянка, тем, естественно, все больше будут нивелироваться такие малые отличия и среднее квадратическое отклонение будет уменьшаться. Для каждого данного опыта невозможно априори установить размеры делянок и количество повторностей, но, имея уже некоторый опыт работы в определенном месте, можно на основании обработки материалов прийти к заключению о желательном размере делянок и о достаточном числе повторностей. Тут как раз теория малых выборок может оказать, огромные услуги. Поэтому вполне понятно (что, кажется, странным Константинову), что теория малых выборок и дисперсионный анализ нигде не ставят вопроса о необходимых размерах делянок и числе повторностей априори, но всегда ограничиваются тем, что указывает, как проверить, что использованное число повторностей и принятые размеры делянок оказались достаточными для получения удовлетворительной надежности выводов.

Но, конечно, теория малых выборок сохраняет свою справедливость лишь при соблюдении (или не существенном отклонении) нормального закона изменчивости генеральной совокупности. Если мы это забудем, то в некоторых случаях можем сделать неправильный вывод, что теория малых выборок недостаточно чувствительна к установлению значимости различий и, следовательно, что мы имеем право, как это предлагает П. Н. Константинов. (1939), с ней просто не считаться, и даже при малом числе испытании пользоваться, например, старым правилом трех сигм. Такой вывод был бы, разумеется, большой ошибкой. При нормальном распределении, как это показано на примерах в главе 3.1 о повторностях, сопоставление разницы со своей ошибкой на основе теории малых выборок может значительно увеличить надежность вывода, и такой результат является правилом, если же распределение явно ненормальное, то сравнение разницы со своей средней ошибкой может дать как будто бы сильное ухудшение степени надежности вывода.

Для иллюстрации возьму специально придуманный пример. Предположим, что вариант *A* дал в пяти случаях такие цифры:: 2; 9,5; 9,5; 9,5; 9,5, сумма 40, среднее 8,0 (M_0=9,5), а вариант *B* дал такие результаты: 10; 10; 10; 10; 20, сумма 60, среднее 12,0 (M_0=10); является ли разница между обоими вариантами существенной или несущественной.

Применяя общий критерий (применимый к любой форме распределения дат), получаем, как и в случае с пиретрумом у Я. В. Чугунина (1937), вероятность случайного возникновения подобного (т. е. не трансгрессивного) распределения, равную 1:128. Таким образом, различие достаточно надежно для варианта *A,* применяя критерий *t* мы получим:

сумма квадратов отклонений дат варианта *A*,
от средней по *A* 45,
то же, для варианта *B* 80.

Средняя ошибка разности равна (ввиду равенства числа дат обоих вариантов):

$$\frac{45+80}{4\cdot 5}, \text{ или } 2,5$$

Таким образом, *t* оказывается равным 4:2,5, или 1,6. Такое *t* даже при бесконечно большой выборке соответствует вероятности случайного возникновения различия лишь немногим меньшей 0,1, а при восьми степенях свободы вероятность оказывается лежащей между 0,2 и 0,1, т. е. указывает на абсолютное отсутствие какой-либо значимости наблюдаемого различия. В этом случае, как мы видим, дело вовсе не в мнимой малой чувствительности теории малых выборок, так как даже пользование классической теорией больших выборок не спасает положения. Какой же вывод правилен? Ведь получив первым способом вероятность, равную 1 : 128, мы как будто никаких ограничений на форму распределения не накладывали, следовательно, наш вывод доложен быть как будто совершенно общим и более правильным, чем вывод, полученный другим методом? На самом деле это не так. Первый метод заключает тоже известное ограничение: именно предполагается полная независимость изменения индивидуумов опыта, дающих исходные даты, только тогда применение правил комбинаторики приводит к безупречным результатам.

Зададим же себе вопрос, имеются ли основания думать, что эта независимость колебаний (что выражается числом степеней «свободы) действительно соблюдена в опыте, давшем такие результаты? Простое рассмотрение цифр заставит нас предполагать, что этой требуемой независимости не было. В самом деле. из пяти дат каждого варианта четыре даты тождественны (они .могли бы быть и нетождественны, но близки, это существенной разницы в рассуждение не вносит), а пятая отстоит от них очень далеко. Это указывает на наличие групп в пределах каждого варианта, отличающихся более резко, чем совокупности обоих вариантов, а следовательно, позволяет усомниться в правильности организации опыта (наличие не устраненной опытом, ясно выраженной гетерогенности поля). Поэтому всякий случай, где простое вычисление вероятности дает лучшую надежность результатов, чем определение *t,* требует тщательного анализа. Очень часто такое расхождение является серьезным указанием на неправильность организации опыта и, следовательно, на ненадежность полученных выводов. Для материала Я. В. Чугунина (1937) это можно утверждать с полной уверенностью.

Для избежания подобных ошибок и будет правильное осознание разницы между систематическими и случайными ошибками, правильное проведение принципа повторностей, о чем говорится в отдельных главах. Но совершенно прочным является путь, на который вступают многие исследователи, именно простое увеличение материала для исследования. Так, например, в упомянутой работе Де Крюи «Стоит ли им жить?» приводится указание на опыт, проведенный в Соединенных Штатах Америки по вопросу о влиянии витаминов на рост детей. Для этого несколько тысяч детей школьного возраста были подвергнуты разным пищевым рационам: в одном случае полноценной пищей, а в другом с недостатком витаминов. Опыт продолжался несколько месяцев и, конечно, привел к результату, что питавшиеся неполноценной пищей дети отстали в росте от питавшихся полноценной пищей. В таком виде этот опыт — бессмысленная жестокость по отношению к детям, которых сознательно держали на неполноценной диете, но, несмотря на много тысяч подопытных индивидов, результат может считаться надежным лишь в случае его правильной организации и устранения возможности систематических погрешностей.

Пример того, что даже сплошное исследование не предохраняет от грубых ошибок, можно заимствовать из статьи Б. С. Ястремского (1937) (см. рис.2). Речь идет о распределении по возрастам населения России по переписи 1892 г. Мы видим ясно выраженные зигзаги: огромные пики, соответствующие десяткам лет, и немного им уступающие для лет, кончающихся на 5. Все объясняется тенденцией (совершенно не имеющей никакой экономической базы) округлять свой возраст, и в отдельных случаях число лиц определенного возраста по переписи может в пять раз превышать истинное. Таким образом, даже сплошное исследование не предохраняет от ошибок и не может быть использовано непосредственно, без обработки.

2.2. ОБ ОБЪЕДИНЕННОЙ ДИСПЕРСИИ

При дисперсионном анализе для оценки размеров изменчивости, связанной с погрешностями опыта, пользуются общей средней ошибкой, основанной на объединенной дисперсии (вариансе,, т. е. квадрате среднего квадратического отклонения). Этот же прием Фишер (1958) применяет и для простейшего случая, для определения средней ошибки разности. Обычно (в случае независимости изменения обоих переменных) ошибку разности определяют по формуле

$$m_d = \sqrt{m_1^2 + m_2^2}$$

и t определяется как отношение разности средних арифметических $\overline{x_1} - \overline{x_2}$ (черточка над буквой обозначает среднюю арифметическую) к этой средней ошибке разности. Так как всякая средняя ошибка равна $\frac{\sigma}{\sqrt{n+1}}$, где σ равно стандартному отклонению, $n+1=n^1$ — число наблюдений, послуживших для определения средних, следовательно, n — число степеней свободы, а

$$\sigma = \sqrt{\frac{\Sigma(x-\overline{x})^2}{n}},$$

то средняя ошибка разности приобретает вид

$$m_d = \sqrt{\frac{\Sigma(x_1-\overline{x_1})^2}{n_1(n_1-1)} + \frac{\Sigma(x_2-\overline{x_2})^2}{n_2(n_2-1)}}$$

Взамен этого Фишер определяет среднюю вариансу для разности, разделяя общую сумму квадратов на сумму степеней свободы, т. е. получая

$$\frac{\Sigma(x_1-\overline{x_1})^2 + \Sigma(x_2-\overline{x_2})^2}{n \ + n_2}$$

и для определения квадрата средней ошибки умножает эту вариансу на сумму обратных величин чисел дат, т. е.

$$\frac{1}{n_1^1} + \frac{1}{n_2^1} = \frac{1}{n_1+1} + \frac{1}{n_2+1}.$$

Формула для определения средней ошибки разности приобретает таким образом следующий вид:

$$m_d = \sqrt{\frac{[\sum(x_1-\overline{x_1})^2 + (x_2-\overline{x_2})^2](n_1+n_2+2)}{(n_1+n_2)(n_1+1)(n_2+1)}}.$$

При равенстве дат обеих сравниваемых величин обе формулы приводят к тождественным результатам, в чем нетрудно убедиться. При большом числе дат разница между ними тоже невелика. Но существенные различия получаются при небольшом и неодинаковом числе дат. Так, в уже приведенном примере с эффективностью пиретрума против плодожорки при определении средней ошибки разности (всех опытных данных и контроля) обычным способом получаем среднюю ошибку, равную

$$\sqrt{7{,}742 + 16{,}955}\text{, или } \sqrt{24{,}657} = 4{,}965$$

Если же применить формулу Фишера, то получим $\sqrt{\frac{1727{,}284 \cdot 18}{16{,}3 \cdot 15}}$, или 6,57 т. е. значительно большую величину. Не следует думать, что формула Фишера дает всегда большую величину: это бывает только тогда, когда изменчивость дат одной величины, основанных на большом числе степеней свободы, превышает изменчивость дат другой величины, основанной на меньшем числе степеней свободы. Большей частью бывает наоборот, и тогда формула Фишера приводит к меньшим размерам , средней ошибки разности, чем обычная формула.

Теперь возникает вопрос: можно ли доверить формуле Фишера (1958) без вникания в те математические доказательства, на которых она основана. Простое размышление покажет, что формула Фишера имеет явные преимущества перед старой. В самом деле, предположим, мы сравниваем две величины (это могут быть, положим, величины двух видов), из которых одна основана на очень ограниченном и трудно добываемом материале, а другая на материале, который мы можем произвольно увеличивать.

Если средняя ошибка второго вида по нашему скудному материалу оказалась довольно большой, то при использовании обычной формулы, очевидно, никакое увеличение материала первого вида не позволит уменьшить среднюю ошибку разности ниже средней ошибки второго вида (подставить для проверки этого в формулу m_1, равное нулю). Напротив, при применении формулы. Фишера средняя ошибка разности может быть сделана меньше средней ошибки вида, основанного на малом числе наблюдений. В тех же случаях, где формула Фишера приводит к большей строгости, эта строгость вполне понятна, так как малое число дат легко могут случайно дать пониженное значение средней ошибки и гораздо более правильным будет принять большие размеры средней ошибки, основанной на большем числе дат.

2.3. РАЗЛИЧНЫЕ КАТЕГОРИИ ОШИБОК ИССЛЕДОВАНИЯ

Классифицировать ошибки исследования можно исходя из различных принципов. Наиболее важными являются два: 1) в зависимости от характера отклонений ошибок от истинного значения; 2) в зависимости от полноты исследования и точности отдельных наблюдений.

Группировка по первому признаку приводит к делению ошибок :на систематические и случайные, по второму — на ошибки репрезентативности и ошибки точности.

2.3.1. Систематические и случайные ошибки

Единственное различие между систематическими и случайными ошибками заключается в том, что систематические ошибки имеют тенденцию варьировать только в одну сторону, случайные же колеблются в обе стороны от истинного значения. В качестве иллюстраций систематических ошибок можно привести и обыкновенно приводят ошибки, происходящие от не совсем правильной калибровки измерительных инструментов, от влияния индивидуальных свойств исследователя. Например, человек с более острым зрением и большим навыком при анализе стеблей на зараженность шведской мушкой найдет больший процент зараженности, чем малоопытный и с менее острым зрением, так как последний легко пропустит, не заметив самых маленьких личинок и т. д. Совершенно ясно, что если мы одну серию наблюдений будем делать, положим, с термометром, дающим ошибку на полградуса в положительную сторону, а другую серию с другим термометром, ошибающимся на полградуса в отрицательную сторону, то, даже если бы эти две серии наблюдений касались совершенно тождественного явления, разница между ними будет один градус; и хотя бы мы проделали миллионы наблюдений, размер систематической ошибки останется тем же самым.

Поэтому первой задачей всякого исследования является устранение систематических ошибок. Как это сделать?

В отношении инструментов наиболее простой и удобный путь — внесение поправок: все точные инструменты снабжаются поправками, которые и вносятся в результаты наблюдений. При астрономических наблюдениях производятся предварительные испытания наблюдателей, и результаты этих испытаний дают поправку для каждого наблюдателя в виде так называемого личного уравнения. Путь поправок обычно наиболее удобный путь, но он не единственный, не всегда приложимый и не всегда наиболее удобный.

Другим путем является превращение систематических ошибок в случайные, иначе говоря, превращение их из односторонних в двусторонние. Это достигается тем, что мы каждую серию наблюдений проводим не одним инструментом или наблюдателем, а несколькими инструментами или наблюдателями. Если в двух сравниваемых сериях наблюдений в каждой серии совершенно в одинаковых отношениях участвовали имеющиеся у нас инструменты или наблюдатели, то при сравнении этих двух серий ошибка, связанная с инструментом или наблюдателем, будет полностью устранена; если же мы, не имея возможности полностью уравновесить наши серии, распределяем инструменты и наблюдателей по жребию (так что на каждую серию падает несколько инструментов или наблюдателей), то систематическая ошибка превращается в случайную.

В отношении инструментов такой путь, вообще говоря, нецелесообразен, так как: 1) он требует нескольких инструментов. там, где при пользовании поправками можно обойтись одним; 2) вследствие перехода систематической ошибки в случайную, размер случайной ошибки значительно увеличивается, и для получения вывода той же надежности требуется провести значительно большее число наблюдений. Но он может стать необходимым, например, в том случае, если данные заводом поправки затерялись или инструмент в силу долгой работы требует проверки, а мы не можем этого сделать. Как известно, в наиболее тщательно проводимых астрономических исследованиях, даже при наличии поправок, наблюдения проводятся серией инструментов, а не одним инструментом (хронометром и т. д.).

При биологических исследованиях вопрос о том, устранить ли систематическую ошибку путем поправки или перевести ее с помощью рандомизации (т. е. распределения по жребию, о чем подробнее будет ниже) в случайную, решается в каждом данном случае в зависимости от обстоятельств. Например, очень большие прения вызывает вопрос о так называемых разведочных посевах. Почва на опытных станциях никогда не бывает тождественна по своему плодородию, и так как это обстоятельство влияет на результаты опытов, вызывая повышенную изменчивость, то рекомендуют за год и более до опыта провести посев на том же участке, и распределение урожая по полю использовать для суждения о различии плодородия почвы на участке с внесением соответствующих поправок в результаты опытов. Этот путь вполне уместен и рационален, и наилучшим способом использования таких разведочных данных является применение линий регрессии в анализе ковариансы, но способ этот не единственный и даже часто не наиболее удобный.

Неудобство его в том, что он требует года предварительных исследований, т. е. удлиняет продолжительность исследования, почти удваивает объем исследования, и очень часто вносимая им поправка не окупает затрату времени: если мы распределим варианты опыта по системе рандомизированных блоков на необследованном ранее поле, то, увеличив повторность, придем к столь же надежным выводам, как и в первом случае, но в два раза более короткий срок. Но, несомненно, на старых опытных станциях следовало бы вести историю всех участков поля и использовать эту историю для внесения поправок в опыты текущего года. Игнорирование такой истории приводит иногда к неожиданным фактам. Например, на Тамбовской опытной станции, по наблюдениям В. Н. Щеголева и собственным (Любищев 19406), неоднократно наблюдалось резкое различие урожаев на участках по всем условиям как будто тождественным. При расспросе старожилов выяснилось, что особо плодородные участки находились на местах, где ранее была распаханная дорога или свалка и т.д..

Уже из этого изложения ясно, что граница между систематическими и случайными ошибками в каждом исследовании не является неподвижной и что мы имеем полную возможность, исходя яз обстоятельств данной работы, перемещать ошибки из одной категории в другую.

Так, например, размеры случайной ошибки могут быть уменьшены: 1) увеличением числа наблюдений, 2) увеличением точности отдельных наблюдений и 3) превращением известной части случайной изменчивости в систематическую с последующим ее устранением. И здесь нельзя сказать для всех случаев, какой путь лучше. Выбор пути диктуется конкретными обстоятельствами. Иногда точность отдельных наблюдений бывает явно недостаточна, например, если мы производим измерения длины тела насекомого или части его столь грубым масштабом, что получается всего два или три класса изменчивости: точность измерений всегда должна быть такой, чтобы получалось не менее десяти классов. Но если получается несколько сот или тысяч классов (например, при измерении с точностью до 0,1 мм при амплитуде колебаний изменчивости в 30—40 см), то такая точность является совершенно излишней и добавочный труд, затраченный на получение такой точности, истрачен совершенно впустую.

Очень часто третий путь — превращение случайных ошибок в систематические с последующим их изучением и устранением — является наиболее эффективным. Вернемся, например, к старому примеру с инструментами. Если одна серия наблюдений проводится одним инструментом, а другая— другим, то случайная ошибка невелика, но весьма вероятна систематическая ошибка (если нам не известны поправки), искажающая результат. Если, напротив, обе серии проводятся обоими инструментами, то систематическая ошибка устраняется, но сильно вырастает случайная, могущая в данном случае не исказить, а смазать результат, сделать его неотчетливым. Наилучшее решение достигается таким образом, что сравнение двух серий производится лишь для наблюдений, проведенных одним и тем же инструментом: тогда и систематическая ошибка оказывается устраненной и случайная не увеличенной. А сравнивая результаты показаний, полученных разными инструментами, мы по ходу опыта можем определить поправку для каждого инструмента.

В биологических условиях мы имеем некоторую аналогию. Очень часто приходится слышать, что опыты по эффективности тех или иных мероприятий, положим в саду, можно проводить лишь при наличии односортных, одновозрастных деревьев и вообще в однородных насаждениях. Вообще говоря, это и не необходимо и недостаточно. Это не является, вообще говоря, необходимым потому, что некоторые приемы обработки только количественно отличаются на деревьях разного возраста и сорта, и, правильно размещая опыт (распределяя по жребию варианты опыта), мы можем гарантировать себя от систематической ошибки. Но случайная ошибка может быть очень велика, и потому такая постановка опыта, вообще говоря, нерациональна на опытных станциях, имеющих дело с малым числом повторностей, но вполне допустима в широких производственных опытах, поставленных в десятках и сотнях точек. Но постановка опытов в однородном по всем известным нам признакам саду не является и достаточной для получения надежных выводов, так как наряду с явными признаками неоднородности сада (сорт, возраст, размеры кроны и т. д.) имеются открытые признаки неоднородности, заключающиеся, как правило, в микроэкологических различиях почвы, рельефа и т. д.

Эти различия приводят к тому, что урожай соседних деревьев часто существенно отличается от урожая совершенно, казалось бы, тождественных по сорту, возрасту и т. д. деревьев, расположенных поодаль. Такие систематические различия часто искажают результат при постановке опытов путем размещения всех деревьев одного варианта вместе (в садоводстве это часто называют «деляночный» метод). Вместе с тем отклонения эти не настолько определены, чтобы мы могли устранить их путем внесения некоторых постоянных поправок, подобно тому, что мы делаем с инструментами. Наиболее удобным путем будет разбитие сада на несколько участков и помещение в каждом участке (так называемом блоке) полного набора исследуемых вариантов. Этим путем мы значительную часть случайной изменчивости выделяем а особую категорию и связываем ее с определенным топографическим различием. Случайная ошибка может оказаться сильно уменьшенной, а надежность опыта — значительно повышенной. Подробно об этом рассказано в главе о рандомизированных блоках и факториальном анализе.

Чем более оказывается изученной область производства исследований, тем большее количество признаков может быть выделено, по которым мы производим группировку нашего материала, и, следовательно, все большая и большая часть случайной изменчивости может быть систематизирована, изучена и устранена. Имеются два основных пути такого уменьшения случайной изменчивости в рамках реорганизации опыта (т. е. без увеличения числа наблюдений и без увеличения точности отдельных наблюдений): 1) если признаки лежат в пределах нашего контроля (т. е. признаку может быть дано значение или искусственно, или путем выбора объекта, обладающего данным значением признака), то работа идет по пути латинского и греко-латинского, и высших квадратов, и факториальной схемы опыта; 2) если признак не подлежит нашему контролю, но может быть учтен (например, метеорологические данные), то, применяя анализ ковариансы .(в простейшем случае используя просто линии регрессии), можно сильно сократить случайную изменчивость.

Мы видим, таким образом, что при одной и той же точности наблюдений и одном числе наблюдений случайная изменчивость может быть сильно сокращена путем рациональной организации опыта. Предел такому сокращению кладется тем, что если мы возьмем очень много критериев классификации (при факториальном анализе или высших квадратах), то опыт становится чересчур громоздким, а если возьмем много учитываемых сопутствующих признаков (для анализа ковариансы), то обработка делается непосильной. Искусство постановки опыта и заключается в. том, чтобы из огромного количества возможных влиятельных факторов выделить для учета ведущие, а все второстепенные сознательно перевести в категорию случайной изменчивости. Это, конечно, достигается не знанием математики, а знанием теории явления, наблюдательностью и умением использовать предыдущий накопленный опыт.

2.3.2. Ошибки репрезентативности и ошибки точности

Ошибки репрезентативности возникают всегда, когда от сплошного обследования какой-либо совокупности (например, всенародной переписи переходим к несплошному (см.: Старовский, 1931). В биологической работе сплошное обследование имеет малое значение (может быть, в охотничьих хозяйствах или при спуске рыбоводных прудов) и ввиду его дороговизны почти никогда не применяется. Ошибки репрезентативности могут быть и систематическими, и случайными, и только в последнем случае можно сказать, что размер ошибки соответствует размеру выборки; как и всегда устранение систематической ошибки может быть достигнуто жеребьевкой или для всей совокупности, или для отдельных типических групп (метод, вполне соответствующий методу рандомизированных блоков). Без наличия случайного отбора мы никогда не можем быть гарантированы от грубейших ошибок в выводе (см. пример из той же работы В. Н. Старовского (с. 121),. где полученные данные в 1,5—2 раза превышали действительные).

Примеры систематических ошибок репрезентативности можно привести следующие: 1) обследование полей только определенного срока посева (положим, исключая поздние сроки) или лишь. по близости от населенных пунктов может резко исказить картину о количественном составе тех или иных вредителей; 2) регистрация данных только о наличии вредителей без регистрации тех пунктов, где обследование не показало наличия вредителя,, этим путем искусственно значительно сокращается число обследованных мест и средняя зараженность может оказаться значительно выше истинной.

Но не следует думать, что если мы произведем сплошное обследование, то этим самым избежим ошибки, или что сплошное обследование всегда дает результат лучше выборочного (подразумевая, согласно В. Н. Старовскому (1931), под выборочным методом ту форму несплошного обследования, которая устраняет систематические ошибки репрезентативности). Сплошной метод. устраняет только ошибки репрезентативности и оставляет нетронутыми ошибки точности, которые сами могут быть систематическими и случайными. Едва ли не наиболее ярким примером таких систематических ошибок точности является проявившееся при многих переписях стремление населения показать свой возраст, округляя до 5 (Ястремский, 1937). Эта ошибка не является систематической по отношению к среднему возрасту всего населения, потому что при округлении может быть и увеличение, и уменьшение числа лет, но по отношению к числу лиц, обладающим возрастом, кратным пяти, она является систематической. Эта ошибка сравнительно легко устраняется путем обработки материала. Труднее приходится с теми систематическими ошибками, где имеет место общая тенденция к сокращению (например, к показанию меньшего возраста по сравнению с истинным). В дан»ном случае правильно поставленное выборочное обследование гораздо меньшего числа лиц, но проведенное с большей точностью (например, по метрическим свидетельствам, а не по личным показаниям) даст несравненно более правильный вывод, чем сплошная перепись.

ГЛАВА 3
О ПОВТОРНОСТЯХ В ПОЛЕВОМ ЭКСПЕРИМЕНТЕ

Вопрос о повторности опыта или наблюдения вызывает, что вполне понятно, очень большую дискуссию в литературе, и существует довольно широко распространенное мнение, что количество повторностей может быть более или менее стандартизировано; с другой стороны, практика многих опытников сводится к тому, что, проведя опыт в определенной повторности, данные по повторностям в конечном счете обезличиваются, выражаются средней арифметической и тем дело заканчивается. Наконец, во многих случаях то, что в опыте называется повторностью, не является повторностью вовсе, так как не соблюдается основное требование — взаимная независимость повторностей. Следует помнить, что повторность только тогда достигает цели, когда каждое повторение независимо от другого, и что введение таких независимых друг от друга повторностей преследует две цели: 1) уменьшение ошибки при оценке результатов опыта; 2) возможность оценить размеры этой ошибки.

Первая цель достигается простым использованием средних арифметических, вторая — только путем использования данных по всем повторностям. Разберем более подробно эти три вопроса: 1) что следует называть настоящей повторностью; 2) о необходимости использования данных по повторностям и 3) о числе повторностей.

3.1. ПОНЯТИЕ О НАСТОЯЩЕЙ ПОВТОРНОСТИ

Настоящей повторностью, как я уже указывал, можно называть повторение опыта только в том случае, если оно будет независимым. Это требование покоится на соображениях простого здравого смысла, которые, однако, часто позабываются.

В самом деле, показания отдельных повторностей могут быть сравнены с показаниями свидетелей на суде, а свидетелей всегда. допрашивают порознь, чтобы они не могли сговориться .или чтобы показание одного не повлияло на показания других. Если же мы имеем, положим, сто совместно допрашиваемых свидетелей, то цена их общему показанию будет равна показанию одного свидетеля.

В чем же заключается зависимость повторностей друг от друга?

Положим, мы производим исследование по влиянию различных способов запарки коконов шелковичного червя на размотку, и для простоты примем только два варианта опыта: один способ — простое подогревание в горячей воде, а другой — в автоклаве под давлением. Положим, что мы размотали по сто коконов обоими способами и нашли, что размотка под давлением вызывает большое количество разрывов. Этот вывод (конечно, при условиях производства опыта) будет справедлив только в том случае, если за весь период исследования оба способа чередовались, а не было сделано так, что сначала разматывали сто коконов одним способом, а потом другим. В этом последнем случае .могло случиться так, что, положим, первая партия попала на период сухой и жаркой погоды, а вторая — на период влажной, конечно, влажность воздуха является весьма существенным фактором, несомненно влияющим на количество разрывов нити. В этом случае при неправильном размещении нашего исследования сделанный нами вывод о преимуществе того или иного способа размотки может оказаться в значительной степени искаженным тем, что не принятое нами во внимание различие (в данном случае различие по влажности воздуха) оказалось связанным с тем различием, которое мы исследуем.

Другим примером грубого нарушения правила независимости повторностей является широко распространенный в садовых опытных станциях так называемый «деляночный» метод организации опыта. Берется, положим, несколько вариантов опрыскиваний. Для каждого варианта отводится по нескольку деревьев, но все эти деревья помещаются рядом, образуя одну делянку. Но как бы ей был однороден сад, в нем всегда имеется неравномерность — по почве, рельефу, уровню почвенных вод, распределению вредителей и т. д., и, как правило, рядом расположенные деревья по всем этим признакам более сходны между собой, чем болев отдаленные. Поэтому, когда все деревья одного варианта расположены в одном месте, а все деревья другого — в другом, то различие между ними обусловлено не только различиями между характером обработки разных вариантов, но часто в гораздо большей степени различиями в условиях сада, и полученные данные сплошь и рядом или вводят нас в заблуждение (когда они отвечают нашим надеждам), или же приводят к таким явным нелепостям, что заставляют забраковать часто очень дорого стоящий опыт.

Забота о независимости повторностей должна постоянно тревожить исследователя, так как часто, казалось бы, ничтожное уклонение от такой независимости может иметь совершенно нежелательные последствия. Приведу опять-таки пример из собственного опыта (Любищев, 1931). На Безенчукской станции в 1931 г. ставились опыты сортоиспытания: опыты были поставлены в шести повторностях с совершенно независимым, на первый взгляд, расположением сортов в отдельных повторностях. Но оказалось, что было сделано одно упущение, в силу чего повторности не оказались независимыми, а именно каждый сорт во всех шести повторностях засевался подряд и благодаря ошибке в установке сеялки (отчего первая серия сортов засевалась пониженной нормой) была сделана грубейшая ошибка в выводах, именно было принято (и статистически доказано), что твердые пшеницы меньше изреживаются проволочниками, чем мягкие. Ошибка была вскрыта только путем тщательного исследования топографических отношений.

Таких форм зависимости повторностей друг от друга бесконечно много: один вариант может быть поручен одному наблюдателю, другой — другому, и тогда индивидуальные особенности наблюдателей (опытность, добросовестность) окажутся смешанными с теми различиями, которые являются предметом исследования.

Забота о сохранении независимости отдельных повторностей и лежит в основе метода рандомизированных блоков. Он заключается в том, что все варианты одной повторности образуют так называемый блок, т. е. собрание делянок, индивидов, частей индивида и т. д., по возможности более сходных между собой по всем признакам; и в пределах каждого такого блока варианты размещаются по отдельным делянкам, деревьям и т. д. по жребию с целью избежать всякой субъективности в выборе объектов опыта.

Из изложенного вовсе не следует, что если по техническим причинам или по каким-либо другим соображениям невозможно работать с соблюдением требования независимости повторностей, то такой материал никакой цены не имеет и не может быть использован. Он может быть использован с применением специальных методов (например, введением сопутствующих признаков, применением топографического подхода и т. д.), но эти методы всегда значительно сложнее методов, применимых тогда, когда исследование поставлено с соблюдением требования независимости. Обычные методы здесь могут привести к грубейшим ошибкам, и надо прямо говорить, что в данном случае мы работаем без повторностей. О приемах использования такого материала будет указано дальше.

3.2. ИСПОЛЬЗОВАНИЕ ДАННЫХ ПО ПОВТОРНОСТЯМ

Существует широко распространенное и совершенно ошибочное мнение, что при одном и том же числе повторностей результат всегда тем надежнее, чем больше оказалась разница между сравниваемыми вариантами, почему в большинстве отчетов опытных станций результаты опытов приводятся только в виде средних арифметических. На простом примере, заимствованном из книги Фишера (1937а), можно показать совершенную ошибочность такого взгляда.

Положим, мы поставили опыт по влиянию того или иного приема на урожай и получили в двух повторениях прибавку в 8 и 9 кг на единицу площади, а в другом случае тоже при двух повторениях получили прибавки в 8 и 18 кг на ту же единицу площади. В среднем в первом случае получили 8,5 кг прибавки, а во втором — 13. Какой результат надежнее? Судя по среднему арифметическому, — конечно, второй, так как там прибавка в полтора раза превышает прибавку первого опыта. Но если мы без всякой арифметической обработки приглядимся к цифрам, то, пожалуй, усомнимся в этом выводе. В самом деле, для первого опыта показания (8 и 9 кг) мало отличаются друг от друга, следовательно, мы вправе заключить, что многочисленные посторонние влияния на урожай сравнительно невелики по сравнению с тем влиянием, которое мы исследуем, и что, следовательно, влияние исследуемого нами приема может считаться довольно обоснованным.

Во втором случае разница между двумя повторностями огромна и, следовательно, наряду с влиянием изучаемого нами приема имеются налицо другие факторы, искажающие картину. И, возможно, вся прибавка объясняется не изучаемым нами приемом обработки, а именно этими, независимыми от нашей обработки факторами. Этот глазомерный вывод может получить и математическое выражение, если мы вычислим среднюю ошибку для обоих случаев и определим вероятность нулевой гипотезы, в данном случае предположения, что никакого влияния наши приемы на урожай не оказывают и что все различия объясняются исключительно «случайными» влияниями не учитываемых нами факторов. Средняя ошибка в первом случае равна 0,5 кг (так как разница обоих дат от средней равна 0,5, то сумма квадратов разниц равна тоже 0,5 и квадрат ошибки равен 0,25), а во втором — 5,0 кг, следовательно, отношение средней арифметической к своей средней ошибке равно для обоих случаев

$$\frac{8{,}2}{0{,}5} \text{ и } \frac{13}{5}$$

или 17,0 и 2,6 (эта величина называется, как известно, *t*). Так как в обоих случаях имелось всего по две даты и, следовательно, имеется всего одна степень свободы, то по таблицам Стьюдента—Фишера можем найти, что вероятность полного отсутствия существенного влияния исследуемых нами приемов в первом случае лежит между 0,05 и 0,01, а во втором — между 0,3 и 0,2. Следовательно, второй опыт приводит к совершенно ненадежным результатам, а на основе первого можно принять с довольно серьезной уверенностью наличие влияния нашей обработки на урожай.

Поэтому при наличии повторностей или данные по повторностям должны быть обработаны как следует, или, если это по той или иной причине невозможно, все данные по повторностям (а не только средние арифметические) должны быть приведены в отчете.

3.3. О ЧИСЛЕ ПОВТОРНОСТЕЙ

Число повторностей определяется двумя особенностями: степенью изменчивости нашего материала и степенью сложности нашего исследования. Как увидим дальше, при достаточной сложности исследования можно обойтись без повторностей вовсе, по крайней мере без явных повторностей, т. е. без соблюдения полной одинаковости контролируемых нами условий: повторность в. широком смысле слова, конечно, сохраняется, но она оказывается скрытой, т. е. элементы повторности можно извлечь из сходных элементов различных дат.

В настоящей главе я не буду касаться сложных опытов, а коснусь лишь весьма существенного вопроса: какова должна быть повторность в тех случаях, если опыт дает наилучшие возможные результаты. Вернее, здесь я коснусь вопроса, какова надежность глазомерных выводов при отсутствии колебаний в результатах. Из чисто педагогических соображений буду оперировать произвольно подобранными простыми числами.

Положим, мы сравнивали два варианта опыта в двух повтор-йостях и в обоих случаях получили превышение варианта *В* над вариантом *А*. Можем ли мы удовлетвориться таким глазомерным выводом и считать, что различие двух вариантов доказано? Есл» мы рассматриваем каждую повторность порознь, то ясно, что при. полном отсутствии существенных различий вариант *В* может превышать вариант *А* и наоборот: при двух испытаниях (повторно-стях) согласованность показаний (превышение *В* или *А* в обоих случаях) имеет вероятность 1/4 для каждого варианта или 1/2 для обоих.

Следовательно, если мы до опыта не ожидали превышения того или иного варианта, а просто сравниваем до того неизвестные нам приемы, то совпадение результатов двух повторностей не имеет абсолютно никакой значимости. Если мы имеем три повторности, то совпадение во всех трех случаях при полном отсутствии существенных различий между вариантами может быть в $1/2^3$ или в 1/8 числа опытов (в том случае, если до опыта мы уже ожидаем превышение варианта *В* над *А)* или в 1/4 числе опытов, если мы просто сравниваем неизвестные нам ранее приемы, по поводу которых теоретически мы никаких предположений не высказывали. Само собой разумеется, что вероятность 1/8 чисто случайного возникновения нашего результата слишком велика, чтобы можно было признать наши выводы (о преимуществе одного варианта перед другим), имеющими какую-нибудь доказательность.

Между тем в практике опытных учреждений очень часто исследователь (при простом опыте и без учета количественной стороны) считает, что превышение результатов одного варианта по сравнению с другим во всех трех повторениях уже «доказывает» преимущество. Для того чтобы такое простое сравнение имело хоть какую-нибудь претензию на значимость, вероятность «случайного» возникновения наблюденного результата никак не должна быть больше 1/20 (или 0,05), а обычно предъявляются более строгие критерии: требование вероятности, не превышающей 0,01. В этом случае при простом опыте и при сравнении в пределах каждой повторности отдельно требуется не менее пятикратной повторности в первом случае (так как $1/2^5$ равна 1/32, т. е. меньше 0,05) и не менее семикратной во втором (так как $1/2^7$ равна 1/128, или меньше 0,01) при условии, что опыт дал «идеальные» результаты, т. е. во всех повторностях один вариант превышал. другой без исключения, и что превышение одного варианта над. другим предвиделось на основании теоретических соображений или на основе предшествовавшего опыта; если такого предвидения не было, то число повторностей надо увеличить на одну.

Все эти соображения справедливы тогда, когда мы сравниваем результаты в пределах каждой повторности отдельно, не объединяя данные по повторностям. При таком объединении в случае «идеального» результата количество повторностей может быть сокращено. Для более наглядной иллюстрации приведу два произвольно подобранных результата опыта, причем в обоих случаях мы имеем по восьми одинаковых цифр, но распределены они по-разному.

		Первый опыт		Второй опыт	
		вариант *А*	вариант *В*	вариант *А*	вариант *В*
повт.	1	5	9	5	10
»	2	7	10	7	11
»	3	8	12	8	12
»	4	11	14	9	14

Если рассматривать все повторности порознь, то, казалось бы, два опыта дают одинаковую надежность, так как всегда вариант *В* превышает вариант *А;* следовательно, вероятность случайного возникновения такого результата (при ожидании, что вариант *В* превысит вариант *А)* равна $(1/2)^4$, или 1/16. Но, присмотревшись внимательнее к обоим опытам, мы видим, что они существенно отличаются. В первом опыте в пределах каждой повторности вариант *В* превышал вариант *А,* но если мы возьмем всю совокупность результатов, то увидим, что результат 4-й повторности варианта *А* превышает 1-ю и 2-ю повторности варианта *В,* таким образом, имеется трансгрессия.

Во втором опыте трансгрессии нет, так как максимальное значение варианта *А* (4-я повторность) меньше минимального значения варианта *В* (1-я повторность). Какова вероятность случайного возникновения (при отсутствии существенного различия вариантов *А* и *В)* во втором опыте? В данном случае мы имеем такой результат: из восьми дат четыре меньших оказались относящимися к варианту *А* и четыре больших — к варианту *В.* Для того чтобы определить вероятность случайного возникновения такого распределения (при отсутствии значимой разницы между обоими вариантами), надо подсчитать общее число возможных распределений и число распределений, подобных наблюденному (в том смысле, что все четыре меньшие даты попали к варианту *А,* а все четыре большие — к варианту *В).* Общее число возможных распределений равно, очевидно, числу перестановок для 8 элементов, т. е. 8; 7; 6; 5; 4; 3; 2; 1. Число же распределений, подобных наблюденному, получается из возведения в квадрат 4; 3; 2; 1, так как и меньшие и большие даты могут образовать по 4; 3; 2 перестановки и каждая перестановка «меньших» дат может комбинироваться с каждой перестановкой «больших» дат. Следовательно, искомая вероятность будет:

$$\frac{4\cdot3\cdot2\cdot4\cdot3\cdot2}{8\cdot7\cdot6\cdot5\cdot4\cdot3\cdot2}, \text{ или } \frac{1}{70}$$

Таким образом, при таком результате, какой мы имеем во втором опыте, достаточно надежный вывод получается уже при четырехкратной повторности, а при пятикратной повторности (10 дат) мы таким образом получим вероятность, равную

$$\frac{5\cdot4\cdot3\cdot2\cdot5\cdot4\cdot3\cdot2}{10\cdot9\cdot8\cdot7\cdot6\cdot5\cdot4\cdot3\cdot2}, \text{ или } \frac{1}{252}$$

т. е. надежность вывода получается чрезвычайно высокая (>0.039).

Но все эти результаты получаются тогда, когда мы принимаем во внимание лишь сопоставление больше — меньше, не обращая внимания на численное значение всех дат. В этом последнем случае надежность вывода может значительно увеличиться при той же повторности. Это ясно уже из такого простого соображения: очевидно, если вариант *A* в четырех повторностях дал результаты 5; 7; 8; 9, а вариант *B* — 10; 11; 12; 14, то (если не обращать внимания на распределение дат по отдельным повторностям) вероятность случайного возникновения такого результата будет не больше 1/70. Но сразу бросается в глаза, что порядок распределения величин варианта *B* соответствует порядку распределения величин варианта *A*. В данном случае наиболее эффективным критерием сравнения будет определение средней разности обоих вариантов, которая для четырех повторностей соответственно равна: 5; 4; 4; 5, в среднем 4,5 со средней ошибкой 0.289. Следовательно, отношение разницы к своей средней ошибке (4,5 к 0,289) равно 15,6. Так как здесь мы имеем всего четыре цифры разностей, то нельзя использовать обычные таблицы интеграла вероятностей, основанные на большом числе дат: здесь надо пользоваться таблицами для малых выборок Стьюдента — Фишера. Однако, пользуясь даже этими таблицами, мы убеждаемся, что вероятность такого t при трех степенях свободы значительно меньше 0.001, т. е. результат исключительно надежен. Даже результат опыта 1 оказывается вполне надежным, так как там разница по повторности оказывается соответственно 4; 3; 4; 3, т. е. в среднем 3.5 с той же средней ошибкой 0.289 и t равно 12.2, т. е. вероятность случайного возникновения такого результата (при отсутствии существенной разницы) лишь немногим превышает 0.001 (для 0.001 t равно 12.94).

Этот пример ясно показывает, что сплошь и рядом чисто глазомерная оценка результатов может дать вполне надежные выводы, но такая оценка требует, во-первых, совершенной согласованности результатов, а во-вторых, приводит к сильно повышенным требованиям в отношении повторности, чем в том случае, если результат будет обработан как следует. Работа «на глаз» — крайне неэкономная работа.

3.4. ВОЗМОЖНОСТЬ РАБОТЫ БЕЗ ПОВТОРНОСТЕЙ В СЛОЖНОМ ИССЛЕДОВАНИИ

Предыдущие рассуждения о числе повторностей касаются только простых по своей организации исследований (опытов или наблюдений), т. е. когда мы сравниваем между собой варианты,. не связанные в одну общую систему, и когда каждое отдельное наблюдение или опыт дают только одну дату. В этом случае справедливы все те требования к числу повторностей, которые вытекают из математической статистики. Но существует широко распространенное недоразумение, разделяемое многими авторами-опытниками, например П. Н. Константиновым (1939), что «теория малых выборок» неприменима в полевом опыте и что, вопреки этой теории, можно пользоваться гораздо меньшим числом повторностей и даже обойтись совсем без повторений. Разобрать это недоразумение и составляет предмет настоящей главы. Действительно, во многих случаях можно работать без повторностей, но это нисколько не умаляет значения теории малых выборок, а только является иллюстрацией того, что всякая теория может прилагаться лишь к той области, для которой она была создана. Разберем несколько таких случаев, где работа без повторностей может привести к достаточно надежному выводу.

а) Выявление эмпирических констант интерполяционных кривых. Этот случай особенно часто встречается в точных науках,. где отдельные наблюдения можно получить с высокой точностью и где, следовательно, неконтролируемые влияния очень слабы. В биологии это встречается реже ввиду обычно большей изменчивости биологических явлений.

Положим, мы изучаем воздействие какого-либо фактора на интересующий нас признак: количества удобрения на урожай, температуры на продолжительность развития насекомого, количества пищи на вес откармливаемых животных и т. д. Нам надо сделать вывод не только о том, что имеет место то или иное изучаемое нами влияние, но и определить размеры и характер этого влияния. 'Положим, теперь, что мы исследуем влияние удобрения на урожай и нашли, что при двух дозах удобрения урожай оказался .выше, чем при одной дозе, а при одной дозе выше, чем при полном отсутствии удобрения. Если результат формируется только таким чисто качественным образом (или если не обращается внимания на размеры прибавок урожая в обоих случаях), то надежность такого вывода невелика. В самом деле, три различных числа можно расположить (число перестановок из трех) шестью различными способами: из них один способ (монотонно возрастающий) является подтверждающим наше предположение. Следовательно, в одном случае из шести результат может быть делом случая, при полном отсутствии реального воздействия удобрения. При четырех испытаниях (без удобрения и с тремя различными дозами удобрения) и вполне монотонном характере полученной зависимости (повышение урожая с каждой новой дозировкой) вероятность нулевой гипотезы (отсутствие влияния удобрения) будет уже 1/24, при пяти — 1/120, при шести — 1/720 и всегда выражается величиной $\frac{1}{n!}$, где n — число различных испытаний.

Мы видим, таким образом, что, хотя здесь повторности нет (так как каждое испытание производится с новой дозировкой), но .благодаря тому, что все различные испытания составляют одну систему, можно получить надежные выводы и без повторностей. На основании таких неповторяющихся данных можно построить .интерполяционную кривую и вычислить соответствующие эмпирические константы.

б) Использование топографических отношений. Этот случай особенно важен для прикладной биологии, и всего удобнее разобрать его, пользуясь графической иллюстрацией (рис.3). .На рисунке 3 показаны четыре схемы результатов опытов, причем в каждом опыте сравнивались только два варианта I и II. С участка каждого варианта было взято

25.проб по одной линии, причем во всех четырех случаях (если брать все пробы, не учитывая их топографических отношений) мы имеем одинаковые средние .для того же варианта и совершенно» одинаковые средние ошибки, так как все различие четырех опытов *A*—*D* заключается лишь в пространственном размещении тождественных 25 дат. Поэтому, ^если мы имеем просто две серии для двух вариантов по 25 дат каждая, то хотя бы между этими средними и было вполне статистически доказанное различие, мы не вправе сделать вывода о доказанности различия между вариантами, так как здесь нет настоящей повторности (все 25 дат одного варианта находятся в одном участке земли) и, возможно, что статистически доказанное различие относится не к различию вариантов, а к различию естественно исторических условий опытного поля. Здесь и помогает использование топографических отношений. Разберем схему *A*: урожай участка II заметно в среднем выше урожая участка I, но *мы* видим, что в пределах обоих участков урожай постепенно повышается слева направо, так что если построить линии регрессии, то линия регрессии участка I непосредственно переходит в линию регрессии участка II. В данном случае мы решительно можем сказать, что на самом деле никакой разницы варианты опыта не обнаруживают, все «статистически доказанное» различие участков целиком сводится к неравномерности . урожайности поля, в данном случае плавному повышению урожаев слева направо.

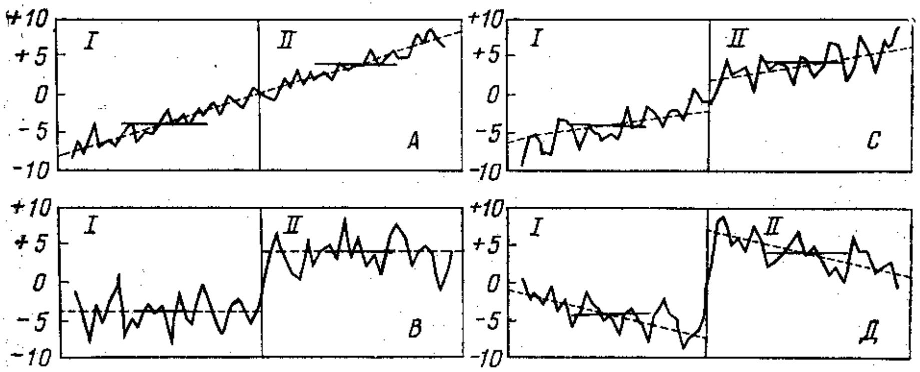

Рис. 3. Четыре случая соотношения двух вариант при одинаковом отличии средних сравниваемых вариантов и при одинаковом распределении около среднего, но при разном расположении в поле. Для всех четырех случаев разность средних равна 8, средняя ошибка разности 0.702

Совершенно другую картину показывает схема 5. В пределах каждого поля урожаи колеблются около одной прямой регрессии и, сопоставив прямые обоих участков, мы получаем разницу, соответствующую разнице средних. Если, следовательно, не было допущено какой-либо грубой ошибки в постановке опыта (положим, участок I был посеян значительно позже участка II), то на основании результатов, подобных схеме *B,* можно сделать надежное заключение о различии вариантов I и II, хотя опыт проведен без повторностей. Схемы *C* и *D* дают некоторое усложнение в схеме C, «истинная разница» вариантов меньше разницы средних арифметических, так как часть разницы относится за счет падения урожайности справа налево, напротив, в схеме *D* «истинная» разница больше разницы средних арифметических, так как естественное падение урожая (независимо от опыта) идет слева направо. Использование топографических отношений может иметь чрезвычайно широкое приложение в экологии и прикладной биологии и позволяет контролировать наши выводы, с одной стороны, *а.* с другой — работать при минимальном количестве повторностей, а иногда и при полном отсутствии их.

Сознательное или бессознательное использование топографических отношений и лежит в основе того, часто повторяемого опытниками утверждения, что объезд полей и при отсутствии повторности может дать опытному работнику более ясное представление о результатах опыта, чем биометрическая обработка большого числа повторностей. Сущность этого утверждения заключается в том, что при обычно практикуемой биометрической обработке совершенно игнорируются топографические отношения, и потому (опять-таки, повторяю, при обычной шаблонной обработке) биометрическая обработка не дает ответа на вопрос, какой схеме рисунка отвечают результаты нашего опыта, а простой обход участка сразу позволит решить вопрос без всякой обработки. Но дело здесь, как и всегда в подобных случаях, не в том, что на глаз можно решить вопрос точнее, чем при помощи математических вычислений, поставленных как следует быть, а в том, что даже очень совершенные орудия исследования не приносят никакой пользы лицу, неумеющему ими пользоваться.

Отсюда совершенно ясно, что и теория малых выборок никакого отношения к данному случаю не имеет, так как теория малых выборок оперирует с элементарными датами, а участок поля при охвате его в целом отнюдь не является элементарной датой. Если же мы искусственно соединяем содержание нашего опытного участка, сводя его к одной элементарной дате (например, валовой урожай всего участка), то тогда, конечно, теория малых выборок вступает в свои права. Но ошибка здесь заключается в том, что мы все богатство содержания нашего опыта подменяем одной цифрой, т. е. крайне расточительно используем наш материал, используя ничтожную часть полученных цифр. К сожалению, такая расточительность является далеко не редким явлением в экологии я прикладной биологии.

в) В сложных опытах. Наконец, работать без повторности можно при сложных опытах по так называемой факториальной схеме Р. Фишера (1937, 1958), используя взаимодействия высших порядков для оценки размеров случайной ошибки. Об этом будет речь в разделе, посвященном факториальной схеме и факториальному анализу.

Подобными сложными опытами, допускающими очень низкую повторность или полное отсутствие повторности, являются широкие географические испытания по срокам посева, районирование сортов и т. д. Если брать каждый такой опыт отдельно в одной точке, то, конечно, требуется повторность, но если объединить опыты многочисленных точек, то при таком объединении и учете экологических различий отдельных точек (по почве,, климату и т. д.) можно получить вполне отчетливые выводы даже при отсутствии повторностей. Такой опыт является также сложным факториальным опытом, и поэтому целесообразно даже при возможности организации повторности в каждой точке делать эти повторности не идентичными, а вводить новый фактор (например, глубина заделки, норма высева и т. д.). Отсутствие синтетического подхода и объясняет тот странный факт, что, несмотря на огромную работу, проведенную на многочисленных опытных станциях, мы до сего времени не имеем обоснованного районирования.

ГЛАВА 4
ТЕОРИЯ И ПРАКТИКА ДИСПЕРСИОННОГО АНАЛИЗА

4.1. О РАНДОМИЗАЦИИ ЭКСПЕРИМЕНТА

Термин «рандомизация» (английское слово *randomization* от слова *random* — случай) введен Р. Фишером, но само понятие рандомизации и необходимость применения рандомизации вытекают с необходимостью из самой сущности опытного дела. Рандомизация заключается в том, что мы вводим распределение по случаю (жребию) всюду, где мы имеем основание опасаться смешения систематических и случайных ошибок. Недостаточно ясное понимание сущности рандомизации привело к тому, что стали вводить жеребьевку там, где она вовсе неуместна, а отсюда возникла реакция, отрицающая необходимость рандомизации и там, где она вполне уместна и необходима.

Предположим, что мы ставим опыт по изучению пяти вариантов в пяти повторностях. Мы, следовательно, имеем 25 делянок, причем если участок достаточно обширен, то естественноисторические условия (почва, рельеф, экспозиция и т. д.) могут быть существенно отличными. В нашем опыте, наряду с различиями, вносимыми испытуемыми нами вариантами, имеется очень много различий, неизбежно сопутствующих каждому опыту: различие в почве разных делянок, рельефа, неизбежность несколько разновременного посева, индивидуальность работников, проводящих опыт, и т. д. Все эти различия в неправильно поставленном опыте могут сыграть вредную роль, исказив или смазав результаты. Если, например, все деревья одного варианта будут помещаться рядом и, следовательно, будут подвергаться воздействиям сходных условий, а другие в ином месте (подвергаясь воздействию других условий), то различие, которое мы припишем и различию вариантов, на самом деле, может быть, было вызвано различием естест-венноисторических условий опытного участка. Это ошибочная постановка опыта, к сожалению, отнюдь не изжитая в нашем опытном деле.

Следующая, лучшая, но все же далеко не удовлетворительная постановка будет заключаться в том, что, наметив 25 делянок (деревьев, индивидуумов и т. д.), мы наши 5 вариантов в пяти повторностях разбросаем между ними по жребию. В этом случае возможность смешения влияния естественноисторических условие с влиянием вариантов опыта будет уже устранена, но будет введена другая погрешность, которая, правда, не приведет к искажению результатов опыта (к тому, например, чтобы лучший вариант был принят за худший или чтобы было принято за доказанное различие между вариантами, не имеющими такового), но может смазать результаты (в том смысле, что отчетливые при правильной организации опыта результаты окажутся неотчетливыми). Это объясняется тем, что при такой постановке вся изменчивость полученного материала распределяется только по двум категориям, а именно: 1) изменчивость, соответствующая влиянию наших вариантов опыта (так как мы имеем пять вариантов, то она будет основана на 5—1, или на четырех степенях свободы), и 2) вся остальная изменчивость (20 степеней свободы), которую мы рассматриваем как случайную и сравнение с которой изменчивости вариантов служит для суждения о степени надежности наших выводов. Но эта изменчивость вызывается не только многочисленными мелкими неучитываемыми влияниями (что, собственно, и характеризует случайную изменчивость), но и рядом влияний, подлежащих учету (почвы, рельефа и т. д.). Влияние этих факторов очень часто весьма значительно, и если изменчивость, вызванную этими факторами, мы объединяем с изменчивостью чисто случайной, то мы значительно увеличиваем размеры случайной ошибки и тем самым наши результаты теряют в своей отчетливости.

Поэтому следует всегда стремиться при работе распределять повторности так, чтобы в пределах повторности подобрать возможно однородные делянки. Полный набор вариантов в одной повторности и образует то, что называется блоком, и рандомизация проводится лишь в пределах блока (для каждого блока отдельно), почему этот метод организации опыта и называется методом рандомизированных блоков и является основным методом в полевой практике. Отдельные блоки между собой могут очень сильно отличаться (и, как будет показано в главе о рандомизированных блоках, даже следует стремиться к тому, чтобы они сильно отличались), но внесенная этими отличиями изменчивость уже не сможет смазать наши результаты при данной системе, так как эта изменчивость (при взятом нами примере 5 повторностей она будет соответствовать 5—1, или четырем степеням свободы) будет выделена в особую категорию, а изменчивость чисто случайного характера, служащая мерилом опыта, уже будет определена по оставшимся 16 степеням свободы. Рандомизация, таким образом,. будет несколько сужена, поскольку часть отличий (между блоками) носит не случайный, а систематический характер.

Еще большее сужение рандомизации достигается в методе так называемого латинского квадрата. Он заключается в том, что полные наборы вариантов образуются не по одному направлению, а по двум, независимым друг от друга, в силу чего (для взятого нами примера) еще 4 степени свободы отходят к категории систематических различий и ошибка определяется уже по 12 степеням»

Таким образом, общим правилом, которым следует руководствоваться при рандомизации, следует считать такое: везде, где мы можем уловить существенные различия в объектах нашего опыта, могущие оказать влияние на результаты опыта, помимо влияния исследуемых вариантов (например, сорт деревьев, возраст, неодинаковое плодородие разных участков поля или сада) следует разделять наш опытный участок на блоки так, чтобы в пределах каждого блока были по возможности мало отличающиеся объекты — в таком разделении на блоки рандомизация полностью отсутствует. Но когда мы подобрали блоки и в пределах каждого блока наметили число по возможности сходных объектов, равное числу намеченных вариантов, то варианты размещаются по отобранным объектам жребием (способы жеребьевки могут быть, конечно, различны, но необходимо, чтобы они были совершенно механическими, исключающими возможность избирательности наблюдателя или какой-либо систематичности). Этот процесс вполне соответствует тому общему правилу, что метод размещения должен соответствовать характеру изучаемой нами изменчивости: неодинаковость объектов в пределах одного блока является источником случайной ошибки, и, следовательно, в размещении испытуемых вариантов по нашим объектам сознательно должен быть введен элемент случайности.

Многим кажется, что строгое требование рандомизации (и при том механической) является чрезмерным педантизмом и что несоблюдение этого требования (размещение объектов на глаз или даже систематическое распределение вариантов в пределах повторности) не может иметь серьезного значения. Поэтому я покажу на примерах, к каким грубейшим ошибкам иногда приводит игнорирование этого требования.

Первый пример возьму из сборника «Физиология больного и поврежденного растения» под редакцией В. Н. Любименко (Л., 1933), причем план работ был намечен тем же В. Н. Любименко. Таким образом, не может быть и речи об отсутствии квалифицированного руководства. Для разбора возьму некоторые данные из работы 3. М. Эйдельман, причем могу отметить, что весь сборник полон методических погрешностей, подобных разбираемой.

Задачей работы 3. М. Эйдельман было определить влияние подрезки листьев на урожай яровой пшеницы. Были проведены три серии основных опытов с удалением 25, 33, 50, 75 и 100% общей поверхности листьев, причем при первой серии удаление происходило в начале кущения (момент появления 3-го листа), при второй — в стадии трубки и при третьей — в начале цветения. Для каждой степени было оперировано по 50 растений. Опыт был поставлен, таким образом, без настоящей повторности, и сама техника опыта вызывает очень серьезные возражения, о которых я здесь распространяться не буду. Разберу просто окончательные данные по урожаю зерна (табл. 4).

Очень часто исследователи ограничиваются приведением процентных данных по отношению к контролю, как это приведено во второй половине таблицы. Результаты имеют «приглаженный» характер: во всех сериях урожай падает по мере увеличения повреждений, и, кроме того, с запаздыванием повреждения влияние его ослабляется. Однако знакомство с литературой и данными других исследователей (в том числе самого В. Н. Любименко в его прежних работах) приводит к недоумению, почему в данном случае влияние повреждения оказалось столь большим. Отчасти что чрезмерное влияние объясняется своеобразной техникой опытов, но в известной степени объясняется неправильным расположением вариантов опыта. Настоящей повторности, как я уже указывал, здесь нет, но для контроля повторность налицо (так как во всех трех сериях контроль по обработке одинаков) и рассмотрение данных контроля позволяет навести серьезную критику на выводы автора. В самом деле, мы видим, что урожай контроля возрастает от I серии к III, и при том очень значительно (почти вдвое): в работе приведена элементарная биометрическая обработка и разница I и III серий оказывается равной 1.41+0.15 г, т. е. статистически вполне достоверна. Это различие серий не предусмотрено опытом: оно, очевидно, явилось следствием неоднородности участка. Различие в сериях столь велико, что урожай растений почти всей третьей серии (за исключением поврежденных на 100%) выше контроля I серии.

Таблица 4

Результаты опытов по влиянию подрезки листьев пшеницы на урожай (данные 3. М. Эйдельман)

Серия	Контроль	Степень повреждения				
		25%	33%	50%	75%	100%
Вес (г) зерна на одно растение						
I серия	1.49	0.96	—	0.42	0.06	0.02
II »	2.34	1.43	1.42	1.16	0.67	0.43
III »	2.91	2.77	2.23	2.01	1.52	1.18
Вес зерна в процентах от контроля						
I серия	100	64.4	—	28.2	4.0	1.3
II »	100	61.3	60.7	49.6	28.6	18.4
III »	100	95.3	76.6	69.0	52.2	40.6

Расположение вариантов в пределах каждой серии было не рандомизированным, а систематическим, т. е. ближайшие степени повреждения располагались по соседству. В работе (как это, к сожалению, принято в литературе) не дано указаний на расположение отдельных серий, но если, положим (что по ряду соображений очень вероятно), II серия шла за III, I за II, то падение урожая при постепенном переходе от контроля к 100%-ной делянке объясняется не только влиянием различной степени обрезки, но и постепенным падением, естественно, плодородия данного участка, вполне доказанного из сравнения различных контролей. Правильно поставленный опыт заключался бы помимо, конечно, введения настоящих повторностей в том, что в пределах каждого блока (набора всех вариантов) варианты размещались бы не в систематическом порядке, а по жребию. Тогда результаты оказались бы, вероятно, не столь «гладкими». Этот пример является хорошим предостережением, что гладкость результатов и их полная согласованность между собой (что мы видим в данном примере для каждой серии, взятой порознь) являются серьезным доводом в пользу надежности выводов лишь в том случае, если были приняты гарантии для устранения систематических ошибок, а такие гарантии заключаются в правильно проведенной рандомизации.

Другой пример я возьму из неопубликованных материалов, любезно предоставленных В. И. Талицким. Дело идет об исследованиях по вредоносности кукурузного мотылька, проведенных на Аджаметской станции в Грузии в 1933 г.

Определение вредоносности производилось таким образом, что на площади в 2500 м^2 растения кукурузы сорта «имеретинский гибрид» на одной трети (800 м^2) систематически очищались от яичек кукурузного мотылька, а на остальных 1700 м^2 не очищались. Затем на очищенной площади было выделено 50 растений (неповрежденных), которые тщательно измерялись по ряду признаков, а на неочищенной было выделено 200 поврежденных растений. При этом оказалось, что средний вес зерен с неповрежденных растений равен 141.7±5.5 г, а средний вес зерен с поврежденных — 72.0±2.3; разница 69.7±6.0 г, или 49.2%±3.3%, абсолютно надежна. Разница, очевидно, не может быть объяснена избирательностью кукурузного мотылька, так как неповрежденные были тоже заражены, но потом очищены, не говоря уже о том, что резкая отрицательная избирательность кукурузного мотылька никем не наблюдалась.

Все прочие условия (почва, срок посева, сорт и т. д.) были одинаковы; отчет о работе по тщательности и документации производит самое благоприятное впечатление, так что с точки зрения обычных взглядов на методику постановки полевого опыта вывод о сильной вредоносности кукурузного мотылька (около 50%, т. е. снижение урожая вдвое) должен считаться доказанным. Но этот вывод противоречит всему прежнему опыту о вредоносности кукурузного мотылька. Развитие кукурузы было исключительно мощное (высота стеблей — около двух метров, стебли очень толстые, поломов не наблюдалось), зараженность невысока (1—3 гусеницы на зараженный стебель), и при таком вегетативном развитии, и таком заражении совершенно невозможно ожидать такой высокой вредоносности. Использование других признаков указало на источники ошибки. По высоте неповрежденные стебли оказались выше поврежденных (212.9 и 180.0 см в среднем), но по числу междоузлии получилось обратное: у неповрежденных растений оказалось в среднем 11.76±0.21 междоузлие на растение, а у поврежденных — 13.30±0.088, разность поврежденных и неповрежденных составляет 1.64±0.23 междоузлия, т. е. тоже абсолютно надежна. Так как заражение происходит в такой момент, когда растение в основном уже сформировалось, то никакое самое сильное заражение не может повлиять на число междоузлии, не говоря уже о том, что мы никак не можем представить, чтобы заражение вызвало увеличение числа междоузлии.

Чем же объясняются эти отличия? Как явствует из описания работы, для изучения неповрежденных растений было выбрано 50 штук с площади в 800 м^2, а для изучения пораженных — 200 с площади 1700 м^2, т. е. в обоих случаях бралась небольшая часть из общей совокупности. Способ выбора модельных растений не указан, но, очевидно, рандомизации не было, и хотя мы не имеем никаких оснований подозревать автора в стремлении к искажению результатов, но вполне естественно, что с пораженного участка отбирались средние растения, а с непораженного такие, которые во всех отношениях производили впечатление здоровых растений. Такими были энергично растущие растения, которые и оказались имеющими более высокий рост и меньшее число междоузлии и в силу большей общей энергии роста принесшими больший урожай.

Сравнение зараженных растений по степеням поражения (было использовано два признака — срок заражения и балл зараженности в ножку початка, как имеющего наиболее серьезное значение) показало совершенно неощутимую вредоносность (что и следовало ожидать от растений подобной мощности и невысокого заражения, как это было в данном случае), и потому можно с полной уверенностью сказать, что если бы отбор растений в качестве модельных был произведен по жребию, то никакой существенной разницы в урожае между поврежденными и неповрежденными растениями не оказалось бы. Доказанное статистически различие в весе урожая поврежденных и неповрежденных растений продолжает быть доказанным, но только приходится изменить самый объект доказательства: доказанной является не вредоносность кукурузного мотылька (конечно, в данном случае; при других условиях заражения вредоносность кукурузного мотылька может быть очень большой), а избирательность наблюдателя.

Практическим выводом является: в работах, подобных разобранной (я здесь не говорю уже об отсутствии настоящей повторности, поскольку все зараженные растения доставляли один участок, все незараженные — другой, в данном случае эта погрешность или не имела значения, или имела ничтожное значение по сравнению с основной ошибкой — избирательностью наблюдателя), необходимо по проведении искусственного заражения либо искусственного очищения определенного участка рядом с контролем или производить сплошное обследование всех зараженных и незараженных растений, или же делать выбор из всей совокупности по жребию или по какой-либо системе, исключающей избирательность наблюдателя (например, брать каждое десятое растение).

Разумеется, и рандомизация не спасет от систематических ошибок там, где варианты нашего исследования распределяются не по нашему произволу. Например, при изучении вредоносности хлебного пильщика, если мы отберем сначала здоровые и пораженные стебли, а затем из обоих партий по жребию отберем определенное число стеблей для исследования, то это нас не спасет от ошибочного вывода, потому что в силу избирательности пильщика вся совокупность зараженных стеблей, как правило, относится к более мощным стеблям, и эту избирательность (имевшую место до начала исследования) никакая рандомизация устранить не может. Для устранения маскирующего влияния избирательности здесь необходимо применить использование признаков (путем построения линий регрессии), независимых от пильщика, т. е. не изменяющихся под влиянием заражения.

Все это показывает, что при наличии достаточно сложного исследования (когда используются многие признаки, а не один) отсутствие рандомизации (даже там, где она применима) может быть исправлено использованием сопутствующих признаков, и этот метод (использование сопутствующих признаков) является единственно возможным там, где рандомизация в полной мере неприменима по существу дела и где имеет место исследование путем наблюдения, а не специально поставленного опыта. Но для того, чтобы результаты опыта, дающего в конечном счете только один признак (например, урожай с гектара), могли претендовать на убедительность, совершенно необходимо, чтобы в постановке опыта присутствовал элемент рандомизации: мы выбираем для обоих вариантов опыта два по возможности сходных участка и затем жребием решаем, какой участок должен быть отведен для того и другого варианта. Если этого не сделано, то , мы не гарантированы от того, что под один вариант систематически (т. е. во всех или большинстве точек, где проводится опыт) будет отведен несколько лучший участок и посев будет произведен лучшими семенами и т. д. Если это имеет место, то никакое увеличение числа повторностей нас не спасет от ошибки и разница в урожае, которую мы припишем испытываемому варианту, на самом деле будет относиться к избирательности организаторов опыта в отношении качества опытных участков.

Резюмируя, можно сказать, что принцип рандомизации при всей его простоте и пользе при правильном применении не является безусловно необходимым, но, с другой стороны, и не является достаточным, так как непродуманное его применение от ошибки не гарантирует. Но там, где он применим и применяется с полной сознательностью, он позволяет сильно стандартизировать опыты, вести работу сразу по сложной схеме, гарантирует от систематических ошибок, упрощает обработку и в целом вносит огромную экономию в исследование.

4.2. ОСНОВНЫЕ ПОНЯТИЯ ДИСПЕРСИОННОГО АНАЛИЗА

Дисперсионный анализ, разработанный Р. А. Фишером и его школой, в настоящее время получил широкое распространение преимущественно в Англии и Америке, но и в нашей литературе имеется ряд его изложений (перевод книги Р. А. Фишера о статистических методах, изложения Н. Ф. Деревицкого, Ю. Л. Поморского, В. И. Романовского), но все эти изложения или слишком кратки и схематичны, или недостаточно популярны, так как Теория дисперсионного анализа требует основательных математических познаний (см.: Романовский, 1947, с. 276). В силу этого возник ряд недоразумений в оценке этого метода.

Главной целью дисперсионного анализа является определение влияния разных условий на испытуемый признак или явление, это достигается путем разложения совокупной изменчивости (дисперсии, выраженной в сумме квадратов отклонений от общего среднего) на отдельные компоненты, вызванные влиянием различных источников изменчивости.

Мы можем различить три основные группы источников изменчивости: 1) систематическая изменчивость, вызываемая вариантами нашего опыта, т. е. то, что нас наиболее интересует; 2) систематическая изменчивость, вызванная колебанием условий опыта; 3) случайная изменчивость, т. е. та остаточная изменчивость, которая не включена в пункты (1) и (2) и вызвана вариантами опыта. Мы можем сделать заключение о степени надежности выводов и о степени влияния тех или иных вариантов опыта. Но и в пределах каждой из этих групп дисперсионный анализ позволяет значительно детализировать и углубить обработку материала, определяя степень надежности наших выводов в отношении каждой степени свободы.

Математическая теория дисперсионного анализа, как я указал, довольно сложна. Но для сознательного пользования этим методом достаточно твердо усвоить то основное положение, на котором весь этот метод базируется. Он основан на теореме аддитивности дисперсии (или вариансы, т. е. квадрата среднего квадратического отклонения). Из элементарных учебников вариационной статистики мы знаем, что квадрат средней ошибки суммы (или разности) двух изменчивых величин равен сумме квадратов ошибок обоих величин, взятых порознь при условии, что обе величины изменяются совершенно независимо друг от друга. То же имеет место и для трех, и большего числа независимых друг от друга переменных. Это общеизвестное правило и иллюстрирует принцип аддитивности дисперсии, так как квадрат средней ошибки есть та же дисперсия, но касающаяся не единичного наблюдения, а средней арифметической из ряда наблюдений, и этот принцип аддитивности и можно формулировать следующими словами: дисперсия, вызванная рядом независимых друг от друга источников изменчивости, равна сумме дисперсий, вызванной всеми источниками изменчивости, взятыми порознь.

Для того чтобы показать сущность дисперсионного анализа на простейшем примере, возьму числовые данные о сравнении двух сортов из книги П. Н. Константинова (1939) и, последовательно обрабатывая этот материал разными методами, покажу надежность каждого из них. Дело идет о сравнении двух сортов (обозначенных *x* и *y)* в шести точках: основной материал вместе с вычислениями отклонений, квадратов отклонений и произведений отклонений от соответствующих средних дан в табл. 5 (знак плюс опускается, ставится только знак минус).

Таблица 5

1	2	3	4	5	6	7	8	9
x	y	α_x	α_y	α_x^2	α_y^2	$\alpha_x\alpha_y$	α_{x-y}	$\alpha_{(x-y)}^2$
62	55	1	3	1	9	3	7	49
45	34	-16	-18	256	324	288	11	121
53	45	-8	-7	64	49	56	8	64
54	45	-7	-7	49	49	49	9	81
73	62	12	10	144	100	120	11	121
79	71	18	19	324	361	342	8	64
366	312	0	0	838	892	858	54	500

В первых столбцах (1 и 2) дан исходный цифровой материал, в остальных — вычисления для разных способов обработки.

1-й способ: определение разности средних. Средние величины будут: для сорта $x=\frac{366}{6}$ или 61, и для сорта *y* соответственно 312:6=52. Средняя ошибка сорта *x* равна $\sqrt{\frac{838}{5\cdot 6}}$ или $\sqrt{27.9333}=5.29$, и средняя ошибка сорта *y*= $\sqrt{\frac{892}{5\cdot 6}}$, или $\sqrt{29.7333}=5.459$. Средняя ошибка разности по широкоизвестной формуле будет равна: $\sqrt{27.9333+29.73333}$, или 7.59.

Таким образом, отношение разности сортов к своей средней ошибке будет равно 9:7.59, или 1.19. Такое *t* даже при большом числе дат, послуживших для определения, показывало бы, что разница между сортами не имеет никакого существенного значения. Здесь же мы имеем всего 12 дат, колеблющихся вокруг двух арифметических средних, т. е. всего 12—2 или 10 степеней свободы, где надо предъявлять повышенные требования к соотношению разности и ее средней ошибке.

Но правилен ли будет этот вывод? Очевидно, нет: простой взгляд на графу 8, где приведена разность между *x* и у, показывает, что она во всех шести случаях положительна, т. е. в том случае, если мы до опыта не ожидали превышения *x* над *y,* вероятность случайного возникновения серии из шести положительных разностей равна $1/2^5$, или 1/32, т. е. уже является серьезным указанием на значимость различия. Дело, как хороша известно, объясняется тем, что обычная формула ошибки разности имеет значение лишь в случае независимости обоих переменных; если же ее мы прилагаем к тому случаю, где переменные не независимы, то мы увеличиваем случайную ошибку за счет присоединения к ней всей изменчивости, вызванной колебаниями условий от точки к точке. Поэтому для того, чтобы найти «истинную» случайную ошибку, независимую от колебания условий в разных точках, мы должны эту изменчивость исключить. Это можно проделать разными способами.

2-й способ: определение средней разности. Простейшим из таких способов будет определение средней разности вместо разности средних. В качестве переменной мы берем не исходные цифры, а разности сортов для каждой точки. Получаем, следовательно, 6 дат вместо исходных 12, приведенные в графе 8. Средняя разность, конечно, оказывается равной 9, как и раньше (54:6), но ошибка этой средней разности будет значительно меньше. Как известно, средняя ошибка определяется так, что возводятся в квадрат все разности от некоторой условной средней и затем вводится поправка. Если каждое отклонение обозначить а, общее число дат — *n,* а разность между условным средним и истинным средним арифметическим — *b*, то средняя ошибка, как известно, равна

$$m=\sqrt{(\frac{\sum a^2}{n}-b^2)\cdot\frac{1}{n-1}}\text{, или } m=\sqrt{\frac{\sum a^2-nb^2}{n(n-1)}}$$

Так как условное среднее мы можем выбрать произвольно, то можно взять его и равным нулю, тогда *b* равно *M,* среднему арифметическому, и отклонения *a* равны абсолютным значениям дат, а *nb* равно Σ*x,* так как $M=\frac{\sum x}{n}$.

Для определения средней ошибки разности и возводим в столбце 9 все разности в квадрат, затем получаем среднюю ошибку разности по формуле

$$\sqrt{\frac{500-486}{6\cdot 5}}\text{, или } \sqrt{\frac{14}{30}}=0.683$$

Отношение средней разности к своей средней ошибке (9:0.683) оказывается уже равным 13.2, что уже указывает на чрезвычайно надежное различие между сортами, несмотря на малое число испытаний. Мы имеем всего 6 цифр (полученных из первоначальных 12), следовательно, пять степеней свободы; по таблицам Стьюдента мы видим, что для получения чрезвычайно высокой надежности (Р, вероятность случайного возникновения различия равна 0,001) при пяти степенях свободы достаточно *t*, равное 6.86. Следовательно, полученная нами величина дает, можно сказать, абсолютную гарантию надежности различия.

Но этот способ имея достоинства простоты, не лишен недостатков. В самом деле, мы видели, что при получении различий между сортами в шести точках, мы получили 6 цифр вместо исходных 12, и, следовательно, изменчивость оказалось соответствующей пяти степеням свободы, а не десяти. Такое снижение числа степеней свободы может сильно отразиться на надежности вывода, так как для получения высокой значимости (вероятность отсутствия существенной разницы равна 0.001) при десяти степенях свободы достаточно *t*, равное 4.59, а при пяти степенях свободы требуется 6.86.

3-й способ: использование корреляции между переменными. Этот способ, также как и первый, основан на определении разности средних, но с тем отличием, что ошибка разности определяется по формуле, приспособленной для зависимой изменчивости между переменными. Именно средняя ошибка разности равна

$$\sqrt{m_x^2+m_y^2-2rm_xm_y}$$

где *r* – коэффициент корреляции между обоими переменными. Для определения коэффициента корреляции введен столбец 7 в табл. 5, и мы получаем этот коэффициент равным $\frac{858}{\sqrt{838\cdot 892}}$, или 0.992. Введя этот коэффициент в формулу, получим ошибку разности, равную

$$\sqrt{27.9333+29.7333-2.0992\cdot 5.29\cdot 5.45}$$

или $\sqrt{0.4670}=0.6830$. Отношение 9:0.683 равно 13.2 и в точности совпадает с величиной 2-го способа, но имеет то преимущество, что соответствует девяти степеням свободы, а не пяти (одна степень свободы из исходных десяти использована для определения коэффициента корреляции). При наличии парных сопоставлений такой корреляционный метод является наиболее удобным и эффективным методом оценки разных вариантов, но, конечно, следует помнить, что и этот метод не является идеальным и что, например, при некоторых случаях криволинейной корреляции он окажется недостаточным для отделения случайной изменчивости от иных форм изменчивости, ее маскирующих. Главным ограничением корреляционного метода является то, что он приспособлен лишь для парных сопоставлений, между тем как дисперсионный анализ является свободным от этого ограничения и может применяться к любому числу сопоставлений.

Однако в целях наглядности применим дисперсионный анализ и к данному примеру.

4-й способ: дисперсионный анализ. Материал для дисперсионного анализа располагается в таком виде:

Сорт x	*Сорт y*	Сумма	Среднее	Сумма квадратов
62	55	117	58,5	6869
45	34	79	39,5	3181
53	45	98	49,0	4834
54	45	99	49,5	4941
73	62	135	67,5	9173
79	71	150	75,0	11282
Сумма 366	312	678		
Среднее 61	52		56.5	

Сумма квадратов				40280

Прежде всего определяем размер общей изменчивости, т. е. сумму квадратов отклонений всех дат от общей средней, равной 56,5. Это всего удобнее делать (при наличии счетной машины) таким образом, что условное среднее берется 0 и сначала возводятся в квадрат все исходные 42 дат. Так как при всех вычислениях неизбежно вкрадываются ошибки и так как здесь такие ошибки могут оказать особенно сильное влияние ввиду значительности поправки, то следует принять за безусловное правило, что все вычисления производятся два раза и не путем простой. проверки первоначально сделанного вычисления (при такой проверке хорошо известно, что сделанная раз ошибка не замечается' при проверке, даже многократной), а обязательно путем вычисления, независимого от первого. В данном случае полезно произвести суммирование квадратов сначала по строкам, потом по< столбцам, что и сделано с нашим материалом. Результат обоих вычислений совпал, что укрепляет уверенность в правильности вычисления. Полученная сумма квадратов 40280 является суммой квадратов расстояния от нуля, а нам надо взять сумму квадратов от арифметической средней, следовательно, надо внести поправку. Эта поправка по общему правилу равна nM^2 или, так как nM равно Σx (понимая под x все даты), то поправка равна $M\Sigma x$, или $\frac{\sum (x)^2}{n}$. Можно пользоваться обоими формулами для взаимного контроля, следовательно, мы получаем поправку, равную 678-56.5, или $\frac{678}{12}$,что равно 38307. Вычтя поправку 38307 из первоначальной грубой суммы квадратов, 40280, мы и получим искомую общую сумму квадратов (общую вариансу или дисперсию) — 1973.

Теперь эту общую дисперсию надо разложить на три категории: связанную с вариантами (здесь сортами), связанную с повторностями и случайную. Это разложение допустимо тогда, когда все категории изменчивости независимы друг от друга. Это соблюдено в данном случае, так как в каждой повторности представлены оба сорта без повторений. Для вычисления суммы квадратов, соответствующей вариантам (в данном случае — сортам), можно опять-таки сначала определить сумму квадратов от нуля, а затем из полученной суммы вычесть ту же поправку. Сумма квадратов от нуля определяется таким образом, что суммируются квадраты сумм обоих сортов и сумма делится на число дат, послуживших для определения каждой сортовой суммы. В данном случае: $\frac{366^2 + 312^2}{6} = 38550$, вычтя поправку 38307, получим сумму квадратов для сортов от общей средней 38550-38307, или 243.

Грубая сумма квадратов (от нуля) может быть получена также путем суммирования произведений по каждому сорту суммы дат и средней арифметической, и это вычисление является хорошей проверкой. В данном случае 366-61+312-52=38550.

Наконец, окончательную сумму квадратов 243 в данном случае очень просто получить непосредственно без поправки: берем разность суммы обоих сортов, возводим в квадрат, делим на общее число дат и получаем $\frac{(366-312)^2}{12} = 243$. Этот последний •способ получения квадрата разности, соответствующей определенной степени свободы, будет широко применяться в дальнейшем при детальном анализе результатов исследования, и он является вместе с тем хорошей проверкой проделанных вычислений. Точное совпадение обоих результатов ясно из того, что, обозначив сумму дат первого сорта через *а,* а второго через *в*, имеем для вычисления с поправкой $\frac{a^2+b^2}{6} - \frac{(a+b)^2}{12}$, что (если провести раскрытие скобок и приведение к одному знаменателю), очевидно, равно $\frac{(a-b)^2}{12}$, эта последняя формула указывает путь вычисления без поправки.

Для вычисления дисперсии, соответствующей повторностям, опять возводим в квадрат суммы всех дат, делим сумму на два (число дат в каждой повторности) и из полученного частного вычитаем ту же поправку, получаем

$$\frac{117^2 + 79^2 + 98^2 + 99^2 + 135^2 + 150^2}{2} - 38307$$

:или 40030-38307=1723. Таким образом, из общей изменчивости, общей суммы квадратов 1973 на долю вариантов приходится 243, на долю повторностей — 1723 и остаток — 7 приходится на долю случайной изменчивости, или ошибки. Ее можно вычислить и непосредственно, а не путем вычитания первых двух компонентов из общей суммы, но это обычно много более сложно и потому практикуется редко. Мы, таким образом, и осуществили первый этап дисперсионного анализа, или анализа варианcы, разложив общую сумму квадратов (общую дисперсию, или вариансу) по трем категориям изменчивости (табл. 6).

Таблица 6

Анализ варианcы

Категории изменчивости	Число степеней свободы	Сумма квадратов	Средний квадрат	θ
Варианты (сорта)	1	243	243	173,57
Повторности	5	1723	344,6	246,14
Ошибка	5	7	1,4	
Всего	11	1973		

Но сравнивать непосредственно эти части мы не можем, так как сортов всего два, а повторностей было шесть; известно же, что дисперсия возрастает пропорционально числу степеней свободы, т. е. числу независимых друг от друга направлений изменчивости. Сортов имеется два: так как мы берем изменчивость около среднего арифметического, то, очевидно, что у сортов имеется только одна степень свободы, так как после того, как один из сортов изменился в том или ином направлении, значение другого сорта определяется точно из значения первого и среднего арифметического для обоих сортов. Для повторностей мы имеем, очевидно, 5 степеней свободы (6—1), а всего имеем 11 степеней свободы (12—1), таким образом, на ошибку остается 5 степеней свободы Для сравнения разных категорий изменчивости мы и должны сумму квадратов данной категории изменчивости разделить на соответствующее число степеней свободы,. и получим средний квадрат, соответствующий данной категории изменчивости. Если, положим, исследуемые нами варианты (сорта) не представляют между собой существенных различий, то изменчивость, вызванная в нашем материале сортами, не будет существенно отличаться от чисто случайной изменчивости или изменчивости, связанной с ошибкой опыта: она может быть несколько меньше или несколько больше, но не превосходить ее во много раз. Вернее говоря, чем больше отношение дисперсии исследуемой нами категории изменчивости и дисперсии ошибки опыта, тем менее вероятно случайное возникновение такого отношения. Отношение это именуется θ (тета), и существуют таблицы Снедекора (приведены у Ю. Л. Поморского, В. И. Романовского, Н. Ф. Деревицкого), по которым можно судить, какова вероятность того, что при данном числе степеней свободы наблюдаемая нами величина θ могла возникнуть в силу чисто случайной изменчивости. В этих таблицах по горизонтальной верхней линии показано число степеней свободы большей вариансы, а по вертикали слева—число степеней свободы меньшей вариансы, и в клетке, находящейся на пересечении этих двух граф, мы находим три цифры: верхняя, наименьшая, показывает величину θ, вероятность случайного возникновения которой при данном числе степеней свободы равна 0.05, вторая — для вероятности 0.01 и третья — для вероятности 0.001. Обычно принимают, что если вероятность случайного возникновения различий больше 1/20, т. е. если наблюдаемые различия могли возникнуть в силу чисто случайных причин чаще, чем один раз на двадцать испытаний, то мы вправе считать полученный результат совершенна несущественным. Поэтому значение тета в верхней строке считается минимальным для того, чтобы признать наличие достаточно существенных указаний на наличие различий.

Следует отметить, что первоначальным критерием Р. А. Фишера была так называемая функция Z, которая является не чем иным, как половиной натурального логарифма θ: так как вычисление θ гораздо проще вычисления Z, то применение в находит сейчас все большее распространение.

В нашем случае мы имеем чрезвычайно высокие значения теты: 173.57 — для сортов и 246.14 — для повторностей. Если возьмем цифры θ для вероятности 0.001, то увидим, что при 1 и 5 степенях свободы достаточна тета, равная 4704, а при 5 и 5 степенях свободы — 29.75. Таким образом, вероятность случайного возникновения различий для обоих исследованных категорий изменчивости (рост и повторность) является исчезающе малой и надежность различий между повторностями доказывается даже сильнее, чем для сортов. Не следует, конечно, думать, что путем дисперсионного анализа мы получаем более надежное различие, чем путем примененных ранее способов (второй и третий): там использовалась функция t — отношение средних ошибок, а здесь функция θ — отношение квадратов средних ошибок, Для сортов, где имеется всего одна степень свободы, требуемая тета, в частности, равна квадрату t и мы, следовательно, получаем $t = \sqrt{173.57} = 13.2$, в точности совпадающее с величиной, полученной вторым и третьим способами. Таким образом, анализ вариансы дает в данном случае то же, что и второй способ, несколько уступая по точности третьему способу (поскольку там ошибка основана на 9, а не на 5 степенях свободы), но зато имеет перед третьим способом два преимущества: 1) дает возможность судить об изменчивости и по повторностям; 2) может применяться не только для парных сопоставлений, но и для таких, где испытывается одновременно большое число вариант, в этом .последнем случае преимущества анализа вариансы становятся совершенно очевидными.

Из таблицы 6 анализа вариансы можно показать и путь, соответствующий первому способу. В этом случае мы смешиваем две категории изменчивости по повторностям и ошибку, т. е. суммируем (согласно основной теореме аддитивности вариансы) 1723 и 7. Получаем 1730, соответствующую 10 степеням свободы или 173 на одну степень свободы (средний квадрат). Деля на 173 средний квадрат по сортам (243), получим θ=1,405 или $t = \sqrt{1.405} = 1.19$, опять-таки совпадающее с полученным первым способом.

Анализ вариансы не является, таким образом, каким-то математическим измышлением, придуманным для того, чтобы «вымучить» из материала выводы, недоступные другим методам это — обобщение и расширение известных методов на случаи, где прежние методы непосредственно не могли прилагаться.

Вместе с тем сразу выясняется и еще крупное достоинство .дисперсионного анализа — он позволяет маневрировать при определении размеров средней ошибки. В самом деле, изменчивость, связанная с повторностью, выделяется нами в особую категорию, потому что часто, как в данном случае, она смазывает различия испытываемых нами вариантов. Этим путем в данном случае мы огромную часть изменчивости, смешанную при вычислении по первому способу с ошибкой, выделяем в особую категорию, и хотя число степеней свободы тоже уменьшается, но выигрыш в надежности выводов получается чрезвычайно большой. Но если бы анализ вариансы показал, что средний квадрат, соответствующий повторностям, не больше или немногим больше, чем средний квадрат ошибки, то это означало бы, что нет существенной разнородности в пределах повторности, что эта изменчивость тоже по существу случайная и, следовательно, мы имеем право объединить обе эти категории: выигрыш будет заключаться в том, что объединенная изменчивость при том же или немногим большем среднем квадрате будет базироваться на большом числе степеней свободы.

Проделанный выше способ вычислений с возведением в квадрат оригинальных дат и с внесением потом значительной поправки наиболее удобен тогда, когда в нашем распоряжении имеются счетные машины (арифмометры или клавишно-счетная машина КСМ с электрическим приводом). Тогда большие размеры поправки нас не могут смущать, и все вычисления проходят очень быстро. Если мы работаем без машины, то следует избегать оперировать с многозначными числами. Но и в данном случае, как и всюду, вполне применим широко распространенный метод вычисления от условного среднего, возможно близкого к точному среднему; при таком вычислении все поправки оказываются минимальными, но необходимо проделать предварительную работу вычитания условной средней из всех дат (отчего наряду с положительными числами окажутся и отрицательные). На прежнем примере покажем, как это делается. Вычитаем из всех наших дат число 56 (так как среднее арифметическое равно 56.5) и получаем табл. 7.

Знак плюс при всех положительных числах опущен. Грубая сумма квадратов от условного среднего 56, как видно, равна 1976. Поправка равна $\frac{-6^2}{12}$ или 6-0,5, т. е. 3. Отсюда точная общая сумма квадратов равна, как и следовало ожидать, 1976—3, или 1973. Поправка, как видим, ничтожно мала по сравнению с грубой суммой квадратов. То же самое проделаем для сортов $\frac{30^2+24^2}{6}-3$ или 246-3=243.

Наконец, для повторностей

$$\frac{5^2+33^2+14^2+13^2+23^2+38^2}{2}-3=1726-3=1723.$$

Мы получили те же самые цифры, как и при первом вычислении, оперируя все время значительно меньшими числами.

Остается поставить вопросы: какие гарантии безошибочности нашего вывода о наличии существенности сортового различия к является ли исследование в такой постановке вполне удовлетворительным?

Таблица 7

Сорт *x*	Сорт *y*	Сумма	Среднее арифметическое	Сумма квадратов
6	-1	5	2,5	37
-11	-22	-33	-16,5	605
-3	-11	-14	-7,0	130
-2	-11	-13	-6,5	125
17	6	23	11,5	325
23	15	38	19,0	754
Сумма 30	-24	6		
Среднее арифметическое 5	4			
Сумма квадратов 988	988			1975

Полученный нами вывод (исчезающе малая вероятность отсутствия сортовых различий), а отсюда убеждение в наличии сортового различия сохраняют свою убедительность только в том случае, если в постановке опыта не было допущено смешения сортового различия с каким-либо иным: если, например, сорт *x* всюду сеялся раньше *y* или на лучших землях или каждый сорт высевался и обрабатывался особым работником и т. д. Тогда статистическая достоверность вывода касается всего комплекса различий и сплошь и рядом ведущую роль в доказанном статистическом различии играет вовсе не исследуемое нами различие, а иное, вкравшееся часто совсем незаметно. Гарантией от такого искажения вывода является последовательная рандомизация в пределах каждой повторности, о чем подробнее будет сказано в главе о рандомизации.

Что касается вопроса об удовлетворительности нашей постановки, то недостатком ее является то, что из сферы исследования выпадает вопрос взаимодействия сортов и окружающих условий. Такое взаимодействие общеизвестно; из двух сортов в одних условиях один сорт оказывается лучше, в других они меняются местами, и вообще нередки такие случаи, что опыт, поставленный в очень широких условиях, приводит как будто к выводу об отсутствии различий, так как в части повторностей один сорт превышает другой, в другой части имеет обратное явление и алгебраической суммой оказывается величина, близкая к нулю. А так как в данном случае наличие резкого отличия условий несомненно (это вытекает из чрезвычайно высокой значимости различий между повторностями), то фактор взаимодействия тоже представляет большой интерес. В данном случае, однако, мы его оценить не можем, так как взаимодействие сорта и повторности и составляет то, что мы называем изменчивостью, связанной с ошибкой опыта, которая сама служит стандартом, с которым сравниваются другие виды изменчивости. Так как эта величина очень мала, то в данном случае взаимодействие, видимо, отсутствует. Там же, где оно выражено, необходима более сложная организация опыта, известная под названием факториальной схемы, о чем речь будет в главе 4.7.

4.3. МЕТОД РАНДОМИЗИРОВАННЫХ БЛОКОВ

Только что разобранный пример представляет простейший случай так называемых блоков. Каждая повторность образует то, что называется блоком — полный набор вариантов опыта. В каждом блоке представлены все варианты в одинаковом количестве, обычно по одной делянке, но могут быть и парные, и вообще множественные делянки. В каждом блоке делянки (или вообще объекты опыта) подбираются возможно однородные, и до рандомизации мы имеем право делать какие угодно выключки. Но, наконец, мы подобрали в каждом блоке нужное нам число делянок или вообще объектов, после этого путем чисто механической рандомизации (жребием, использованием специальных таблиц и т. д.) размещаем наши варианты по отобранным объектам. Таким образом, основной принцип очень прост, но на некоторых деталях следует остановиться.

Прежде всего, по какому признаку образовывать блоки. Наиболее распространенный — чисто территориальный принцип (я имею в виду прежде всего агрономию, где впервые метод рандомизированных блоков и был применен Р.

Фишером (1937*а,в*). Опытный участок разбивают на несколько отделов, блоков, стараясь их формировать так, чтобы в пределах каждого блока была возможная однородность почвы, рельефа и других условий. Каждый блок разделяется на одинаковые делянки (с полным разрешением делать до опыта какие угодно выключки). Если мы этим путем достигнем цели и выберем блоки, внутренне однородные, вся межблоковая изменчивость будет отделена нами от чисто случайной и точность наших выводов сильно повысится. Но блоки можно формировать и по другому признаку. Например, опыты мы проводим в чрезвычайно разнородном (по возрасту, степени развития крон и т. д.) саду, где собраны в полном беспорядке разные сорта, положим, яблонь. Очень часто (но далеко не всегда, конечно) изменчивость по этим признакам гораздо сильнее, чем изменчивость по чисто территориальному признаку. В этом случае будет целесообразно образовывать яблоки по признаку сорта, возраста и т. д. В этом и заключается удача или неудача организации опыта, если мы образуем блоки JIO территориальному признаку. Но в пределах одного компактного блока объекты настолько разнородны, что невозможно выбрать сколько-нибудь однородные объекты, в силу этого изменчивость внутри блока окажется очень высокой, а она-то и служит мерилом случайной изменчивости.

Принцип рандомизации гарантирует нас от неверного вывода, но благодаря неудачной организации результат окажется смазанным и вывод ненадежным. Как будет показано дальше, во многих случаях, если нам неясно, какой фактор из окружающих условий оказывает особо сильное влияние на изменчивость (помимо, конечно, используемых нами различий), имеется возможность вести рандомизацию в двух и более направлениях — методами латинского и греко-латинского квадратов.

Второй вопрос заключается в том, как велики могут быть различия между блоками. Предположим, что ставятся опыты по влиянию различных удобрений на растения. Многие опытники склонны считать, что такие опыты мы должны ставить (во всех повторностях) в однородных, типичных для данного района условиях (в смысле почвы, обработки, метеорологических условий года и т. д.). Это стремление к типичности объясняется тем, что при обычной обработке результатов и работе в разных условиях, где изменчивость, вызванная условиями, не выделена в особую категорию, мы получаем такую высокую изменчивость для ошибки исследования, что результаты оказываются смазанными. Но от этой смазанности как раз и охраняет при правильном проведении метод рандомизированных блоков.

Однако работа в узких «типичных» условиях имеет тот недостаток, что выводы, сколь бы они ни были достоверны для данных условий, не могут быть приложены за пределами этих условий. И если, положим, «типичные» условия господствуют (что очень часто бывает) лишь на половине обслуживаемого района или еще меньше, то исследованный прием или вообще не может быть рекомендованным для «нетипичной» части обслуживаемого района (отграничить типичную часть от нетипичной, конечно, не «легко), или, если он будет рекомендован, приведет к совершенно неожиданным результатам.

Правильным выводом будет работа во всей амплитуде условий, доступных исследователю с выделением блоков так, чтобы каждый блок составлял возможно однородное целое. В этом случае, если мы получим вывод о преимуществе того или иного приема во всей серии блоков, то это даст нам право широко рекомендовать данный прием (по крайней мере в пределах изученных нами условий); если же мы существенного различия исследованных нами приемов не получим, то, конечно, не следует сразу принимать, что этих различий нет. Необходимо обратить внимание на возможность взаимодействия между изучаемым нами фактором и блоковыми условиями. Уже сопоставление различий между вариантами в разных блоках (в особенности расположенными по какому-либо возрастающему или убывающему признаку, например по влажности почвы участка, количеству осадков года и т. д.) может дать намек на существование таких взаимодействий. Последовательное изучение таких взаимодействий дается так называемым факториальным анализом, о чем речь будет впереди. Наличие взаимодействия (в простейших опытах) может быть доказано при наличии или повторности целой системы (дублирование всей схемы рандомизированных блоков) или же повторности в пределах одной системы блоков (парные блоки и т. д.).

В качестве примера применения метода рандомизированных блоков возьму данные по определению коэффициента полезного действия опылителей на тыкве, собранные А. Н. Невкрыто. Задачей работы являлось определить, какое число посещений одного цветка достаточно для полного опыления завязи (известно, что опыление может быть неполное, что повлечет лишь частич^ ное развитие семян). Наблюдения проводились путем допущения определенного числа посетителей (главным образом домашней пчелы) до цветка, после чего цветок изолировался. Для числа посетителей допускались варианты: 0 (полная изоляция), 5, 15, 45 и открытые совсем цветки, когда принималось 135 посетителей на основе данных о посещении вообще. Здесь, как и вообще часто применяется, применена не простая равномерная шкала для измерения вариантов, а шкала на основе геометрической прогрессии. Почему здесь неудобно применять простую равномерную шкалу, например, принять 0, 10, 20, 30 и т. д. посещений? Потому что если взять мелкие интервалы, что получится слишком большое число вариантов, трудно осуществимое при наличном количестве сил, так как одновременно можно наблюдать за одним только цветком. Если же взять крупные, например 0, 30, 60, 90, 120, то между первым и вторым вариантом (вероятно, наиболее интересный интервал, так как именно здесь можно ожидать нахождение достаточного количества опылителей) окажется слишком большой промежуток. Принятие геометрической прогрессии дает при небольшом числе вариантов узкие интервалы в начале, все возрастающие—в менее интересной части.

С другой стороны, геометрическая прогрессия для математической обработки более удобна, чем какой-либо произвольный ряд цифр, так как путем логарифмирования она превращается в арифметическую прогрессию. Результат для каждого цветка отмечался баллом: 0 — полное отсутствие завязности, 1 2 и 3 от сомнительного и слабого завязывания до полного завязывания. Наблюдение велось в двух местах: в степи и на пойме (село Матвеевка Полтавской обл.). Каждое наблюдение получено, как средний балл опыления для 2—4 цветков (обычно 3). Каждый день проводилось исследование всех пяти вариантов (тремя наблюдателями, чередовавшимися в отношении вариантов по жребию): распределение вариантов между выбранными для исследования цветками тоже велось по жребию. Блоком в данном исследовании являлся день, в который приводились все пять вариантов, так как, естественно, можно было ожидать (хотя это не подтвердилось), что коэффициент полезного действия будет меняться в зависимости от погоды и сезона. Работа проводилась в 1937 г.

Так как при полной изоляции во всех 15 случаях оказался полный нуль завязности, то после такого подтверждения известного уже взгляда о полной самостерильности тыквы в дальнейшем этот вариант не исследовался. В 52 же случаях допущения хотя бы одной пчелы (минимум к двум цветкам, так как редкие случаи отсутствия завязности у единичных цветков имелись) всегда наблюдался тот или иной балл завязности. Такой отчетливый результат, конечно, ни в какой математической обработке не нуждается.

Для поймы имеем шесть дней наблюдения (блоков), для степи — семь дней. И в том и другом случае общий период наблюдения одинаков (с конца июля до середины августа), но дни не совпадают, поэтому объединение материала с тем, чтобы производить сравнение в тот же день для степи и поймы, невозможно. -Но так как оказалось, что коэффициенты полезного действия в степи и пойме различны (может быть различна фауна опылителей, степень их активности, характер их работы — за пыльцой или нектаром и т. д.), то целесообразно произвести анализ для поймы и степи отдельно, а потом, в случае отсутствия существенных различий, их объединить. С точки зрения освоения техники вычислений такой путь удобен и тем, что можно познакомиться с техникой такого объединения первоначально изолированного материала, где вовсе не требуется все вычисления производить заново.

Проделаем поэтому сначала всю обработку по материалам для поймы (средняя завязность) (табл. 8).

Таблица 8

Варианты (число посетителей)

Дни	5	15	45	136	Сумма	Среднее
26-28.VII	2,0	2,0	3,0	3,0	10,0	2,500
31.VII	1,0	1,0	2,3	3,0	7,3	1,825
l.VIII	0,7	2,0	3,0	2,0	7,7	1,925
14.VIII	1,5	2,5	3,0	1,5	8,5	2,125
15.VIII	2,7	2,7	2,0	3,0	10,4	2,600
18.VIII	0,8	2,3	2,0	2,8	7,9	1,975
Сумма	8,7	12,5	15,3	15,3	51,8	
Среднее	1,4500	2,0833	2,580	2,500		2,158

Как видим, средний балл завязности для всего материала равен 2,158. Средние баллы по вариантам и дням сильно колеблются, и, естественно, возникает вопрос, в какой мере такие колебания являются существенными и могущими дать основание для тех или иных надежных выводов.

Для этого проделаем анализ дисперсии. Определяем общую сумму квадратов (возводим в квадрат все даты) 2,0; 2,0; 3,0 и т. д. (для проверки проделываем это действие два раза, один раз—по столбцам, другой—по строкам). Получаем грубую сумму квадратов от нуля до 124,88. Поправка равна $\frac{51.8^2}{24}$, или 51,8-2,15833=111,80167, разность 13,07833 и даст общую сумму квадратов. Сумма квадратов по дням получается:

$$\frac{10.0^2+7.3^2+7.7^2+8.5^2+10.4^2+7.9^2}{4}-111.80167$$

или (10,0-2,500+7,3-1,825+7,7-1,925+8,5-2,126+10,4-2,600+7,9-1,975)-111,80167=113,85-111,801167=2,04833.

Общий принцип такого вычисления заключается в следующем: сумма дат для каждого дня возвышается в квадрат, все квадраты суммируются и сумма делится на число дат, послуживших для образования каждой суммы по дням. Второй способ, дающий, конечно, совершенно тождественный результат к потому очень удобный для проверки вычисления, заключается в умножении суммы дат для каждого дня на соответствующую среднюю величину и в суммировании всех произведений. Из каждой такой суммы вычитается та же поправка, так как в обоих случаях надо найти изменчивость около общего арифметического среднего.

Вполне естествен вопрос о том, сколько знаков надо вычислять и можно ли при таком способе вычислений пользоваться логарифмической линейкой. Уже последний пример показывает,, что мы из грубой суммы 113,85 вычли поправку 111,80167; разность равна 2,04833. При таком вычитании мы потеряли две первые значащие цифры (значащей цифрой называется цифра, отличная от нуля, независимо от ее положения, например, число 0,000026 имеет две значащие цифры, хотя общее число цифр семь) и если бы мы вычисляли с точностью до трех знаков, то у нас только первая цифра разности оказалась бы определенной точно. Логарифмическая линейка (обычных размеров) дает уже третью цифру с погрешностью. А так как при изложенном способе вычислений (грубая сумма квадратов берется от нуля) поправка, как правило, очень велика, то мы можем получить даже первые цифры ненадежными. В конечных результатах (средние квадраты ошибки) достаточно трех значащих цифр, определенных точно, но, имея в виду потерю цифр при поправках, желательно вести вычисление с точностью до 5—6 цифр. К этому присоединяется еще одно обстоятельство: разное способы вычисления суммы квадратов дают прекрасный контроль за вычислениями, и, конечно, контроль тем надежнее, чем большее число цифр показывает совпадение при независимых вычислениях. Поэтому целесообразно вести вычисления с запасом, используя всю или почти всю ширину интервала счетной машины, тем более что такое увеличение числа цифр почти не влияет на скорость вычислений. Разумеется, при отсутствии счетной машины следует прибегать к вычислениям не от нуля, а от другого условно среднего, близкого к среднему арифметическому, как показано в предыдущем примере.

Проделав совершенно аналогичное вычисление для вариантов опыта, получим

$$\frac{8.7^2+12.5^2+15.3^2+15.3^2}{6}-111.80167=4.88500$$

Сумма квадратов для ошибки получается вычитанием из общей суммы квадратов сумм для дней и для вариантов. Мы получаем следующий анализ варианcы (табл. 9).

Таблица 9

Категории изменчивости	Число степеней свободы	Сумма квадратов	Средний квадрат	θ
Варианты	3	4,88500	1,62833	3,97473
Повторности (дни)	5	2,04833	0,40967	
Ошибка	15	6,14500	0,40967	
Всего	23	13,07833		

Как уже было указано раньше, число степеней свободы на единицу меньше числа соответствующих дат: будут ли это исходные даты или средние по повторностям и по вариантам. Число степеней свободы для ошибки (которая, как увидим дальше, может рассматриваться как взаимодействие повторностей и вариантов) при методе рандомизированных блоков всегда равно произведению чисел степеней свободы вариантов и повторностей, а также, конечно, равно разности между общим числом степеней свободы и суммой степеней свободы вариантов и повторностей. Это ясно из простого алгебраического равенства. Если число вариантов т, а число повторностей n, то общее число дат будет mn, а общее число степеней свободы будет $mn—1$.

Тогда число степеней свободы для ошибки будет

$$mn - 1 - (m - 1) - (n - 1) = mn - m - n + 1 = (m - 1) - (n - 1).$$

При произведенном нами анализе в силу редкого случайного совпадения средний квадрат, соответствующий дням (повторностям), оказался в точности равен среднему квадрату, соответствующему ошибке, следовательно, различие между днями не вносит никакой изменчивости сверх случайной в изучаемый нами признак. Поэтому изменчивость, связанную с повторностью, мы можем присоединить к «ошибке» и получить (в силу имевшего» место в данном случае полного совпадения средних квадратов) ту же величину среднего квадрата ошибки, но основанную не на 15, а на 20 степенях свободы. Отношение среднего квадрата для вариантов к среднему квадрату ошибки (тета) равно 3,97473, и при данном числе степеней свободы (3 степени свободы большей варианcы и 20 для меньшей) такой размер теты соответствует вероятности отсутствия различий, лежащих между 0.05 и 0.01,. т. е. мы можем считать различие между вариантами, установленными достаточно четко.

Теперь возникает вопрос о выяснении существенности различия между отдельными вариантами. Совершенно очевидно, что между вариантами с 45 и 135 посещений мы никакого различия ни при какой обработке найти не можем, так как там средние абсолютно тождественны, но неясно, одинаково ли существенно различие между вариантами 5 и 15, с одной стороны, и 15 и 45 — с другой. Это можно проделать путем сравнения соответствующих средних и деления разности на ошибку разности. Но, в данном случае нам важно не сравнение отдельных средних, а выяснение вопроса, является ли существенным увеличение завязности при увеличении числа посетителей.

Наилучшим методом для решения таких вопросов, которые могут быть сформулированы еще до исследования (а в данном случае, очевидно, мы можем только ожидать повышения завязности от увеличения числа посетителей, но ожидать ее падения не можем), является разложение суммы квадратов, соответствующих нашим вариантам по отдельным степеням свободы. Для каждой степени свободы вычисляется разность, соответствующая определенному противопоставлению, разность возводится в квадрат и делится на определенным способом вычисленный делитель: сумма квадратов для всех степеней свободы при правильном разложении должна в точности совпадать с суммой квадратов для вариантов, вычисленных ранее. В этом и заключаются преимущества такого разложения: 1) получается ответ на систему вопросов, поставленных заранее, 2) достигается проверка.

Такое разложение является обобщением вычисления разницы между вариантами при наличии двух вариантов, что было показано выше. Но мы, конечно, можем сравнивать не только два варианта, а какое угодно число. Например, мы должны выяснить различие в завязности между первым вариантом (пять опылителей) и остальными тремя. Для этого мы должны из средней арифметической для трех высших вариантов вычесть сумму, соответствующую первому варианту, т. е. проделать такое вычисление:

$$\frac{12.5+15.3+15.3}{3}-8.7, \text{ или } \frac{12.5+15.3+15.3-3\cdot 8.7}{3}$$

Чтобы получить разность, соответствующую данному контрасту, мы, следовательно, должны поставить такие коэффициенты перед соответствующими суммами (1, 1, 1 и 3), чтобы они в сумме были равны нулю. Другое требование, предъявляемое к коэффициентам, заключается в ортогональности (или независимости) наборов коэффициентов для всех степеней свободы. Оно гласит: сумма попарных произведений соответствующих коэффициентов .любых двух степеней свободы должна быть тоже равна нулю. Более подробно приемы образования ортогональных коэффициентов будут показаны в отдельной главе, здесь же я ограничусь приведением таблицы разложения по трем степеням свободы с указанием, что принятое нами разложение (одно из бесчисленного числа возможных) удовлетворяет требованиям ортогональности.

В самом деле, если мы перемножим коэффициенты для первой и второй степени свободы, то получим

-3,0+(+1, -2)+(+1+1)+(+1+1)=0,

точно так же и при других двух возможных комбинациях.

Получаем следующее разложение по трем степеням свободы (табл. 10).

Таблица 10

Степень свободы	Коэффициенты для сумм вариантов				Делитель Δ	Разность δ	$\frac{\delta^2}{\Delta}$	θ
	5	15	45	195				
I	-3	1	1	1	72	17,0	4,01389	9,7979
II	0	-2	1	1	36	5,6	0,87111	2,1264
III	0	0	-1	1	12	0,0	0,0000	0
Суммы вариантов	8,7	12,5	15,3	15,3	Сумма		4,88500	11,9243

Разности (δ) вычисляются таким образом: сумма, соответствующая каждому варианту, умножается на соответствующий коэффициент и все произведения складываются (принимая, конечно, во внимание знак соответствующего коэффициента), например, 8,7-(-3)+12,5-1+15,3-1+15,3-1=17,0. Эти разности возводятся в квадрат и делятся на делитель (Δ), образованный таким способом: сумма квадратов всех коэффициентов умножается на число дат, послуживших для определения каждой суммы, в данном случае 6, поскольку каждый вариант основан на шести днях наблюдения. Получаем, например, для первой степени свободы:

$$(3^2+1^2+1^2+1^2)\text{-}6=72$$

Сумма квадратов разностей $(\frac{\delta^2}{\Delta})$ соответствующих каждой степени свободы, в сумме должна дать ранее вычисленную сумму для вариантов — 4,88500; такое совпадение подтверждает правильность вычислений. Тета определяется делением квадрата разности на средний квадрат разности для ошибки (0,40967). Определение тета также легко проверяется, так как сумма тета, определенных для каждой степени свободы отдельно, должна равняться ранее определенной тета, умноженной на число степеней свободы, т. е. 3,97473-3, или 11,9242, — разница, как видим,. лишь в последнем знаке.

Беря опять таблицу значений тета для разных уровней значимости (т. е. для разных значений вероятности случайного возникновения наблюдаемых различий), мы должны уже пользоваться первым столбцом слева, т. е. тем столбцом, вверху которого стоит 1 (число степеней свободы большей вариансы), так как наши квадраты разности (вариансы) вычислены для каждой степени свободы порознь. Отыскивая строку, соответствующую* 20 степеням свободы, находим там, как всегда, три цифры: 4,35, 8,10 и 14,82, соответствующие *P* (вероятности случайного возникновения такой вариансы) 0,05, 0,01 и 0,001. Тета для первой степени свободы (9,7979) больше 8,10. Следовательно, вероятность случайного возникновения такой разности при отсутствии реального различия меньше 0,01. Иначе говоря, мы можем считать достаточно точно установленным, что пяти посещений безусловно недостаточно для полного опыления завязи. Что касается второй степени свободы (противопоставления 15 посещений большему числу), то увеличение завязности здесь при повышении числа посещений не является установленным, тета 2,1264 далеко не достигает даже низшего уровня значимости. Следовательно, возможно, что вполне достаточное число посетителей находится очень близко к 15.

Таблица 11

Варианты (число посетителей)

Дни	5	15	45	135	Сумма	Среднее
24.VII	1,5	1,5	3,0	3,0	9,0	2,250
26—27.VII	2,0	2,0	3,0	3,0	10,0	2,500
29.VII	3,0	3,0	3,0	2,0	11,0	2,750
30.VII	1,3	3,0	3,0	3,0	10,3	2,575
2.VIII	1,5	1,5	3,0	2,5	8,5	2,125
13.VIII	2,3	0,7	2,0	2,0	7,0	1,750
17.VIII	1,7	2,7	2,3	3,0	9,7	2,425
Сумма	13.3	14.4	19.3	18.5	65.5	
				Общее среднее *M*		2,33929

Перейдем теперь к обработке подобных же данных по степному участку, исходный материал представляется в виде табл. 11 (и здесь, как и в первом случае, материал двух дней объединен).

Проделывая вычисления совершенно так же, как в первом случае, получаем:

Общая сумма квадратов от нуля	166,19
Поправка на *M*	153,22321

Общая сумма квадратов от *M*	12,96679
Сумма квадратов для вариантов от нуля	156,99857
Поправка на *M*	153,22321

Сумма квадратов для вариантов от *M*	3,77536
Сумма квадратов для дней от нуля	155,8575
Поправка на *M*	153,22321

Сумма квадратов для дней от *M*	2,63429

Получаем анализ вариансы (табл. 12).

Таблица 12

Категории изменчивости	Число степеней свободы	Сумма квадратов	Средний квадрат	θ
Варианты	3	3,77536	1,258453	3,45457
Дни	6	2,63429	0,439048	
Ошибка	18	6,55714	0,364286	
Всего	27	12,96679		

И здесь изменчивость по дням лишь немного превосходит изменчивость для ошибки, поэтому и здесь можно было бы объединить обе категории изменчивости, но в данном случае мы не получим никакой выгоды, так как хотя число степеней свободы увеличится с 18 до 24, но и средний квадрат также увеличится.

И здесь тета для вариантов указывает на вероятность отсутствия существенной разницы, меньше 0,05 '(очень близко к тому, что видели в первом случае). Разложение суммы квадратов для вариантов представлено в табл. 13.

Вывод получается сходный с выводом по пойме. Совершенна очевидно, во-первых, что никакого значения увеличение числа посещений с 45 до 135 не имеет (имеется даже небольшое, возможно случайное, уменьшение числа завязности при переходе от 45 посетителей к 135). Пяти посещений, очевидно, мало, но в отличие от данных первого анализа и 15 посещений дают существенно меньший эффект по сравнению с большим числом их. Так как, за исключением этого отличия, мы имеем полную согласованность данных между поймой и степью и это отличие указывает не на существенное разногласие, а лишь на большую определенность результатов по степи и так как средние величины в обеих стадиях практически тождественны (общие средние 2,339 и 2,158, т. е. разность равна 0,181 при средней ошибке разности около 0,2), то оба материала мы можем объединить. Подобные объединения часто приходится производить и целесообразно показать технику такого объединения.

Таблица 13

Степень свободы	Коэффициенты для сумм по вариантам				Делитель Δ	Разность δ	$\frac{\delta^2}{\Delta}$	θ
	5	15	45	135				
I	-3	1	1	1	84	-12,3	1,80107	4,944
II	0	-2	1	1	42	9,0	1,92857	5,294
III	0	0	-1	1	14	-0,8	0,04571	0,125
Сумма по вариантам	13,3	14,4	19,3	18,5		Сумма	3,77535	10,36

При объединении материала не нужно производить заново самую трудоемкую часть вычисления — суммирование квадратов -всех дат. Мы поступаем следующим образом. Для определения общей суммы дат складываем суммы дат по пойме и степи: 51,8 и 65,5, получаем 117,3, что соответствует 52 датам (24 даты первой стации и 28 второй).

Поправка от общей средней $(\frac{117.3}{52} = 2.255769)$ будет равна $\frac{117.3^2}{52} = 117.3 \cdot 2.25576923 = 264.60173 \cdot$

Для вычисления общей суммы квадратов отклонений от нуля мы складываем вычисленные ранее суммы квадратов от нуля (т. е. просто абсолютные значения) 124,88 и 166,19 и вычитаем поправку. Получаем

сумма квадратов от нуля — 291,07
поправка —264,60173
общая сумма от *M* — 26,46827

Для дней мы опять-таки складываем суммы квадратов от нуля и вычитаем поправку:

113,85 + 155,8575 — 269,7075
поправка —264,60173
сумма квадратов для дней от *M* —5,10572.

Для вариантов, конечно, таким образом поступать нельзя, так как новые суммы для вариантов основаны каждая на 13 днях и не могут быть получены простым суммированием прежних сумм. Поэтому для вариантов мы сначала производим суммирование .прежних сумм:

Варианты:	5	15	45	135	Всего
Пойма	8,7	12,5	15,3	15,3	51,8
Степь	13,3	14,4	19,3	18,5	65,5
Всего	22,0	26,9	34,6	33,8	117,3

Эти новые суммы мы и возводим в квадрат, суммируем, делим на 13 (поскольку каждая сумма основана на 13 датах) и вычитаем ту же поправку:

сумма квадратов от нуля —272,86231
поправка — 264,60173
сумма квадратов от *M* — 8,26058

Получаем анализ варианса для объединенного материала (табл. 14).

Таблица 14

Категории изменчивости	Число степеней свободы	Сумма квадратов	Средний квадрат	θ
Варианты	3	8,26058	2,75353	7,5658

Дни	12	5,10572	0,42548	1,1691
Ошибка	36	13,10197	0,36394	
Всего	51	26,46827		

Разложение по степеням свободы для вариантов (по той же схеме, как и раньше, мы опускаем распределение коэффициентов. так как оно повторяет прежние данные) дает следующее-(табл. 15).

Таблица 15

Степень свободы	Разность δ	Делитель Δ	$\frac{\delta^2}{\Delta}$	θ
I	29,9	156	5,50314	15,1209
II	14,6	78	2,73282	7,5089
III	-0,8	26	0,02461	0,0676

Объединение материала значительно повысило надежность выводов. Для среднего квадрата по вариантам мы имеем тету, равную 7,5658. По таблицам при трех степенях свободы большей варианcы и 30 степенях свободы меньшей для вероятности отсутствия существенной разницы, меньшей 0,001, достаточна тета, равная 7,05. Таким образом, можем считать, что значимость имеет место для первой степени свободы вариантов (противопоставление 5 посещений остальным); и здесь для *P,* равной 0,001» достаточна тета, равная 13,29 (при 30 степенях свободы ошибки: для 36 степеней тета, конечно, уменьшится); мы же имеем 15,12. Противопоставление варианта 15 посещений более высоким тоже оказывается весьма существенным: здесь *P* близка 0,01 (тета для 30 степеней свободы ошибки равна 7,56, для 60—7,08, наша величина 7,51 немногим уступает 7,56). Более точное определение теты для 36 степеней свободы можно произвести следующим путем: 60/30 равно 1+24/36, или 1+2/3. Поэтому к величине теты для 60 степеней свободы (7,08) следует прибавить две трети разницы теты для 30 и 60 степеней свободы (7,56-7,08=0,48), т. е. 0,32. Мы получаем 7,08+0,32=7,40, что определенно меньше 7,51.

Мы видим, таким образом, какое существенное повышение надежности наших выводов получается от объединения материала. Это, конечно, имеет место только тогда, когда выводы, основанные на частных материалах, согласованы друг с другом. Резкое повышение надежности неудивительно, так как вероятность случайного возникновения какого-либо сложного результата (в данном случае — повторение тех же выводов на независимом материале) равна произведению вероятностей простых результатов. Это очень часто забывают многие исследователи. Проделывается, например, ряд сходных опытов, или получается несколько аналогичных сопоставлений. При каждом опыте или сопоставлении получается результат малой надежности, но совершенно одинакового характера. Часто делают совершенно ошибочное заключение, что общий вывод не доказан. Между тем если все выводы одного характера (т. е. во всех случаях один вариант превышает другой), то мы имеем полное право объединить весь материал (способы такого объединения, конечно, могут быть различны) и тогда совокупность результатов разных опытов, каждый из которых не дает надежного вывода, может привести к выводу совершенно надежному.

В таком неумении комбинировать результаты частных исследований кроется один из существенных источников неосновательных жалоб на то, что биометрическая обработка часто приводит к ненадежности выводов даже там, где эти выводы просто бросаются в глаза. При правильной обработке такие случаи невозможны: напротив, биометрически можно вскрыть существенное различие и там, где на глаз мы такого различия не находим.

4.4. РАЗЛОЖЕНИЕ ВАРИАНСЫ ПО СТЕПЕНЯМ СВОБОДЫ

При разборе примера с числом опылителей в главе о рандомизированных блоках уже было проделано разложение варианcы (суммы квадратов), соответствующей исследованным вариантам по отдельным степеням свободы. Такое разложение и помогает нам (в особенности широко им уместно пользоваться при факториальном анализе) извлечь все выводы из нашего материала и убедиться, какие из намечающихся выводов являются надежными, а какие основаны, вероятно, только на случайной изменчивости. Задача настоящей главы заключается, с одной стороны, в изложении техники такого разложения, а с другой стороны, в упоре на тот момент, что такое разложение целесообразно только тогда, когда оно дает ответ на определенные биологически осмысленные вопросы.

Когда имеется всего два варианта, то число степеней свободы равно единице и, следовательно, сумма квадратов для вариантов совпадает со средним квадратом. Разложение по степеням свободы предполагает минимум три варианта (две степени свободы). Однако уже при двух степенях свободы можно разложить нашу сумму квадратов бесчисленным числом способов, так как, уже было указано, чисто математически налагается только требование независимости (ортогональности) всех серий коэффициентов, что выражается в двух положениях: 1) сумма коэффициентов каждой серии должна быть равна нулю; 2) сумма попарных произведений коэффициентов любых двух серий тоже должна быть равна нулю.

Для большей наглядности возьмем произвольный числовой пример и покажем, как можно разложить этот материал по степеням свободы. Положим, суммы по трем вариантам (обозначим их через y_1, y_2 и y_3, причем каждая сумма получилась сложением двух дат) оказались равными 6, 8 и 10. Сумма квадратов отклонений около общей средней (равной, очевидно, 4, так как всего имеется шесть дат) равна

$$\frac{6^2+8^2+10^2}{2}-\frac{24^2}{6}=4$$

Покажем, как эту же величину можно получить суммированием квадратов для двух степеней свободы. Так как число вариантов три, то, очевидно, все коэффициенты не могут быть равными, так как при равенстве коэффициентов не может быть удовлетворено первое требование о равенстве нулю суммы всех коэффициентов.

Простейшим приемом будет противопоставление одного из вариантов двум другим, т. е. принятие коэффициентов 2, -1 и -1. Отсюда уже вытекают с необходимостью коэффициенты для другой степени свободы. В самом деле, обозначим эти неизвестные коэффициенты k_1, k_2 и k_3. Тогда согласно изложенным двум правилам мы должны иметь $k_1+k_2+k_3=0$ (имея в виду наличие уже у нас коэффициентов 2, -1 и -1).

Суммируя оба уравнения, получим $k_1=0$, т. е. первый коэффициент должен быть равен нулю. А тогда, очевидно, из первого уравнения $k_2=k_3$, и, принимая простейшее значение, единицу, получим значение коэффициентов 0,1, -1 (знаки коэффициентов можно изменить, это нисколько не влияет на результат, так как разница возводится в квадрат). Получаем табл. 16.

Делитель получается, как уже было указано, от суммирования квадратов всех коэффициентов данной степени свободы и умножения на число дат, послуживших для определения соответствующей суммы, а каждая разность от умножения сумм на соответствующие коэффициенты и суммирования, принимая во внимание, конечно, знаки.

Таблица 16

	y_1	y_2	y_3	Делитель Δ	Разность δ	$\frac{\delta^2}{\Delta}$
Сумма	6	8	10			
I степень свободы	2	-1	-1	12	-6	3,00000
II степень свободы	0	1	-1	4	-2	1,00000

Сумма 4,00000

Но даже это простейшее разложение может быть произведено в данном случае тремя способами, смотря по тому, где мы поставим коэффициент 2 для первой степени свободы. Например, его можно поставить против второго варианта. Мы тогда аналогичным образом получим табл. 17.

В данном случае вся изменчивость оказалась сосредоточенной в одной второй степени свободы, первая же степень дала разницу, равную нулю.

Таблица 17

	y_1	y_2	y_3	Делитель Δ	Разность δ	$\frac{\delta^2}{\Delta}$
Сумма	6	8	10			
I степень свободы	-1	2	-1	12	0	0.0000
II степень свободы	-1	0	-1	4	4	4.0000

Сумма 4.0000

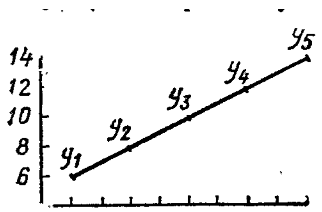

Рис. 4. Геометрическая интерпретация метода сумм — линейная комбинация пяти компонентов y_0, y_1, y_2 y_3, y_4, y_5 методе сумм (ось ординат: общая сумма пяти компонент; абсцисс: частные вклады пяти компонентов в виде их прямолинейной комбинации)

Прежде чем перейти к другим способам разложения суммы квадратов по двум степеням свободы, рассмотрим геометрический и биологический смысл этого первого, простейшего разложения. Это иллюстрируется рис. 4 и 5. На рисунке 4 показано графически прямолинейное возрастание y от y_1 до y_3 (пока не будем принимать во внимание y_4 и y_5). Тогда наша вторая степень свободы y_1—y_2 (во втором случае) показывает размер возрастания зависимой переменной на всем интервале наблюдения. Первая же степень свободы указывает на степень прямолинейности такого возрастания. В самом деле, по формуле для трапеции $y_2 = \frac{y_1+y_3}{2}$, отсюда $-y_1+2y_2-y_3=0$. То есть в случае прямолинейного возрастания величин от первой к третьей разность, соответствующая первой степени свободы, или должна равняться нулю (как это имеет место у нас в виду действительно строгой прямолинейности такого возрастания), или несущественно отличаться от нуля; суждение о несущественности отклонения от нуля достигается, как уже указывалось, сравнением со средним квадратом ошибки.

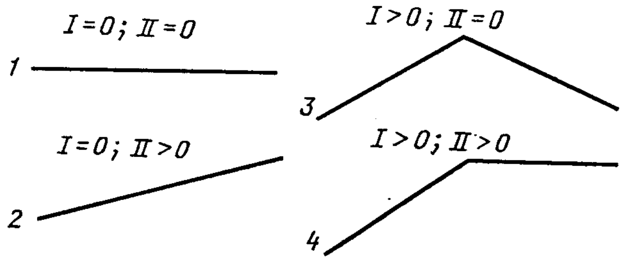

Рис. 5. Некоторые типичные случаи изменения y, которые дают различные значения для обеих степеней свободы (объяснения см. в тексте)

На рисунке 5 показаны некоторые типичные случаи изменения y, которые дают различные значения (принимая во внимание, конечно, только существенные различия) для обеих степеней свободы.

Линия 1 — прямая линия, параллельная оси абсцисс (реально будет обычно не прямая линия, но ломаная с незначительными отклонениями): в этом случае и обе степени свободы не дадут существенных отклонений от нуля, что и выражается формулой I=0, II=0.

Линия 2 — наклонная прямая. I степень свободы (критерий прямолинейности) даст 0; II степень свободы — критерий прямолинейной зависимости двух переменных — отлична от нуля: I=0, II>0.

Линия 3 — ломаная линия, причем y_1 и y_3 практически равны. Отклонение от прямой реально: I>0, II=0 (так как зависимость криволинейная).

Линия 4 — линия явно уклоняется от прямой, и y_1 не равно y_3. Обе степени свободы дают разности, отличные от нуля.

Разложение по двум степеням свободы может, таким образом, в случае наличия существенной разности между вариантами решить вопрос о принадлежности нашего материала к одному из трех типов: 2—4, что во многих случаях имеет определенный биологический смысл. Например, если мы проводим обработку каким-либо веществом растений (борьба с вредителями, удобрение и т. д.), то результат 2 (прямая пропорциональность) показывает, что испытуемое вещество увеличивает урожай пропорционально дозировке и, следовательно, есть перспектива, что увеличение дозировки может дать дальнейшее повышение урожая.

Результат 4 показывает, что разницы нет между второй и третьей дозировкой, следовательно, нет основания ожидать существенного увеличения урожая при дальнейшем увеличении дозировки: надо искать других путей. Наконец, результат 3 показывает, что третий вариант (y_3) показывает ухудшение со вторым, следовательно, имеется оптимум применения данного вещества и дальнейшее увеличение не принесет пользы, а только вред.

Следует лишь отметить, что принятое нами разложение для случая 4 указывает наличие криволинейной зависимости, но не позволяет точно проверить наличие или отсутствие разницы между вторым и третьим вариантами. Для проверки этого пригодно разложение, исследованное нами в самом начале, т. е. коэффициенты для второй степени свободы 0,1 и -1. Для .случая 4 вторая степень свободы даст разницу, несущественно отличающуюся от нуля, а первая степень свободы отличается от нуля.

Таким образом, разложение по двум степеням свободы преследует цель выяснить характер зависимости между переменными, но ответ может заключаться только в выборе между прямолинейной и криволинейной зависимостью. Судить о характере криволинейной зависимости при трех вариантах невозможно, так как через три точки можно провести любую кривую, имеющую три параметра.

Но во многих случаях проверку прямолинейности при трех вариантах необходимо производить, пользуясь иными коэффициентами, чем только что разобранные. Это имеет место всегда, когда интервалы между независимыми переменными неодинаковы. Возьмем опять рис. 4 и положим, что у нас три варианта, помеченные на чертеже, y_1, y_2, и y_3, причем расстояние между третьим и вторым вариантами в три раза больше, чем между вторым и первым. Тогда, если мы желаем проверить гипотезу о пропорциональности дозировки и прибавки к урожаю, мы получаем следующее соотношение:

$y_5—y_2=3(y_2—y_1)$. Отсюда $y_5—4y_2+3y_1=0$ и мы получаем первую серию коэффициентов: 3, -4, 1. Чтобы получить серию коэффициентов l_1, l_2, l_3, необходимо использовать уже указанные два правила, получаем два уравнения:

$$l_1+l_2+l_3=0,$$
$$3l_1-4l_2+l_3=0,$$

отсюда $l_1=5/2\ l_2$, так как один коэффициент произволен, то, беря простейшее значение l_2 равное 2, получим $l_1=5$ и $l_3=-7$. Используя наш произвольный числовой призер, получаем новое разложение (табл. 18).

Получили, как и следует ожидать, совершенно ту же сумму. Вообще, для одной из степеней свобод можно взять какой угодно выбор коэффициентов (только чтобы их сумма равнялась нулю) и тогда, применяя использованные уже два уравнения, вычислить набор ортогональных коэффициентов (табл. 19).

Таблица 18

	y_1	y_2	y_3	Делитель Δ	Разность δ	$\frac{\delta^2}{\Delta}$
Сумма	6	8	10			
I степень свободы	3	-4	1	52	-4	0.30769
II степень свободы	5	2	-7	156	-24	3.69231
					Сумма	4,00000

Таблица 19

	y_1	y_2	y_3	Делитель Δ	Разность δ	$\frac{\delta^2}{\Delta}$
Сумма	6	8	10			
I степень свободы	5	-4	-1	84	-12	1.7142857
II степень свободы	1	2	-3	28	-8	2.2857143
					Сумма	4,0000000

Это разложение интересно тем, что в отличие от предыдущих разложений на обе степени свободы приходятся очень близкие квадраты разности. Можно подобрать такие коэффициенты, что общая сумма квадратов распределится поровну между обеими степенями свободы. Таким образом, чисто математически можно добиться двух крайних случаев: или сосредоточения всей изменчивости на одной степени свободы (как это мы видели при втором разложении), или равномерного или почти равномерного распределения дисперсии. Этого, конечно, можно добиться и при большем числе степеней свободы, но если вопрос подвергать чисто математической трактовке (без связи с реальной действительностью), то это может привести к ошибочным выводам. Разберем опять-таки произвольный пример.

Положим, у нас в опыте было 9 вариантов при четырехкратной повторности. Тогда мы имеем всего 35 степеней свободы, из которых 8 соответствуют вариантам, 3 — повторностям и 24 — ошибке. Предположим, что тета для вариантов оказалась равной 2,0. При таком значении, теты и данном числе степеней свободы средняя изменчивость не выходит за рамки случайной, так как для минимального уровня значимости (вероятность отсутствия существенных различий, равная 0,05) требуется тета, равная 2,38. Но если отношение среднего квадрата вариантов к среднему квадрату ошибки равно 2,0 то отношение суммы квадратов вариантов будет равно 16,0, и, умело производя разложение по степеням свободы, можно добиться того, что на одну из степеней свободы ляжет по крайней мере половина всей дисперсии, что и даст тету для этой степени свободы, равную 8. Но такая тета, (при одной степени свободы для большей варианcы и24 —для меньшей) уже соответствует вероятности отсутствия различия меньшей 0,01, так как для этого достаточно тета, равная 7,82.

Правилен ли будет вывод о значимости различия, соответствующей этой степени свободы? Он будет неправильным в том случае, если мы подбирали коэффициенты исключительно с целью сосредоточения большей части изменчивости на одну степень свободы: в таком случае это будет неправильное применение математических критериев, рассчитанных на проверку априорных предположений, а не чисто эмпирических контрастов. Напротив, такой вывод будет вполне правильным, если сосредоточение изменчивости на одной степени свободы явилось следствием проверки определенной, заранее поставленной гипотезы. Возвращаясь к случаю с опылителями, мы могли взять не 4 варианта количества посещений, а 9, но выбрать их, положим, такими: 10; 25; 40; 55; 70; 85; 100; 115 и 130 посещений. Проанализированный пример привел к заключению, что число посещений, достаточное для. полного опыления, лежит где-то около 20—25 посещений.

Поэтому при взятом нами распределении вариантов мы должны ожидать существенного различия только между первым и всеми остальными вариантами, а отнюдь не между высшими 8 вариантами. Вполне естественно, что наличие одного существенного контраста среди семи несущественных приведет к тому, что этот существенный контраст затеряется среди несущественных и средний квадрат разности для вариантов будет, по всей вероятности,. несущественно отличаться от квадрата ошибки. Но в данном случае мы имеем полное право извлечь этот единственный существенный контраст из массы несущественных путем противоположения в данном случае первого варианта всем остальным. Поэтому единственным биологически обоснованным разложением по степеням свободы среди бесчисленного числа математически возможных в данном случае будет следующее (привожу только ортогональные наборы коэффициентов):

Варианты	1	2	3	4	5	6	7	8	9
I степень свободы	-8	1	1	1	1	1	1	1	1
II “ “	0	-7	1	1	1	1	1	1	1
III “ “	0	0	-6	1	1	1	1	1	1
IV “ “	0	0	0	-5	1	1	1	1	1
V “ “	0	0	0	0	-4	1	1	1	1
VI “ “	0	0	0	0	0	-3	1	1	1
VII “ “	0	0	0	0	0	0	-2	1	1
VIII “ “	0	0	0	0	0	0	0	-1	1

Варианты, конечно, располагаются в данном случае по возрастанию числа посещений, и если имеет место влияние числа посещения на завязность ,(это влияние может и не проявляться, если бы, например, оказалось, что уже 10 посещений хватает для полного опыления и что, следовательно, всякое превышение этого числа не имеет никакого значения), то, очевидно, оно должно всего сильнее сказаться при противопоставлении первого минимального варианта всем остальным и затем постепенно падать с каждой степенью свободы.

Такое разложение будет, таким образом, биологически вполне обоснованным, и полученные выводы будут полноценны. Если же мы просто разложим наши результаты в возрастающем порядке завязности и потом станем их исследовать по указанной схеме, то выводы уже не будут иметь того значения, как в первом случае, и, настаивая на их значении, мы впадаем в грех ползучего эмпиризма. Почему? Потому, что если мы распределяем варианты по их эмпирическим значениям, то при полном отсутствии существенного различия между вариантами любой из 9 вариантов может оказаться наименьшим. И беря вариант в опыте, давший наименьшую завязность, мы выбираем один случай из девяти независимых, следовательно, производим как бы 9 испытаний. Но чем больше испытаний, тем больше возможность того, что произойдет событие маловероятное: за время существования рулетки в Монте-Карло был случай, когда «красное» вышло 17 раз подряд, но было бы очень странным, если бы человек, ставя 17 раз подряд на красное, все 17 раз подряд выиграл.

Указанное разложение в виде противоположения одного из вариантов всем остальным является одним из простейших и применимо к любому числу степеней свободы, только надо помнить, что если мы подходим к материалу чисто эмпирически, совершенно не имея никаких априорных предположений, то критерии значимости должны быть значительно более строгими, чем в том случае, если разложение соответствует определенным теоретическим представлениям. Нетрудно показать, что если мы противополагаем один вариант всем остальным, то во всех последующих степенях свободы этот первый вариант уже отсутствует и коэффициенты для него, следовательно, равны нулю. В самом деле, имеем, положим, n коэффициентов, обозначаемых

$$k_1, k_2, k_3, \ldots, k_n$$

Для первой степени свободы мы берем их значения:

$$+(n-1),-1,\ -1,\ \ldots\ ,-1.$$

Следовательно, для любой другой степени свободы должны иметь место равенства:

$$(n-1)\left\{\begin{array}{l} k_1+k_2+k_3+\ldots+k_n=0 \\ -k_1-k_2-k_3-\ldots-k_n=0 \end{array}\right\}$$

Складывая оба равенства, получим $nk_1=0$, или $k_1=0$. Уже это стандартное разложение может быть достигнуто значительным числом способов: в самом деле, любой вариант может быть избран в качестве первого: n возможностей, в качестве 2-го n—1, 3-го n—2, всего n(n-l) (n—2) ... 3, 2, 1, или n! разных способов разложения (для пяти вариантов, например 5-4-3-2, имеем 120 способов разложения по степеням свободы) одной и той же суммы квадратов.

Остановимся теперь на разложении суммы квадрата по трем степеням свободы (при четырех вариантах). Здесь кроме того стандартного разложения, которое уже было применено в главе о рандомизированных блоках, очень распространено разложение:

I степень свободы	1	1	-1	-1
II » »	1	-1	1	-1
III » »	1	-1	-1	1

Это разложение — простейшее, поскольку все коэффициенты единицы, и играет большую роль при факториальном анализе, почему сейчас на нем мы задерживаться не будем. Остановлюсь на применении разложения по степеням свободы для выяснения формы зависимости между переменными. Для выяснения того, .какой набор коэффициентов может служить для проверки прямолинейности, обратимся опять к рис. 4. Если имеем 4 точки (концы ординат y_1, y_2, y_3 и y_4)> то, очевидно, при прямолинейной зависимости и равных интервалах мы имеем y_4-y_3=y_2-y_1, или y_1-y_2-y_3+y_4=0,что приводит к набору коэффициентов 1, —1, —1 и 1, ;служащих для вычисления разницы (если квадрат разности существенно превышает квадрат ошибки, то зависимость не прямолинейная).

Но наличие четырех точек позволяет не только решить вопрос о непрямолинейности зависимости, но и выяснить степень отклонения от какой-либо намеченной кривой с тремя параметрами (для суждения о совпадении с кривой с четырьмя параметрами четырех точек, конечно, недостаточно).

Предположим, что мы хотим проверить, соответствует ли зависимость наших двух переменных параболической зависимости, выражаемой формулой:

$$y=ax^2+x+c$$

тогда при наличии четырех точек и равном расстоянии между значениями независимой переменной x можем этой переменной придать значение

$$x_1=0,\ x_2=1,\ x_3=3$$

Получим четыре уравнения:

	Первая разность	Вторая разность
$y_1 = c$		
$y_2 = a + b + c$	$a+b$	$2a$
$y_3=4a+2b+c$	$3a+b$	$2a$
$y_4=9a+3b+c$	$5a+b$	

Вычитая последовательно из каждого следующего уравнения предыдущее и проделывая это 2 раза, получаем, как и следует ожидать, равные вторые разности. Следовательно,

$$(y_3-y_2)-(y_2-y_1)=(y_4-y_3)-(y_3-y_2),\ \text{или}\ y_1-3y_2+3y_3-y_4=0.$$

Следовательно, набор коэффициентов 1, —3, 3, —1 может служить для выяснения степени соответствия нашего материала параболической зависимости. Этот набор, как легко проверить, ортогонален с набором для проверки прямолинейности регрессии (1, —1, —1, 1), и потому оба набора могут быть включены в одну систему. Коэффициенты для третьей степени свободы вычисляются так, чтобы они оказались ортогональными к первым двум степеням: такое разложение может и не иметь биологического значения, но проделать его целесообразно для контроля вычисления.

Мы имеем три уравнения:

$$k_1+k_2+k_3+k_4=0, \qquad (1)$$

$$k_1-3k_2+3k_3-k_4=0, \qquad (2)$$

$$k_1-k_2-k_3+k_4=0, \qquad (3)$$

из (1) и (3) получаем k_1=-k_4, k_2=-k_3, отсюда из (2): k_1=3k_2 и k_4=3k_3

Беря для k_2 простейшее — единицу, получаем серию коэффициентов: 3, 1, —1,—3. Вся система ортогональных коэффициентов получает вид:

I степень свободы	1,	—1,	—1,	1 (критерий прямолинейности)
II » »	1,	—3,	3,	—1 (критерий параболичности)
III » »	3,	1,	—1,	—3 (дополнение для контроля вычислений)

Рассуждая аналогичным образом, можно приспособить разложение к проверке любой гипотезы, сформулированной математически, при условии, конечно, что число вариантов достаточно.

При числе вариантов больше четырех мы, во-первых (там, где это биологически имеет смысл), можем применять стандартное разложение. Разложение при факториальной схеме будет разобрано в своем месте.

Здесь я коснусь только случаев критерия прямолинейности и сопоставлений при качественных различиях вариантов. В обоих случаях мы разбиваем наши варианты на группы и в пределах каждой группы производим самостоятельное сравнение.

Положим, надо выяснить степень отклонения от прямолинейности в случае с девятью вариантами. Тогда можно поступать так (это, конечно, не единственный возможный способ), что сначала выясняем наличие отклонения от прямолинейности в грубом масштабе, объединяя по три соседних варианта, а потом в пределах каждой тройки производим отдельное исследование.

Получаем такую систему:

I степень свободы	1	1	1	-2	-2	-2	1	1	1
II » »	1	1	1	0	0	0	-1	-1	-1
III » »	1	-2	1	0	0	0	0	0	0
IV » »	1	0	-1	0	0	0	0	0	0
V » »	0	0	0	1	-2	1	0	0	0
VI » »	0	0	0	1	0	-1	0	0	0
VII » »	0	0	0	0	0	0	1	-2	1
VIII » »	0	0	0	0	0	0	1	0	-1

Другой случай: мы имеем семь вариантов, из коих четыре образуют одну группу (положим, опыление против вредителей), а остальные — другую (положим, опрыскивание). Тогда целесообразно сначала противопоставить одну группу другой, а потом в пределах каждой группы производить сравнение. Первая степень свободы используется для сравнения групп, три — для различий в пределах опыления и остальные две — для сравнения в пределах опылителей. Первое сравнение производится путем уравновешивания коэффициентов (чтобы сумма их равнялась нулю), а остальные — смотря по характеру опыта: или мы имеем разные дозировки, или качественно различные яды. Беря первый случай, получаем, примерно, такое разложение:

№ варианта	1	2	3	4	5	6	7
I степень свободы	3	3	3	3	-4	-4	-4
II » »	3	-1	-1	-1	0	0	0
III » »	0	2	-1	-1	0	0	0
IV » »	0	0	1	-1	0	0	0
V » »	0	0	0	0	2	-1	-1
VI » »	0	0	0	0	0	1	-1

Как нетрудно проверить, все серии коэффициентов будут ортогональны. Так и должно быть, потому что кроме первой степени свободы в сравнении участвуют варианты только одной из групп, а в этом случае мы, очевидно, должны заботиться об ортогональности только в пределах данной группы, так как все остальные коэффициенты равны нулю и потому, будучи помноженными на любые коэффициенты, дадут всегда нули.

Последовательно применяя эти приемы, можно любое количество вариантов исследовать на наличие или отсутствие существенных различий, и так как при этом мы всегда должны получить полное совпадение сумм квадратов, вычисленных двумя независимыми способами, то это является хорошей проверкой вычислений. Проверкой же никогда не следует пренебрегать, так как даже опытные вычислители от ошибок не застрахованы.

4.5. ЛАТИНСКИЙ КВАДРАТ

Метод рандомизированных блоков, как ясно из вышеизложенного, преследует две цели:

1) с одной стороны, он дает оценку относительного значения применяемых обработок и вообще вариантов опыта, неискаженную смешением с различиями разных участков опыта;

2) с другой стороны, гетерогенность поля нашего опыта выделяется в особую категорию изменчивости и не смешивается с ошибкой опыта, отчего несмотря на различие естественноисторических условий различных блоков ошибка опыта остается неувеличенной.

Но нередки случаи, когда гетерогенность поля имеет несколько направлений. Положим, опыт закладывается на участке, имеющем падение рельефа с одной стороны к другой. Так как рельеф имеет значение для естественного плодородия, то построение блоков следует вести так, чтобы каждый блок занимал полосу вдоль горизонталей, следовательно, отдельные блоки отличались *бы* друг от друга по расположению на склоне. Этим путем различия, связанные с рельефом, не окажут влияния на результаты опыта, так как каждый вариант опыта заключает все модальности рельефа и каждая модальность рельефа заключает все варианты опыта. Но одновременно с рельефом участки поля могут отличаться, положим, по предшественнику и может случиться, что границы по предшественникам идут под прямым углом к грани-дам по рельефу. Тогда этот элемент гетерогенности не сможет привести к искажению выводов в силу рандомизации опыта, так как изменчивость, связанная с предшественниками, будет присоединена к изменчивости, связанной с ошибкой опыта. Эта последняя приобретает большие размеры и потому выводы потеряют в своей отчетливости и потребуют для той же четкости большее число повторений. Ясно, что желательно найти такой метод, который и эту часть изменчивости отделил бы от ошибки опыта. Этот метод и предложен в виде метода латинского квадрата.

Не следует думать, что «направления» в латинском квадрате обязательно являются топографическими или что мы обязательно должны иметь квадратные участки земли. Слово «квадрат» в термине латинский квадрат имеет единственный смысл: число вариантов опыта должно быть равно числу повторностей; само собой разумеется, что допустима тождественность некоторых вариантов, но число вариантов (считая тождественные варианты при закладке опыта за разные варианты) в этом смысле обязательно равно числу повторностей. Например, при полевом опыте одно из направлений может быть топографическим, соответствовать блокам в методе рандомизированных блоков, но в пределах каждого блока может быть подобран, положим, набор разных сортов, постоянный для всех блоков и одинаковый с числом повторностей. Тогда сортовое различие даст другое «направление» латинского квадрата. Возможны случаи, особенно в лабораторной обстановке, тогда вообще никаких топографических различий не будет. Одно направление, положим, будет соответствовать времени, один день будет соответствовать одной строке латинского квадрата, а в пределах каждого дня, положим, будут проводиться несколькими работниками параллельные опыты. Если число дней опыта равно числу работников и числу вариантов опыта, то результаты опыта могут быть обработаны по методу латинского квадрата и изменчивость, связанная с днем проведения опыта и индивидуальностью работников, не окажет никакого влияния на правильность и точность выводов и даже может быть сама изучена по ходу обработки.

Сущность организации работы по методу латинского квадрата очень проста. Взяв принятые два направления изменчивости, опыт располагают так, чтобы в каждой строке и в каждом столбце были представлены все варианты опыта, и так как число вариантов равно числу повторностей, то, следовательно, ни в одной из строк и ни в одной из повторностей не будет повторения вариантов. Этому удовлетворяет, например, такое расположение вариантов (беря, положим, латинский квадрат 5-5):

A	*B*	*C*	*D*	E
B	*C*	*D*	*E*	*A*
C	*D*	*E*	*A*	*B*
D	*E*	*A*	*B*	*C*
E	*A*	*B*	*C*	*D*

Такое расположение является латинским квадратом, и некоторыми опытниками рекомендовался именно такой диагональный квадрат. Но применение такого квадрата заключает в себе ту опасность, что вариант *E* лежит по диагонали, а остальные варианты ей параллельны. Поэтому, если различие по плодородию по тем или иным, непредвиденным исследователем причинам идет наискось, то вариант *E* получит систематическое различие по сравнению с другими вариантами, совершенно не связанное с обработкой, но ошибочно приписанное ей. Напротив, так как важная часть гетерогенности поля будет смешана с различием вариантов,. то доля изменчивости, падающая на ошибку опыта, будет уменьшена: это уменьшение будет не истинным, основанным на уточнении опыта, а ложным, основанном на неправильной организации опыта.

Для устранения первой ошибки другое систематическое расположение квадрата было предложено Кнут Виком: здесь уже те же варианты соседних строк или столбцов отличались на два шага, а не на один. Пример такого расположения имеем:

A	*B*	*C*	*D*	*E*
D	*E*	*A*	*B*	*C*
B	*C*	*D*	*E*	*A*
E	*A*	*B*	*C*	*D*
C	*D*	*E*	*A*	*B*

В данном случае варианты опыта расположены идеально равномерно по полю, и поэтому совершенно исключена возможность, чтобы те или иные компоненты гетерогенности поля могли спутать сравнения между вариантами. Но оказывается, что такое систематическое распределение по сравнению с рандомизированным обладает другим недостатком, гораздо менее наглядным: если даже компонент ошибки опыта при данной расположении менее влияет на сравнение вариантов, чем при рандомизированном расположении, то зато увеличивается компонент ошибки, служащий для оценки размеров ошибки, отчего надежность вывода уменьшается. Это теоретическое заключение Р. Фишера получило экспериментальное подтверждение в работе О. Тедина. Поэтому правильное применение метода латинского квадрата заключается в использовании правильно рандомизированных квадратов. Такая рандомизация сводится к тому, что для опыта путем рандомизации выбирается один квадрат из всех возможных. Остановимся поэтому на вопросе о числе латинских квадратов разных порядков.

Совершенно ясно, что 2-2 латинских квадратов может быть только два:

A	*B*	*B*	*A*
B	*A*	*A*	*B*

но, конечно, такая простая схема не используется в опыте.

Для 3-3 квадратов имеется всего 12 различных латинских квадратов. Именно мы имеем только один квадрат в так называемом стандартном положении, т. е. когда и в первой строке, к в первом столбце буквы расположены в алфавитном порядке:

A	*B*	*C*
B	*C*	*A*
C	*A*	*B*

Нетрудно видеть, что никакой другой стандартный квадрат 3-3 невозможен. Оставив без изменения первую строку и первый столбец (иначе квадрат перестанет быть стандартным), мы видим, что, например, в центре квадрата нельзя поставить ни букву *B* (так как *B* уже имеются и во второй строке, и втором столбце), ни букву *A*, так как тогда останется на третьем месте второй строки буква *C*, которую поставить невозможно, так как в третьем столбце уже имеется буква С. Но мы можем преобразовать наш стандартный квадрат таким образом, что сначала, не трогая первой строки, переменим положение двух других, а затем в каждом из полученных двух квадратов, не трогая первого столбца, переменим положение двух других столбцов. Мы получим четыре квадрата:

1) *A B C*	2) *A B C*	3) *A C B*	4) *A C B*
B C A	*C A B*	*B A C*	*C B A*

C A B   *B C A*   *C B A*   *B A C*

Все эти четыре различных квадрата имеют общую черту: первое место и по строкам, и по столбцам занимает буква *А*. Если .мы теперь для всех первых четырех квадратов заменим букву *А* на *В*, *В* на *С* и *С* на *А*, то получим новые четыре квадрата, отличных от первых четырех, а если вновь заменим подобным же так называемым циркулярным способом *В* на *С*, *С* на *А* и *А* на *В*, то получим еще четыре квадрата. При новой циркулярной замене мы вернемся к первым четырем квадратам. Отсюда ясно, что полученные нами 12 квадратов исчерпывают все возможное разнообразие латинских квадратов 3-3. Приведем оставшиеся восемь квадратов

5)	*B*	*C*	*A*	6)	*B*	*C*	*A*	7)	*B*	*A*	*C*	8)	*B*	*A*	*C*
	C	*A*	*B*		*A*	*B*	*C*		*C*	*B*	*A*		*A*	*C*	*B*
	A	*B*	*C*		*C*	*A*	*B*		*A*	*C*	*B*		*C*	*B*	*A*
9)	*C*	*A*	*B*	10)	*C*	*A*	*B*	11)	*C*	*B*	*A*	12)	*C*	*B*	*A*
	A	*B*	*C*		*B*	*C*	*A*		*A*	*C*	*B*		*B*	*A*	*C*
	B	*C*	*A*		*A*	*B*	*C*		*B*	*A*	*C*		*A*	*C*	*B*

Число 12 получается как произведение 1-2-2-3, что означаем 1 — стандартный квадрат, 2 — число возможных перемещений строк, исключая первую, 2 — то же для столбцов и 3 — циркулярные перемещения (благодаря наличию трех вариантов).

Латинский квадрат 3-го порядка применяется в опыте редко: во-первых, потому что число вариантов мало и получается очень мало степеней свободы для суждения об ошибке, именно

строки	2 степени свободы
столбцы	2 » »
варианты	2 » »
ошибка	2 » »
Всего 8 степеней свободы	

Во-вторых, объем рандомизации невелик (всего 12 возможных квадратов), и, наконец, все квадраты оказываются диагональными. Широкое применение он имеет для разбора сложных комбинаторных соотношений при факториальном анализе. В опытах применяются латинские квадраты не ниже 4-го порядка (и редко выше 8-го). Для латинского квадрата 4-го, 5-го и вообще *п-го* порядка мы имеем следующее распределение числа степеней свободы:

	4-4	5-5	6-6	s-s
строки	3	4	5	s–1
столбцы	3	4	5	s–1
варианты опыта	3	4	5	s–1
ошибка	6	12	20	(s–1)(s–2)
Всего	15	24	35	$s^2–1$

Покажем самым элементарным путем, как определяется число латинских квадратов 4-го порядка. Сначала определим число квадратов в стандартном положении. Написав по алфавиту первую строку и первый столбец, мы на второе место по столбцам и строкам ставим сначала тоже букву *А*, а затем следующие совместные с требованиями латинского квадрата. Таким образом, без труда увидим, что стандартных положений для 4-4 будет только четыре, а именно:

1) *А В С Е*   2) *А В С Е*   3) *А В С Е*   4) *А В С Е*
В А Е С   *В А Е С*   *В С Е А*   *В Е А С*
С Е А В   *С Е В А*   *С Е А В*   *С А Е В*
Е С В А   *Е С А В*   *Е А В С*   *Е С В А*

Из каждого из этих стандартных квадратов, сохраняя в первоначальном положении либо первую строку, либо первый столбец, можем получить 6 новых, меняя лишь столбцы (беря расположения *ВСЕ, ВЕС, СВЕ, СЕВ, ЕВС, ЕСВ},* и 6, меняя только строки: при комбинировании перемещений столбцов и строк получим из каждого стандартного квадрата 6-6, или 36 квадратов. Наконец, меняя буквы циркулярно, получим из каждого квадрата четыре, следовательно, получаем 4-6-6-4, или 576 латинских квадратов 4-го порядка. Конечно, было бы очень утомительно выписывать все 576 возможных латинских квадратов и по жребию выбирать один из них, да в этом нет и надобности. Правильный путь состоит в следующем: сначала по жребию выбирают один из четырех стандартных квадратов. Затем в данном стандартном квадрате по жребию же производят перемещение строк: написав, положим, на трех листках буквы *В, С* и *Е* и перетасовав, выложить их один за другим. Третьим шагом будет перемещение столбцов в квадрате, где уже было сделано перемещение строк и, наконец, четвертым и последним шагом будет установление того, какая буква соответствует каждому варианту опыта. Таким образом, путем таких четырех последовательных рандомизаций, которые в совокупности отнимут несколько минут, достигается безупречный выбор одного из 576 возможных квадратов.

С увеличением порядка квадратов число возможных квадратов чрезвычайно возрастает. Так, для квадрата пятого порядка имеется уже 56 квадратов в стандартном положении, и всего имеется 94080 различных латинских квадратов. Для шестого порядка число их измеряется почти 500 миллионами. В данном случае при таком огромном разнообразии можно и не требовать обязательно педантичной рандомизации, включая выбор одного из стандартных квадратов; достаточно ограничиться тремя этапами рандомизации, т. е., взяв произвольный исходный квадрат (хотя бы даже диагональный), произвести сначала рандомизированное перемещение строк, затем проделать такое же рандомизированное перемещение столбцов и, наконец, по жребию определить соответствие букв и вариантов опыта. Само собой разумеется, что, как редкий случай, в результате рандомизации может оказаться и диагональный квадрат, но еще более удивительно, если такой редкий случай совпадет как раз с диагональным распределением плодородия. Если же практиковать диагональный квадрат как правило, то такое совпадение будет не таким редким.

В качестве конкретного примера применению латинского квадрата возьму данные из работы Хервея и Хартцелля (Hervey, Hartzell, 1931) по выяснению сроков посева на заражение кукурузы кукурузным мотыльком. Кукуруза была посеяна на 25 делянках, расположенных в виде квадрата, и посевы производились через 10 дней, начиная со 2 мая до 11 июня. В приведенной табл. 20 показано расположение делянок, причем римские цифры обозначают сроки посева, а арабские — число гусениц на 25 стеблей, цифры округлены для удобства вычислений.

Из данных табл. 20 помимо средних по столбцам и по строкам Извлечем данные по зараженности, по срокам посева.

Мы видим, что сроки посева дают значительное различие зараженности: максимальная зараженность 2-го срока в пять раз с лишним превышает минимальную зараженность пятого срока. Но с другой стороны, рассматривая зараженность по строкам, видим, что зараженность второй строки в три раза слишком превышает зараженность первой. Так как в каждой строке имеется полный набор всех пяти сроков посева, то, следовательно, различия зараженности по строкам являются следствием гетерогенности поля. Отсюда ясно, что если превышение зараженности в три с лишним раза заведомо лежит в пределах ошибки опыта, то неясно, лежит ли в пределах ошибки и пятикратное превышение зараженности.

Таблица 20

Зараженность кукурузы кукурузным мотыльком в зависимости от сроков посева

										Сумма	Среднее
I	3	II	8	III	4	IV	4	V	0	19	3,8
II	30	IV	13	V	9	III	2	I	6	60	12,0
IV	2	III	14	I	5	V	6	II	25	52	10,4
V	0	I	4	IV	4	II	11	III	3	22	4,4
III	11	V	2	II	16	I	6	IV	2	37	7,4
Сумма	46		41		38		29		36	190	
Среднее	9,2		8,2		7,6		5,8		7,2		7,6

I срок посева	сумма	24	среднее	4.8
II « «	«	90	«	18.0
III « «	«	34	«	6.8
IV « «	«	25	«	5.0
V « «	«	17	«	3.4
Всего		190		7.6

Вычисления проводятся по той же схеме, как и по методу рандомизированных блоков с тем отличием, что вводится еще одна категория с четырьмя степенями свободы.

Вычисляем сначала общую сумму квадратов для исходных дат: 3, 8, 4, 4 и т. д. — и из этой суммы квадратов вычитаем поправку:

$$\frac{190^2}{25} = 1444{,}0$$

Получаем общую сумму квадратов от общей средней:

2764,0—1444,0=1320,0.

По строкам получаем

$$\frac{19^2 + 60^2 + 52^2 + 22^2 + 37^2}{5} = 1703{,}6 \text{ (поправка-1444,0).}$$

Сумма квадратов по строкам от общей средней — 259,6. То же по столбцам:

$$\frac{46^2 + 41^2 + 38^2 + 29^2 + 26^2}{5} = 1475.6 \text{ (поправка-1444.0).}$$

Сумма квадратов столбцов от общей средней — 31,6.

Наконец, по срокам посева получаем

$$\frac{24^2 + 90^2 + 34^2 + 25^2 + 17^2}{5} = 2149.2 \text{ (поправка-1444,0).}$$

Сумма квадратов по срокам посева от общей средней — 705,2.

Вычитая из общей суммы, получим сумму квадратов для ошибки. Как будет показано дальше, в разделе о греколатинском квадрате, эту сумму квадратов тоже можно вычислить непосредственно, что является хорошей проверкой вычислений. Получаем следующий анализ варианса (табл. 21).

Таблица 21

Категории изменчивости	Число степеней свободы	Сумма квадратов	Средний квадрат	θ	p
Строки	4	259,6	64,900	2,407	>0,05
Столбцы	4	31,6	7,900		
Сроки посева	4	705,2	176,300	6,538	<0,01
Ошибка	12	323,6	26,967		
Всего	24	1320,0			

Из таблицы 21 уже ясно, что гетерогенность поля по строкам далеко не достигает даже минимального уровня значимости *(P* равно 0,05), для чего требуется при данном числе степеней свободы (4 и 12) тета, равная 3,26. Что же касается сроков посева, то тут различие имеет вполне существенное значение.

Тета по столбцам и не вычислялась, так как средний квадрат для столбцов заметно меньше среднего квадрата ошибки. Может возникнуть вопрос, не является ли такое различие существенным, т. е. указывающим на какие-то уравнительные факторы по столбцам, приводящие к изменчивости много меньшей чисто случайно? Для выяснения этого надо большую вариансу (в данном случае средний квадрат ошибки 26,967) разделить на меньшую (средний квадрат для столбцов 7,900) и оценить степень значимости подобной теты. Получим 3,42, что далеко не достигает даже минимального уровня значимости: для 12 и 4 степеней свободы минимально значимая тета, равная 5,91.

В данном случае различие зараженности между вторым сроком посева и остальными выражено настолько резко, что ясно, что это и есть единственное существенное различие. В других случаях приходится убеждаться в этом, оценивая различие разных вариантов. Методически это производится или вычислением средней ошибки разности двух вариантов, или разложением по степеням свободы. Покажем оба приема для данного случая.

Средняя варианса, соответствующая средней ошибке единичного наблюдения всего опыта, как видно из табл. 21, равна 26,967; отсюда квадрат средней ошибки для среднего по сорту (поскольку такое среднее основано на пяти датах) равен $\frac{25.967}{5}$. Квадрат средней ошибки разности двух средних равен, как известно, сумме квадратов средних ошибок обоих средних. Так как в данном случае средняя ошибка является общей для всего опыта, то вместо суммы приходится взять двойную среднюю ошибку для сорта. Получим, что средняя ошибка разности равна

$$\sqrt{\frac{26.976 \cdot 2}{5}} = \sqrt{10.7868} = 3.285.$$

Для того чтобы узнать размер разности между двумя сортами, соответствующей принятым трем уровням значимости (Р, равное 0,05, и 0,01 и 0,001), мы должны взять *t*-критерий для этих трех уровней значимости, основанный на 12 степенях свободы, и перемножить на вычисленную нами среднюю ошибку разности (3,285). Получим:

для *P*	0,05	0,01	0,001
для *t*	2,179	3,055	4,317
3,285 (*t*)	6,94	9,72	14,2

Разность зараженности 2-го срока и 5-го (14,6%) удовлетворяет высшему уровню значимости *(P* меньше 0,001), разница 2-го срока и остальных (не меньше 11,2) удовлетворяет, как видим, тоже достаточно высокому уровню значимости, и (Р заметно меньше 0,01) разница между сроками посева, кроме 2-го, не имеет никакой значимости.

Такой метод, впервые он был дан Стьюдентом для оценки разности в сортоиспытании, имеет широкое применение. Его недостаток заключается в том, что если мы при помощи его определяем не заранее намеченные различия, а подбираем различия чисто эмпирически, то при большом числе сортов или вариантов. надо иметь в виду, что число независимых сопоставлений эмпирических величин можно считать равным числу испытаний, и ясно,. что если мы проделаем большое число испытаний, то маловероятное событие (случайное возникновение достаточно большой разницы) становится уже более вероятным. В данном случае число сроков невелико и, кроме того, мы имеем, чтр и средний квадрат для сроков посева соответствует очень высокому уровню значимости.

Другой метод сравнений, уже применявшийся в главе о рандомизированных блоках, заключается в разложении по степеням свободы. Этот метод, как известно, обладает тем большим преимуществом, что, пользуясь системой ортогональных коэффициентов, мы получаем хорошую проверку проделанных вычислений. Получаем следующее разложение (табл. 22).

Распределение контрастов по степеням свободы за отсутствием какого-либо теоретического подхода проводим по эмпирическим данным: сначала противопоставление наиболее высокого по зараженности второго срока остальным, затем наименее зараженного пятого срока первому, третьему и четвертому и т. д. В результате видим, что только первая степень свободы дает существенное различие, но при том очень высоко существенное (так как для P^ равного 0,001, достаточна тета, равная 18,64). При такой высокой значимости и небольшом числе степеней свободы даже эмпирический контраст может считаться доказанным. Биологически наш результат следует, очевидно, толковать так: и очень ранние, и очень поздние сроки посева мало заражаются, подвержен сильному заражению только один срок, вероятно, совпадающий с массовым лётом кукурузного мотылька.

Таблица 22

Сроки посева	I	II	III	IV	V	Делитель Δ	Разность δ	$\frac{\delta^2}{\Delta}$	θ	P
Сумма	24	90	34	25	17					
1	—1	4	—1	—1	—1	100	260	676,00	25,068	<0,001;
2	1	0	1	1	3	60	32	17,067	0,633	
3	—1	0	2	—1	0	30	19	12,033	0,446	
4	—1	0	0	1	0	10	1	0,100	0,004	

Всего: 705,200 26,151

Проверка вычисления: θ = 6,538-4 =26,152.

Как было уже сказано выше, применение латинского квадрата далеко не ограничивается чисто топографическими квадратами. Но и при чисто топографическом распределении серий опыта в поле расположение делянок не обязательно имеет форму квадрата. Предположим, например, что мы имеем продолговатый прямоугольный участок земли и, что тот или иной компонент гетерогенности поля (зараженность насекомыми, естественное плодородие) падает справа налево. Тогда при системе рандомизированных блоков мы разделим весь участок на четыре блока и в пределах каждого блока наметим по четыре делянки, причем расположение делянок возьмем продольное: этим путем мы постараемся в пределах каждого блока захватить все разнообразие условий и сделать, таким образом, делянки внутри блока по возможности однородными. Но можно постараться устранить гетерогенность поля не только между отдельными блоками, но и в пределах блока. Тогда, помня, что зараженность падает справа налево, мы разобьем делянку уже не вдоль, а поперек, чтобы создать максимальное различие внутри блока, но это максимальное различие не будет нам мешать, так как оно будет выделено в особую категорию благодаря организации опыта по схеме латинского квадрата. Мы получим, например, такой план участка:

B	*E*	*A*	*C*	*C*	*A*	*E*	*B*	*A*	*B*	*C*	*E*	*F*	*C*	*B*	*A*

Таким образом, здесь благодаря особенностям данного участка расположение по системе рандомизированных блоков более напоминает квадратное расположение, чем расположение по схеме латинского квадрата. Конечно, применение латинского квадрата по первой схеме не привело бы к неверным выводам, но меньшая часть изменчивости, связанной с гетерогенностью поля, была бы выделена от изменчивости, характеризующей ошибку опыта, и, следовательно, выводы могли бы оказаться менее четкими.

4.6. ГРЕКО-ЛАТИНСКИЙ И ВЫСШИЕ КВАДРАТЫ

Дальнейшее развитие идеи, лежащей в основе латинского квадрата, приводит к греко-латинскому и высшим квадратам. Эти схемы очень полезны в том случае, когда мы желаем одновременно исследовать действие большого числа факторов, но ввиду ограниченности имеющегося в нашем распоряжении времени, средств или участка земли не можем осуществить полную факториальную схему, позволяющую изучить не только различия, вызываемые отдельными факторами, но и все взаимодействия факторов. Отказавшись от изучения взаимодействия и ограничившись только прямыми контрастами, можно провести работу на гораздо меньшем материале, пользуясь схемой греко-латинского квадрата. Поэтому эта схема особенно полезна при проведении ориентировочных опытов для решения того, какой из различных, имеющихся в нашем распоряжении, факторов является наиболее существенным.

Греко-латинский квадрат выводится из некоторых латинских квадратов путем прибавления греческой буквы в каждую клетку, причем в отношении греческих букв должно быть соблюдено то же правило, как и для латинских, т. е. в каждой строчке и в каждом столбце должна быть по одному разу каждая греческая буква и притом каждая греческая буква должна сочетаться с каждой .латинской буквой.

Для греко-латинского квадрата 3-3 имеем следующее стандартное расположение:

A	*B*	*C*
B	*C*	*A*
C	*A*	*B*

Каждый из 1 латинских квадратов может быть преобразован -в греко-латинский квадрат, и так как греческие буквы могут быть перемещены шестью различными способами, то всего мы получим *72* греко-латинских квадрата 3-3. Этим путем все степени свободы будут уже исчерпаны, так как мы имеем:

строки	2 степени свободы
столбцы	2 » »
латинские буквы	2 » »
греческие буквы	2 » »
Всего	8 степеней свободы

Для греко-латинского квадрата 4-4 возможно кроме греческих букв введение еще индексов, так что 15 степеней свободы могут .быть разложены на пять наборов по три степени свободы каждый: каждая из таких систем (строки, столбцы, латинские буквы, греческие буквы, индексы) является ортогональной по отношению ко всем остальным, т. е., например, каждый индекс встречается во всех строках, во всех столбцах, при каждой латинской и при каждой греческой букве. Мы имеем такое исходное положение:

*A*1α	*B*2β	*C*3γ	*D*4δ
*B*4γ	*A*3δ	*D*2α	*C*1β
*C*2δ	*D*lγ	*A*4β	*B*3α
*D*3β	*C*4α	*B*1δ	*A*2γ

Путем перемещения строк, столбцов, латинских букв и т. д. можно получить всего 6912 греко-латинских квадратов 4-4.

Для латинского квадрата 5-5 можно также получить вполне ортогонализованный квадрат, используя уже два индекса и распределяя 24 степени свободы по следующим шести вполне независимым сериям по 4 степени свободы: строки, столбцы, латинские буквы, греческие буквы, первые и вторые индексы. Пример такого высшего квадрата приводим (всего имеется для 5-5 огромное число — несколько миллиардов решений):

*A*1α1	*B*2β2	*C*3γ3	*D*4δ4	*E*5ε5
*B*3δ5	*C*4ε1	*D*5α2	*E*1β3	*A* 2γ4
*C*5β4	*D*lγ5	*E*2δ1	A3ε2	*B*4α3
*D*2ε3	*E*3α4	*A*4β5	*B*5γ1	*C*1δ2
*E*4γ2	*A*5δ3	*B*lε4	*C*2α5	*D*3β1

Для латинского квадрата 6-6 имеется 9408 в стандартном положении, но ни один из них не дает греко-латинского квадрата. Все нечетные латинские квадраты дают греко-латинские квадраты, и для всех квадратов с простым числом бином $p(p^2—1)$ степеней свободы может быть разложен на p+1 независимых серий по p—1 степеней свободы каждая. Такие же сполна ортогонализованные квадраты могут быть образованы для случаев 8-8 и 9-9, они приведены в книге Р. Фишера (1937b).

Взяв для примера разобранный выше случай со сроками посева кукурузы и влиянием их на зараженность кукурузным мотыльком, можно сказать, что тот же опыт мог бы быть использован, например, для изучения еще следующих факторов: 1) греческие буквы — сорт кукурузы, 2) первые индексы — способ посадки (гнездами, поодиночке, на разном расстоянии), 3) вторые индексы — число рыхлении междурядий и другие методы обработки. Фактически этого не было, но мы обработаем полученные цифры как будто опыт был поставлен при полном ортогонализованном греко-латинском квадрате 5-5. Это позволит познакомиться с методикой подобных вычислений и даст одновременно сумму квадратов ошибки, ранее полученную как остаток путем непосредственного вычисления.

При закладке опыта по сполна ортогонализованному квадрату 5-го порядка естественно исходят из данной выше схемы путем последовательного рандомизированного перемещения строк, столбцов и сопоставления разных модальностей тех или иных факторов разным значениям букв и индексов. В нашем случае мы исходим из уже преобразованного латинского квадрата и, следовательно, должны с нашей исходной схемой проделать такое же преобразование. Нетрудно видеть, что латинский квадрат в работе Хервея и Хартцелля получен из диагонального квадрата

A	*B*	*C*	*D*	*E*
B	*C*	*D*	*E*	*A*
C	*D*	*E*	*A*	*B*
D	*E*	*A*	В	*C*
E	*A*	*B*	*C*	*D*

Путем следующих преобразований: 1) переменой мест 3-го и 4-го столбцов; 2) переменой мест 4-й и 5-й строк и 3) заменой букв на римские цифры: *A* — I, *B* — II, *C* — IV, *D* — III, *E* —V. Проделаем то же самое с исходной схемой сполна ортогонализованного квадрата, а затем, во избежание путаницы, заменим первые индексы латинскими буквами в алфавитном порядке. Тогда данные Хервея и Хартцелля (Hervey, Harzell, 1931) принимают такой вид:

I*A*α1 3	II*B*β2 8	III*D*δ4 4	IV*C*γ3 4	V*E*ε5 0
II*C*δ5 30	IV*D*εl 13	V*A*β3 9	III*E*α2 2	I*B*γ4 6
IV*E*β4 2	III*A*γ5 14	I*C*ε2 5	V*B*δ1 6	II*D*α3 25
V*D*γ2 0	I*E*δ3 4	IV*B*α5 4	II*A*ε5 11	III*C*β1 3
III*B*ε3 11	V*C*α4 2	II*E*γ1 16	I*D*β5 6	IV*A*δ2 2

В первой строчке каждой клетки помещены четыре значка, обозначающие принадлежность данной делянки к одной из пяти модальностей четырех изучаемых факторов, а во второй строчке — изучаемый признак, в данном случае число гусениц кукмота на 25 стеблей. Нетрудно убедиться, что расположение знаков сполна ортогонализовано, иначе говоря, соблюдена полная независимость размещения модальностей всех шести направлений изменчивости (считая кроме четырех факторов также строки и столбцы за направления изменчивости): именно каждая модальность любого направления комбинируется со всеми пятью модальностями любого из остальных направлений и нет ни одного повторения. Вычисление подтвердит правильность ортогонализации, так как только в случае полной ортогонализации получится совпадение сумм.

По столбцам, строкам и срокам посева суммы квадратов уже были вычислены ранее. Нам остается вычислить сумму квадратов по трем искусственно введенным факторам, обозначенным латинскими буквами, греческими буквами и арабскими цифрами. Для этого мы со всего квадрата собираем значения, соответствующие буквам *A* (это будет 3+9+14+11+2, или 39), *B* и т. д. Подобным же образом для греческих букв (3+2+25+4+2, или 36) и аналогично для арабских цифр. Получаем следующие суммы:

A	39	α	36	1	41
B	35	β	28	2	17
C	44	γ	40	3	53
D	48	δ	46	4	25
E	24	ε	40	5	54
Сумма	190		190		190

(все суммы, конечно, равняются 190).

Сумма квадратов для фактора обозначена латинскими буквами:

$$\frac{39^2+35^2+44^{+}+48^{+}+24^{+}}{5}$$ =1512,4 (поправка—1444,0).

Сумма квадратов от общей средней — 68,4.

Таким же образом получим для греческих букв:

Сумма квадратов .1479,2—1444,0, или 35,2, и для арабских цифр:

1664,0—1444,0, или 220,0.

Сумма 68,4+35,2+220,0 в точности равна 323,6. Она была получена ранее путем вычитания как суммы квадратов для 12 степеней свободы, соответствующих ошибке опыта.

При полной ортогонализации квадратов уже не остается ничего на долю ошибки опыта, и потому судить об абсолютной значимости различий модальностей исследуемых факторов мы уже не именем возможности, но сопоставление суммы квадратов для разных -факторов дает возможность оценить относительное значение факторов. В данном случае имеем:

	Число степеней свободы	Сумма квадратов
строки	4	259,6
столбцы	4	31,6
сроки посева	4	705,2
факт. латинских букв	4	68,4
факт. греческих букв	4	35,2
факт. арабских цифр	4	220,0
Всего:	24	1320,0

Так как число степеней свободы во всех случаях одинаково, то мы можем сравнивать непосредственно, не вычисляя среднего квадрата. Значимость сроков посева известна была ранее: сравнение крайних по значению сумм из числа остальных дает отношение 259,6:31,6 (минимальная сумма квадратов), равное 8,21, т. е. несколько больше теты (6,39) для минимального уровня значимости (Р равной 0,05) 4 и при 4 степенях свободы. Принимая во внимание чисто эмпирический характер такого отношения, утверждать о существенном различии изменчивости оставшихся пяти групп невозможно. Схема греко-латинского квадрата может с удобством применяться во всех случаях, где мы хотим выяснить относительное значение разных факторов с тем, чтобы после такого опыта перейти к детальному изучению.

4.7. ФАКТОРИАЛЬНАЯ СХЕМА ОПЫТА

Факториальная схема опыта, разработанная тем же Р. А. Фишером (1937а, 1958), представляет из себя наиболее совершенный метод постановки сложных опытов и отвечает целому ряду самых насущных запросов научной и производственной практики. Господствующим до недавнего времени требованием было: в каждом опыте сравниваемые варианты должны отличаться только по одному признаку, все остальные же условия должны быть тождественны. Этот принцип единственной разницы утверждался как совершенно обязательное условие каждого научного исследования, но уже большое количество научных работников высказывалось, что этот принцип не удовлетворяет современным запросам. Изучается, например, влияние температуры и влажности на развитие насекомых. Руководствуясь принципом единственного отличия,, авторы часть опытов строят при переменной температуре и постоянной важности, а часть при постоянной температуре и переменной влажности. Таким путем удается получить результаты о влиянии изолированных факторов, но биолога интересуют факторы не в их изолированном действии, а в их совокупном действии и во взаимодействии друг с другом.

Другой пример: некоторые опытные станции по нескольку лет изучали сначала разную ширину междурядий при той же норме высева, а затем разные нормы высева при той же ширине междурядий. Положим, что в обоих случаях получены оптимальные ширина междурядий и норма высева: значит ли, что комбинация этих оптимальных изолированных факторов окажется также оптимальной комбинацией: в этом совершенно нельзя быть уверенным, так как очень часто взаимодействие факторов смещают оптимумы при совместном изучении факторов.

Но помимо того что изучение взаимодействия наряду с изучением изолированного действия факторов диктуется самой природой исследования, само соблюдение «единственного различия» оказывается во многих случаях основанным на совершенно произвольных допущениях. Обратимся к указанному примеру с влиянием температуры и влажности: при разных температурах берут обыкновенно, соблюдая принцип единственного различия, одинаковую относительную влажность. Но предположение, что одинаковая относительная влажность при разных температурах действительно уравнивает фактор влажности, является недоказанным и даже оспариваемым некоторыми экологами. Ряд авторов утверждают, что одинаковая влажность достигается при уравнении «дефицита влаги», а не процента относительной влажности. Какой бы критерий мы ни брали, с определенной точки зрения то, что мы считаем одинаковой влажностью, таковой не является.

Возьму другой пример, имеющий огромное практическое значение. Производится сравнительное испытание сортов пшеницы. По принципу единственной разницы все сорта ставятся в совершенно одинаковые условия в отношении норм высева, глубины посева и т. д. Правильно ли это? Нетрудно показать, что это соблюдение единственной разницы и невозможно, и нежелательно. В самом деле, что означает одинаковая норма высева? Это выражение имеет целый ряд смыслов: одинаковое количество по весу на 1 га, или (как это применяется в Америке и Англии) одинаковый объем (бушели) зерна на 1 га, или (как это обычно принято в сортоиспытании) одинаковое количество зерен на единицу площади, причем и в этом последнем случае мы можем принимать в расчет процент всхожести семян и не принимать его и т. д. Все эти критерии дадут разные нормы. Какая же норма окажется правильной?

Очевидно, вести сортоиспытание нужно так, чтобы каждый сорт сравнивался при той норме, которая для него является оптимальной, а нормы по сортам, конечно, окажутся равными (имеется взаимодействие двух факторов: сортового различия и нормы высева). То же самое мы имеем и с глубиной заделки. Глубина заделки, очевидно, определяется двумя соображениями: чем глубже заделать, тем лучше семя использует влагу, но если заделать слишком глубоко, то зерну не хватит энергии пробить почву или это отразится на его росте. Отсюда понятно, что известно всем опытникам, что мелкие семена надо заделывать на меньшую глубину, чем крупные, и, сравнивая два сорта с разной величиной семян, надо каждый сорт сравнивать при оптимальной для него глубине заделки. Поэтому при сортоиспытании целесообразно сравнивать сорта на фоне меняющихся факторов (норма, глубина и т. д.). Этим мы, с одной стороны, избежим ошибочного вывода браковки хорошего сорта, испытанного не в оптимальных для него условиях по сравнению с сортом худшим, для которого условия испытания были оптимальными, а с другой стороны, сможем изучить взаимодействие факторов, входящих в опыт.

Но не будет ли такая постановка чрезмерным усложнением схемы, которое потребует значительного увеличения объема опыта? Нет, так как введение таких новых факторов может быть проведено за счет использования повторения опыта. Чем сложнее опыт, тем меньше может быть взята повторность, и, как мы покажем ниже, при опытах достаточной сложности можно вообще обойтись без повторности, в особенности если объединить материал по нескольким пунктам, что будет показано в особой главе.

Сущность факториальной схемы и заключается в том, что, взяв известное число подлежащих исследованию факторов, каждый в определенном числе модальностей, опыт ставят так, что испытываются все мыслимые комбинации исследуемых факторов. Например, в простейшем случае, если испытываются два фактора,. каждый из которых образует только две модальности, то опыт ставят со всеми четырьмя (2x2) возможными комбинациями обоих факторов. При наличии трех факторов, каждый в двух модальностях, получаем 8 комбинаций, иначе говоря, 8 вариантов опыта, при наличии двух факторов с тремя модальностями каждый получаем 3x3, или девять вариантов опыта и т. д.

Принцип единственного различия и при факториальной схеме,. и последующем факториальном анализе не устраняется вовсе, он только подымается на новую высоту. В самом деле, мы имеем, положим, опыт с влиянием трех факторов *А, В* и С, причем каждый фактор представлен всего двумя модальностями, которые для трех факторов и обозначим $a1$ и $a2$, $b1$ и $b2$, $c1$ и $c2$. Тогда, очевидно, поставив опыт с влиянием всех трех факторов по факториальной схеме, будем иметь восемь вариантов опыта: 1) $a1b1cl$, 2) $alblc2$, 3) $a1b2c1$, 4) $alb2c2$, 5) $a2blcl$, 6) $a2blc2$, *7)* $a2b2c1$, 8) $a2b2c2$. И для того чтобы выяснить влияние фактора *А*, мы. будем сравнивать варианты 1—4 (где фактор *А* представлен в виде $a1$) с вариантами 5—8 (где он представлен в виде $a2$); для сравниваемых серий по четыре варианта единственным различием будет различие по фактору *А,* так как по другим факторам в каждой серии будут совершенно тождественные наборы комбинаций значений факторов *В* и *С.* Но в каждой сравниваемой серии не будет тождества всех других факторов. Какое это может иметь влияние на результат? Если между фактором *А* и остальными факторами нет взаимодействия, иначе говоря, если действие фактора *А* проявляется совершенно независимо от того, в каком виде имеются другие факторы, то результат будет тот же,, как и при работе, соблюдая принцип «единственного различия», но с одним и довольно важным различием: испытав фактор *А* в разных комбинациях с другими факторами и убедившись, что во всех случаях он проявляет одинаковое действие, мы приобретаем уверенность в его универсальности и можем быть смелее в распространении наших выводов за пределы нашего опытного участка, чем в том случае, если весь опыт мы вели в однородных условиях.

Но, может быть, и часто бывает, что между факторами имеет место взаимодействие, например, удобрение при достаточной влажности вызывает прибавку урожая, при засухе, напротив, приведя к избытку солей в почве, может снизить урожай. При высокой температуре можно давать значительное искусственное орошение с благоприятным эффектом, при низкой температуре сезона, что бывает часто на севере, избыточное увлажнение может оказаться вредным и т. д. В этом случае, если мы ограничимся одним сопоставлением одиночно действующих факторов, мы можем получить или полную смазанность результатов (если в пределах опыта захвачены условия, соответствующие благоприятному и неблагоприятному проявлению данного фактора), или в разных опытах получить противоречие данных. Разрешение вопроса дается изучением взаимодействия факторов.

Если при сравнении модальностей фактора *А* мы сравнивали варианты 1—4 с вариантами 5—8, то для фактора *В* мы должны таким же образом сравнить варианты 1, 2, 5, 6 с вариантами 3, 4, 7, 8, так как обе эти серии отличаются между собой только по модальности фактора 5. Для фактора *С* мы сравниваем варианты 1, 3, 5 и 7 с вариантами 2, 4, 6 и 8.

Таким образом, один и тот же метод путем различной группировки может дать нам выводы в отношении изолированных факторов (в данном случае по трем разным факторам), которые при обычной работе методом «единственного различия» мы могли бы получить лишь путем постановки трех самостоятельных опытов, где изучалось бы действие каждого фактора в отдельности. Но опыт дает возможность оценить и размеры взаимодействия между факторами.

Возьмем сначала два фактора *А* и *В,* при двух модальностях получается четыре возможные комбинации: 1) $a1b1$, 2) $a1b2$, 3) $a2b1$, 4) $a2b2$. Чтобы получить сравнение изолированно по фактору *А,* мы должны, очевидно, сравнить варианты с модальностью $a1$ с вариантами с модальностью $a2$, получаем: $(a1b1+a1b2)—(a2bl+a2b2)=(al—a2)(b1—b2)$. Здесь, конечно, скобки означают не реальное умножение, а символическое — комбинирование разных модальностей.

Совершенно таким же образом для сравнения по фактору *В* получим

$$(a1b1+a2b1)—(a1b2+a2b2)=(al+a2)(b1—b2).$$

Как же надо алгебраически изобразить взаимодействие двух факторов? Взаимодействие имеет место в том случае, если действие фактора *А* в присутствии одной модальности фактора *В* не равно тому же действию в присутствии другой модальности фактора *В.* Но действие фактора *А* в присутствии модальности $b1$ фактора *В* выражается разностью $a1b1—a2b1$ и такое же действие в присутствии $b2$ разностью $a1b2—a2b2$. Если нет взаимодействия, иначе говоря, если действие фактора *А* не зависит от значения фактора *В,* то эти две разности показывают различия только порядка ошибки опыта. Если же взаимодействие есть, то разность $(a1b1—a2b1)—(a1b2—a2b2)$, или $a1b1—a2b1—a1b2+a2b2$ при сравнении с ошибкой опыта и покажет степень взаимодействия этих двух факторов. Но эту же разность можно символически переписать: $(a1—a2)(b1—b2)$.

Применяя тот метод вычисления, который был описан при разложении по степеням свободы, и сравнивая наши три степени свободы, получим такие выражения:

прямой контраст $A = a1b1 + a1b2 — a2b1 — a2b2$

$B=a1b1—a1b2+a2b1—a2b2$

взаимодействие $AB=a1b1—a1b2—a2b1+a2b2$

Или, выписывая одни коэффициенты, мы получим такое разложение по трем степеням свободы:

	$a1b1$	$a1b2$	$a2b1$	$a2b2$
A	1	1	-1	-1
B	1	-1	1	-1
AB	1	-1	-1	1

Мы видим, что коэффициенты для взаимодействия получаются попарным перемножением коэффициентов для соответствующих прямых контрастов.

Если мы имеем три фактора, каждый в двух модальностях, то сравнение по фактору A получим из такой разности:

$$(a1b1c1+a1b1c2+a1b2c1+a1b2c2)—(a2b1c1+a2b1c2+a2b2c1+a2b2c2)=(a1—a2)(b1+b2)(c1+c2).$$

Для фактора B получим таким образом: $(a1+a2)(b1—b2)\ (c1+c2)$. Для фактора C $(a1+a2)\ (b1+b2)\ (c1—c2)$.

Взаимодействий первого порядка (между двумя факторами будем иметь три различных сочетания из трех элементов по два, а именно:

$$AB=(a1—a2)(b1—b2)(c1+c2)$$
$$AC=(al—a2)(bl+b2)(cl—c2)$$
$$BC=(al+a2)(bl—b2)(cl—c2)$$

Но наряду со взаимодействием двух факторов здесь будет взаимодействие всех трех факторов, очевидно, единственно возможное, так как можно сделать только одно сочетание из трех элементов, если каждый элемент представлен в одной модификации. Что означает взаимодействие трех факторов? Это означает, что взаимодействие двух факторов происходит иначе в зависимости от того, в какой форме присутствует третий фактор. Так как взаимодействие двух факторов выражается такой формулой $a1b1—alb2—a2bl+a2b2$, то, очевидно, взаимодействие трех факторов измеряется так: $(a1b1—alb2—a2bl+a2b2)\ c1—(albl—alb2—a2bl+a2b2)\ c2$ или $(al—a2)(bl—b2)(c1—c2)$. В развернутом виде получаем выражение: $a1b1c1—a1b1c2—alb2cl+alb2c2—a2b1c1+a2b1c2+a2b2c1—a2b2c2$.

Знаки нетрудно писать сразу: члены, сумма показателей модальностей факторов которых нечетная (например, 1,1,1 или 2,2,1), — положительны, там, где четная — отрицательны.

Используя опять-таки общую схему для разложения по отдельным степеням свободы, мы получаем табл. 23.

Коэффициенты для каждой степени свободы выводятся из указанных выше сопоставлений (путем развертывания произведения в многочлен), но они также легко получаются в таблице перемножением коэффициентов двух соответствующих строк. Например, коэффициенты для ABC могут быть получены перемножением (попарным) коэффициентов для строк *A, B и C* или AB и С, или AC и B, или, наконец, BC и A.

Само собой разумеется, что набор коэффициентов при таких разложениях является ортогональным, и при подсчете суммы квадратов по отдельным степеням свободы мы получаем величину, в точности равную сумме квадратов для вариантов опыта.

Таблица 23

A	$a1$				$a2$			
B	$b1$		$b2$		$b1$		$b2$	
C	$c1$	$c2$	$c1$	$c2$	$c1$	$c2$	$c1$	$c2$
A	1	1	1	1	-1	-1	-1	-1
B	1	1	-1	-1	1	1	-1	-1
C	1	-1	1	-1	1	-1	1	-1
AB	1	1	-1	-1	-1	-1	1	1
AC	1	-1	1	-1	-1	1	-1	1
BC	1	-1	-1	1	1	-1	-1	1
ABC	1	-1	-1	1	-1	1	1	-1

Если все факторы представлены только двумя модальностями каждый, то, очевидно, каждый фактор имеет только одну степень свободы и разложение может быть единственным — все коэффициенты будут единицами, все различие будет в знаках. Если же тот или иной фактор представлен в нескольких модальностях, то для него разложение может быть произведено по степеням свободы различным образом, как это уже было указано в главе о разложении по степеням свободы. Положим, имеем два фактора, один, представленный в двух модальностях, а другой — в трех. Тогда для изолированных факторов, естественно, разложение будет вестись так, как если бы другой фактор не имел никакого влияния. Для фактора A, имеющего две модальности, мы получим сопоставление

$$(a1—a2)\ (b1+b2+b3),$$

Для фактора B, представленного в трех модальностях, имеем две степени свободы. Обычно противопоставляются крайние варианты и выясняется прямолинейность зависимости. Получаем, следовательно, два символических выражения:

$$B1=(a1+a2)(b1—b3)$$
$$B2=(a1+a2)\ (B1—2b2+b3)$$

Как следует понимать взаимодействие между факторами A и B? Очевидно, взаимодействия будут иметь место для двух степеней свободы, которые и обозначим соответственно $AB1$ и $AB2$. Для $AB1$ — взаимодействие совершенно того же характера, как и в предыдущем случае: просто наличие модальности $b2$ не принимается во внимание. Для взаимодействия же $AB2$ имеем сопоставление:

$$(b1—2b2+b3)a1—(b1—2b2+b3)a2=(a1—a2)(b1—2b2+b3),$$

или в развернутом виде: $a1b1—2b1b2+a1b3—a2b1+2a2b2—a2b3$.

Получаем для данного случая табл. 24 коэффициентов для разложения по пяти степеням свободы.

Как и всегда, коэффициенты для взаимодействия получаются перемножением соответственных коэффициентов для прямых контрастов. Это правило распространяется на взаимодействия любых порядков. Поэтому при сложных опытах достаточно наметить коэффициенты по прямым контрастам (если, конечно, имеются факторы более чем в двух модальностях, при двух модальностях соответствующий контраст является единственно важным), а коэффициенты для всех остальных степеней свободы получаются простым перемножением этих основных.

Таблица 24

A	*a*1			*a*2		
B	*b*1	*b*2	*b*3	*b*1	*b*2	*b*3
A	1	1	1	-1	-1	-1
*B*1	1	0	-1	1	0	-1
*B*2	1	-2	1	1	-2	1
*AB*1	1	0	-1	-1	0	1
*AB*2	1	-2	1	-1	2	-1

Если мы, положим, имеем 4 фактора, каждый в двух модальностях, и 4 повторности опыта, значит, 16 вариантов и 64 даты, то анализ дисперсии будет иметь такой вид:

действия отдельных факторов	4 степени свободы
взаимодействия двух факторов	6 » »
» трех »	4 » »
» четырех »	1 » »
повторности (блоки)	3 » »
ошибка	45 » »
Всего:	63 » »

Если опыт не разбивался по блокам, то мы получим 48 степеней свободы для ошибки. При организации опыта по методу единственного различия для получения той же точности результатов необходимо для каждого сравнения получать 50 дат, так как тогда сравнение двух средних будет основано тоже на 48 степенях свободы. Так как мы имеем четыре фактора, то, следовательно, для получения выводов из сравнения одних изолированных факторов надо иметь 200 дат против 64 дат в нашем опыте. При этом эти 200 дат не дадут нам никакого понятия о характере взаимодействия факторов, для этого потребуется организовать специальные опыты. Мы видим, таким образом, что факториальная схема не только дает чрезвычайно удобный метод для суждения о взаимодействиях, но и крайне экономизирует опыт. Если мы ставим опыт с тремя факторами, каждый в трех модальностях, то получаем всего 27 комбинаций — 27 вариантов опыта. Они будут разбиты на следующие категории:

действия отдельных факторов (*A*l, *A*2, *B*1, *B*2, *C*1, *C*2)	6 степеней свободы,
взаимодействия двух факторов (*A*1*B*1, *A*1*B*2, *A*1*C*1, *A*1*C*2, *A*2*B*1, *A*2*B*2, *A*2*C*1, *A*2*C*2, *B*1*C*1, *B*1*C*2, *B*2*C*1, *B*2*C*2)	12 » »
взаимодействия трех факторов (*A*1*B*1*C*1, *A*1*B*1*C*2, *A*1*B*2*C*1, *A*1*B*2*C*2, *A*2*B*1*C*1, *A*2*B*1*C*2, *A*2*B*2*C*1, *A*2*B*2*C*2)	8 » »
Всего:	26 » »

Чем больше количество факторов исследуется одновременно, тем сложнее оказывается опыт и тем более можно сократить число повторностей. Если мы в последнем опыте с 27 комбинациями трех факторов возьмем только двукратную повторность, то получим 27 степеней свободы для суждения о размерах ошибки опыта. Дисперсия, соответствующая этим 27 степеням свободы, может быть вычислена или, как это обычно делается, по разности дисперсий для всех 54 дат и для 27 вариантов, или путем суммирования 27 квадратов попарных разностей тех же вариантов двух повторностей. Для того чтобы без помощи факториального анализа получить такую же надежность для каждого сравнения, работая по методу «единственного различия», надо было бы поставить 27 отдельных опытов с 28 испытаниями (чтобы получить 26 степеней свободы для ошибки каждого сравнения), т. е. получить в 14 раз больше дат, чем при опыте, поставленном по факториальной схеме. При очень сложных опытах можно вообще обойтись без повторности, используя для суждения об ошибке взаимодействия высших порядков. Положим, мы имеем опыт с шестью факторами, каждый в двух модальностях. Делая все возможные комбинации, получаем 64 варианта, 63 независимых сравнения нашего опыта могут быть классифицированы таким образом:

отдельные факторы	6 степеней свободы
взаимодействия двух факторов	15 » »
» трех »	20 » »
» четырех »	15 » »
» пяти »	6 » »
» шести »	1 » »
Всего:	63 » »

Объединив взаимодействия высших порядков, мы получим 42 степени свободы для суждения о надежности наших выводов. Такой прием основан на том, обычно оправдывающемся предположении, что взаимодействия трех и больше факторов обычно не являются существенными, но в каждом данном опыте мы можем проверить справедливость этого основного положения. Для этого следует исследовать взаимодействия первого порядка (между двумя факторами). Если окажется, что ни одно из взаимодействий первого порядка не обнаруживает значительной величины, то это показывает на слабую связь между исследуемыми факторами и мы можем принять, что наш прием вполне оправдался. Если, положим, одно из взаимодействий между двумя оказалось существенным, то тогда надо вычислить взаимодействия третьего порядка, в которые входят два фактора, обнаружившие взаимодействие: таких будет, очевидно, четыре (к двум взаимодействующим факторам прибавлять каждый из четырех оставшихся). Сравнивая эти контрасты с суммой квадратов для остальных 38 степеней свободы, сможем заключить о наличии существенных взаимодействий второго порядка. Опыт такого рода конечно неприменим, если имеется большое число взаимодействий высшего порядка, по существенности превышающих первичные эффекты но такие случаи встречаются редко, и, как было уже указано, при обработке данных опыта мы можем заключить, что опыт был вставлен по непригодной для данного случая схеме (без повторностей).

Сокращение и полное устранение повторностей в сложных опытах вовсе не означают отказ от самого принципа повторности абсолютно необходимого в каждом опыте. Абсолютная, явная повторность — повторение совершенно тождественных вариантов опыта — заменяется повторностью открытой, состоящей в наличии сходных признаков у несходных вполне вариантов Эта скрытая повторность и используется в полной мере путем факториального анализа. Факториальная схема организации опыта и факториальный анализ обладают, следовательно, следующими крупны-ми преимуществами по сравнению с господствовавшей ранее схемой «единственного различия».

1. Этим путем мы получаем возможность оценивать факторы не только в их изолированном действии, но и во взаимодействии.

2. В организации опыта достигается огромная экономия, так как один и тот же материал служит для оценки большого числа независимых сравнений.

3. Благодаря постановке опыта в разнообразных условиях каждый вывод получает гораздо большую область приложения, чем при работе в узких рамках «единственного различия».

4. Число повторностей может быть сокращено за счет включения в исследование второстепенных факторов: каждый из основных интересующих нас признаков (например, сортовые различия и т. д.) может быть сравнен с другими, конкурирующими, не «при прочих равных условиях», а в условиях, наиболее благоприятствующих каждому из сравниваемых сортов или приемов работы и т. д.

5. Наконец, обработка материала, как вообще при анализе дисперсии, но в особенности при факториальном анализе допускает чрезвычайно много приемов независимой проверки: этим путем мы убеждаемся в том, что не допустили в вычислениях никакой ошибки.

Все это показывает, что высказываемое нередко в литературе мнение, что теория малых выборок (на которой целиком покоится весь дисперсионный анализ) приводит к требованию чрезмерного увеличения числа повторностей, целиком основано на недоразумении. Такое увеличение повторностей необходимо лишь для изолированных сопоставлений, основанных на каком-либо единственном измерении. Как только получается система дат или увеличивается число исследуемых признаков, то количество повторностей, необходимых для надежного вывода, не увеличивается по сравнению с обычными требованиями, а уменьшается вплоть до возможности полного ее устранения.

В качестве иллюстрации использования факториальной схемы приведу данные Е. В. Карлаш по влиянию различных факторов на длину размотанной нити у китайского дубового шелкопряда. Исследовалось влияние кормового растения (мягкий и жесткий дуб) и освещения. Кроме того, естественно, надо было принять в соображение пол куколки. Пол обозначаем буквой *А*, освещение — *В,* породу дуба — С (схема трехфакторного эксперимента). Так как две повторности, то всего было получено 16 цифр. Каждая цифра представляет собой среднюю из величин для нескольких индивидов данного садка. Количество индивидов колебалось, поэтому вес всех 16 цифр неодинаков. Это обстоятельство, конечно, увеличивает случайную ошибку, но не вводит систематической, при условии, конечно, если не было селективной смертности для отдельных вариантов: исходное количество гусениц во всех садках было одинаковым, но благодаря смертности и неодинаковому числу представителей обоих полов произошла расхождение. При исследовании, очевидно, предполагается, что погибшие и уцелевшие экземпляры не отличаются друг от друга по исследованному признаку, приводим исходные данные (табл. 25).

Таблица 25

А	Самцы				Самки				Сумма
В	освещ.		затенен.		освещ.		затенен.		
С	мягк.	жестк.	мягк.	жестк.	мягк.	жестк.	мягк.	жестк.	
1 повт.	422	382	698	384	533	280	546	320	3565
2 повт.	400	275	418	434	478	244	495	361	3105
Сумма	822	657	1116	818	1011	524	1041	681	6670

Производим сначала общий анализ дисперсии.

Сумма квадратов исходных дат равна 2983744,00 (поправка $\frac{6670^2}{16} = 2780556.25\cdot$

Общая сумма квадрата от общей средней — 203187,75.

У нас имеется всего 8 вариантов. Получаем сумму квадратов от нуля:

$$\frac{822^2 + 657^2 + 1116^2 + 818^2 + 1011^2 + 524^2 + 1041^2 + 681^2}{2} = 2933026{,}00 \text{ (поправка—2780556,25).}$$

Сумма квадратов вариантов от общей средней — 152469,75.

Таким же образом, возводя в квадрат 3565 и 3105, деля сумму квадратов на 8 и вычитая ту же поправку, получаем для повторностей 13225,00. В результате получаем следующий анализ варианcы (табл. 26).

Таблица 26

Категории изменчивости	Число степеней свободы	Сумма квадратов	Средний квадрат	θ	*Р*
Варианты	7	152469,75	21781,39	4,067	<0,05
Повторн. (блоки)	1	13225,0	13225,00		
Ошибка	7	37493,00	5356,14		

Всего	15	203187,75

Мы видим, что средний квадрат для вариантов удовлетворяет низшему уровню значимости (для 6 и 8 степеней свободы большей вариансы и 7 степеней меньшей, для Р, равной 0,05, достаточна тета соответственно 3,87 и 3,73), для отдельных степеней свобод можно ожидать более существенных контрастов.

Ввиду того что здесь имеются три фактора, каждый в двух модальностях, то система наборов ортогональных коэффициентов имеет, очевидно, вид, приведенный на с. 123. Перемножая для каждой степени свободы суммы для каждого варианта на соответствующий коэффициент, складывая, получаем разницы соответственно каждому контрасту. Делитель во всех случаях будет общий, так как коэффициенты все — единицы, отличающиеся только знаками, и так как число повторностей — 2, то общий делитель равен 16. Сумма квадратов разностей для всех степеней свободы равна сумме квадрата для вариантов, вычисленной ранее (табл. 27).

Таблица 27

Контраст	Разность δ	δ^2	$\frac{\delta^2}{16}$	θ	*P*
A	156	24 336	1521,00	0,28	
B	—642	412 164	25760,25	4,81	ок. 0,05
C	1310	1 716 100	107256,25	20,02	<0,01
AB	—268	71 824	4489,00	0,84	
AC	—384	147 450	9216,00	1,72	
BC	—6	36	2,25	0,00	
ЛBC	—260	67 600	4225,00	0,79	
Сумма		2 439 516	152469,75	28,46 /	

Сумма теты проверяется путем сравнения с тетой вычисленной для среднего квадрата вариантов, умноженного на число степеней свободы: 4,067·7=28,479.

Из всех семи сопоставлений серьезную значимость показывает только сопоставление по породе дуба (фактор *С):* для вероятности отсутствия разницы, равной 0,01, достаточна (при 1 и 7 степенях свободы) тета, равная 12,25, а для 0,001—29,22. Контраст по освещению не достигает даже низшего уровня значимости (Р — 0,05), но не очень далек от него (тета — 5,59).

В данном случае, так как делители для всех степеней свободы одинаковы, можно и не проделывать вычисления $\frac{\delta^2}{\Delta}$ и тета для всех случаев. Можно ограничиться вычислением просто квадрата разности, что и сделано в табл. 27. Проверка производится делением суммы этих квадратов 2439516 на общий делитель 16, что и дает в точности 152469,75. Для того же, чтобы оценить значимость различий, умножаем наш средний квадрат ошибки на тета, соответствующую трем уровням значимости (табл. 28).

Таблица 28

Уровень значимости	θ для 1 и 7 степеней свободы	5356,14
0,05	5,59	29941
0,01	12,25	65613
0,001	29,22	156506

С этими величинами сравниваем непосредственно квадраты разниц, деленные на 16, и приходим, конечно, к тем же выводам. При большом числе степеней свободы этот прием предпочтительнее вычислений теты по отдельным степеням свободы. В данном случае можно избежать деления на 16; тогда надо полученные нами величины теты помножить на 16, т. е. 5356,14·16.

При интерпретации результатов надо никогда не забывать, чего именно касается исследуемое нами сопоставление. В данном случае мы исследуем влияние взятых факторов на длину размотанной нити, но отсюда было бы неправильно сделать заключение о влиянии этих факторов на длину нити вообще, так как длина размотанной нити зависит в данном случае не только от длины нити вообще, но и от способности разматывания нити: длинная, но легко обрывающаяся нить может показать меньшую длину, чем разматывающаяся короткая.

Второе, на что следует обратить внимание, это на постоянное различие теоретических и эмпирических сопоставлений. Если мы подходим к материалу без всяких теоретических предпосылок, заставляющих нас ожидать определенное влияние того или иного фактора или взаимодействия факторов, то степень значимости сопоставлений будет ниже, чем при наличии теоретически ожидаемых и обнаруженных в опыте различий. Предположим, что мы исследуем 6 факторов, каждый в двух модальностях: всего будет 64 варианта или 63 степени свободы. Ясно, что даже при полном отсутствии какого-либо реального отличия между вариантами мы, имея как бы 63 независимых испытания, вправе ожидать появления различий соответствующей вероятности 0,05 или даже 0,01, без того, чтобы мы имели основание видеть в этих чисто эмпирических различиях что-либо существенное.

4.8. ФАКТОРИАЛЬНЫЙ ДИСПЕРСИОННЫЙ АНАЛИЗ

Другим примером факториального анализа возьму данные 11-летних опытов со сроками посева озимых пшениц на Одесском опытном поле. Данные приводятся А. А. Сапегиным (Константинов, 1939, с. 37). Опыты велись в двух повторениях (I и II), и испытывались сроки посева — август и сентябрь. Урожаи приводятся в пудах (исходные цифры округлены до пуда). Привожу данные в табл. 29.

Таблица 29

Год	Повторн.	Август *x*	Сентябрь *y*	Год	Повторн.	Август *x*	Сентябрь *y*
1897	I	41	18	1903	I	142	135
	II	32	20		II	130	125
1898	I	131	116	1904	I	41	57
	II	88	99		II	9	60
1899	I	62	73	1905	I	86	58
	II	64	82		II	93	57
1900	I	86	89	1906	I	167	143
	II	66	81		II	139	131
1901	I	125	102	1907	I	105	54
	II	134	97		II	97	46
1902	I	199	171	сумма	I	1186	1016
	II	180	163		II	1032	961
				Общая сумма		218	1977
					Средняя	100,81818	89.86364

Общая сумма для всего опыта — 4195, общая средняя — 95,3409

Простое рассмотрение цифр показывает, что хотя большей частью августовские сроки посева дают лучший урожай, чем сентябрьские, но бывает и обратное отношение. Поэтому необходимо исследовать, являются ли такие колебания следствием того, что просто нет существенной разницы между урожаем этих двух разных сроков посева или же имеет место взаимодействие между условиями года и сроком посева? Простое рассмотрение цифр заставляет нас опять-таки склониться в пользу второго предположения, так как уже грубое сравнение цифр позволяет сделать два существенных вывода:

1) только в одном году (1898 г.) повторности дают разные показания (I повторность — преимущество августовского срока, II — сентябрьского), в остальные годы показания повторностей совпадают: в семи случаях августовский посев дает лучший урожай в обоих повторностях, в трех — сентябрьский;

2) случаи превышения урожая сентябрьского срока относятся к годам с сравнительно низкими урожаями, что опять-таки дает намек на наличие взаимодействия.

При таком распределении и при наличии резких колебаний урожая по годам неудивительно, что если мы будем просто сравнивать средние по срокам и вычислим их разницу со средней ошибкой (основанной на ошибке обоих сроков посева), то получим: 10,95454+13,9935, что дает *t,* равную 0,78283 — абсолютно никакого намека на разницу. Само собой разумеется, что этот результат может объясняться тем, что мы для определения средней ошибки смешали несколько категорий изменчивости; 1) по годам; 2) по повторностям; 3) вследствие взаимодействия условий каждого года и срока посева; 4) собственно ошибку, т. е. результат действия всех неучитываемых нами факторов. Для того чтобы разделить общую дисперсию по всем этим категориям изменчивости, проделаем соответствующий анализ. Сначала определим общую изменчивость. Для этого, как всегда, возведем все наши даты (41, 32, 131 и т. д., всего 44 цифры) в квадрат и сложим. Получим сумму квадратов для августовских севов 275650 и для сентябрьских — 216093 (эти суммы используются нами и при непосредственном вычислении ошибки), всего 491743.

Поправка $\frac{4195^2}{44}$ или 399955,113.

Получаем общую сумму квадратов от общего среднего 91787,887. По вариантам (срок посева) получаем

$$\frac{2218^2+1977^2}{22}-399955{,}113\text{, или } 1320{,}022.$$

По годам берем 11 цифр, каждая из которых составляет сумму четырех дат для года (получаем ряд: 111, 434, 281, 322 и т. д.) и вычисляем:

$$\frac{111^2+434^2+281^2+322^2}{4}-399955{,}113 \text{ или } 81876{,}131.$$

(Общее правило: берем квадраты сумм и делим общую сумму квадратов на число дат, послуживших для получения каждой суммы).

По повторностям оперируем таким же образом с суммами дат для обоих повторностей: для первой получаем 1186+1016, или 2202, для второй 1032+961, или 1993.
Получаем

$$\frac{2202^2+1993^2}{22}-399955{,}113 \text{ или } 992{,}750.$$

Теперь вычислим дисперсию, соответствующую взаимодействие между годом и сроком посева. Так как сроки посева дают одну степень свободы, а годы — 10, то взаимодействие соответствует тоже 1x10, или десяти степеням свободы. Можно, конечно, вычислять дисперсию на каждую степень свободы отдельно и потом получить сумму, но в данном случае нас интересует лишь вопрос о наличии факта взаимодействия, а не детальное вычисление. Поэтому ограничимся вычислением общей суммы сразу. Это можно делать двумя путями: одновременное вычисление обоими способами дает, как всегда, хороший контроль вычислений.

По первому способу составим ряд цифр, показывающих разность урожаев августовского и сентябрьского сроков посева для обоих повторностёй сразу. Получаем цифры для последовательных лет: (41+32)—(18+20)=+35; (131+88)—(116+99) =+4 и т. д. Получаем ряд: +35, +4, —29, —18, +61, +45, +12, —67, +64, +32, +102. Общая сумма +241 характеризует общий перевес урожаев августовского срока над сентябрьским и равна, очевидно (контроль), разности сумм августовских и сентябрьских сроков посева 2218—1977=241.

Если нет взаимодействия между сроками посева и годом, то колебания разности урожаев для всех лет вокруг общей средней будут порядка случайной ошибки: если есть взаимодействие, то дисперсия будет существенно превышать случайную. Поэтому суммируем квадраты всех 11 полученных цифр: +35, +4, —29, —18, +61 и т. д. Сумму квадратов делим на 4 (число дат, послуживших для определения каждой цифры) и из полученного частного (который, очевидно, представляет собой сумму квадратов отклонений средних величин по годам от нуля) вычитаем поправку, равную $\frac{241^2}{44}$. Получаем дисперсию, соответствующую всем десяти степеням свободы взаимодействия годов и сроков,. равную 7077,250—1320,023=5757,227.

Другой способ заключается в следующем. Так как у нас имеется два срока и 11 лет продолжительности опыта, то всего получаем 22 варианта, дающих 21 степень свободы. Из этих 21 степеней свободы одна, которую мы уже вычислили, соответствует срокам посева, 10 других — годам (тоже уже вычислена) и оставшиеся 10 — взаимодействию срока посева и года. Следовательно, надо вычислить суммарную дисперсию, соответствующую всем 21 степеням свободы и из нее вычесть две уже вычисленные нами дисперсии.

Цифры для 22 вариантов будут: 41+32=73; 18+20=38; 131+88=219 и т. д., возводим все эти 22 цифры в квадрат, суммируем, делим на 2 и из частного вычитаем обычную поправку 399955,113, получаем 88953,387, вычитая дисперсию по срокам (1320,022) и по годам (81876,137), получим остаток (взаимодействие) — 5757,228, совпадающий, как и следует, с ранее вычисленным значением (разница — единица в последней цифре объясняется, конечно, неизбежным округлением в последнем знаке). Вычтя из общей дисперсии величины для четырех категорий изменчивости, уже выясненные, получим дисперсию, соответствующую ошибке, и мы получаем такой анализ (табл. 30).

Таблица 30

Категории изменчивости	Число степеней свободы	Сумма квадратов	Средний квадрат	θ	*P*
А срок посева	1	1320,022	1320,022	15,08	<0,001
В годы	10	81876,137	8187,614	93,36	<0,001
С повторности	1	992,750	992,750	11,32	<0,01
АВ взаимодействие лет (сроков)	10	5757,228	575,723	6,56	<0,001
Ошибка	21	1841,750	87,702		
Всего	43	91787,887			

Мы видим, что результат получился исключительной отчетливости. Разумеется, на первом плане стоит изменчивость, связанная с годами. Так как при 10 степенях свободы большей вариансы и 21 меньшей для получения вероятности отсутствия существенной разницы, меньшей 0,001, достаточна тета, равная приблизительно 5,0, то величина 93,36 показывает совершенно исчезающе малую вероятность отсутствия разницы по годам. Вычисление здесь делалось, конечно, не для того, чтобы еще раз убедиться в этой общеизвестной истине, а для того, чтобы исключить дисперсию, связанную с этим самым сильным источником изменчивости. И мы видим, что при правильной обработке этот источник изменчивости отнюдь не спутал данных для суждения о других источниках изменчивости. Для сроков посева (теты для Р, равной 0,01 и 0,001, соответственно равны 8,02 и 14,62) мы видим высокую надежность различия (Р меньше 0,001), для повторностей тоже высокую (Р лежит между 0,01 и 0,001) и для взаимодействия мы также имеем полную уверенность в существенности этого явления.

Мы видим, что наличие минимальной (двукратной) повторности совсем не является препятствием для получения надежных выводов, основанных на теории малых выборок, как думают некоторые авторы, основываясь на непонимании методов применения этой теории. В качестве курьеза не могу не отметить, что именно данный случай, давший, как видим, исключительные по своей четкости результаты, использован П. Н. Константиновым (1939) в качестве примера «неудовлетворительных опытных данных в силу плохой браковки». Без всякой браковки получаются вполне отчетливые результаты.

На этом разобранном примере постараемся посмотреть, чему соответствуют в проведении дисперсионного анализа разные степени изменчивости, что соответствует случаю, указанному вначале, когда путем простого сравнения средних мы получили как будто полное отсутствие существенной разницы между сроками посева. Этому соответствует объединение в одну категорию всех источников изменчивости, кроме срока посева (табл. 31).

Таблица 31

Категории изменчивости	Число степеней свободы	Сумма квадратов	Средний квадрат	θ
Срок посева	1	1320,022	1320,022	0,6128
Прочие	42	90467,865	2153,997	
Всего	43	91787,887		

Как видим, при объединении всех источников изменчивости мы получили такой же неудовлетворительный результат, как и при простом сравнении средних. Вернее, результат тождественный, так как тогда мы определяли *t* — отношение разности к своей средней ошибке, а здесь мы определяем θ — отношение квадратов разности к ошибке. Следовательно, в данном случае, когда число степеней свободы для вариантов равно единице, тета в точности должна быть равна t^2. Так оно и получается:

$$0{,}78283^2=0{,}6128.$$

Заметим, что при вычислении средней ошибки в начале главы были использованы вычисления при дисперсионном анализе. Квадрат средней ошибки для августовского срока равен

$$\frac{275650 - \frac{2218^2}{22}}{462} = 112{,}630,$$

для сентябрьского срока $\frac{216093 - \frac{1977^2}{22}}{462} = 83{,}187.$

Отсюда средняя ошибка разности равна

$$\sqrt{112{,}630 + 83{,}187} = \sqrt{195{,}817} = 13{,}9935$$

Возьмем теперь менее грубое определение ошибки: выделим изменчивость, связанную с годами и повторностями, но взаимодействие лет и сроков посева соединим с ошибкой, получим новую таблицу (табл. 32).

При соединении взаимодействия лет и сроков с ошибкой результат, что вполне понятно, чрезвычайно потерял в своей убедительности. Осталось резко выраженным различие по годам, но это совсем неинтересно, так как это общеизвестно, по срокам же посева получили вывод, лишь немного возвышающийся над минимальным уровнем значимости (для Р, равной 0,05, тета нужна 4,17, для 0,01—7,55). Естественно, при таком подходе можно сказать, что опыт не дал вполне отчетливых результатов.

Таблица 32

Категории изменчивости	Число степеней свободы	Сумма квадратов	Средний квадрат	θ	Р
Срок посева	1	1320,022	1320,022	5,38	<0,05
Годы	10	81876,137	8187,614	33,40	<0,001
Повторности	1	992,750	992,750	4,05	ок. 0,05
Прочие	31	7598,978	245,128		
Всего	43	91787,887			

На этом .примере мы ясно видим, что дисперсионный анализ не представляет какого-то насилия над материалом; стремления путем математических выкладок получить из материала вывод, вовсе не вытекающий из него. Напротив, и этот метод, как все математические приемы, при правильном применении является методом, позволяющим получить надежный вывод и там, где на глаз мы не вполне уверены в надежности: это и есть обычный здравый смысл, только облеченный в точную форму. Отсюда понятно, что приводимые часто примеры, где при биометрической обработке как будто не получилось надежности там, где всякому ясна разница, основаны во всех без исключения случаях на недоразумениях: авторы не сумели правильно использовать методы обработки данных.

4.9. ОБЪЕДИНЕНИЕ РЕЗУЛЬТАТОВ СЛОЖНЫХ ОПЫТОВ, ПРОВЕДЕННЫХ В НЕСКОЛЬКИХ ПУНКТАХ

Очень многие исследования проводятся по единообразной схеме в целом ряде пунктов, и *конечной задачей исследования является всегда синтез всех этих данных*, а не простое сопоставление изолированных результатов.

Такой синтез в полном виде предполагает по крайней мере три элемента: 1) использование повторяющихся элементов в каждом отдельном исследовании для получения большей надежности общего вывода по сравнению с выводами отдельных пунктов; 2) установление закономерной зависимости изучаемого признака с географическими координатами, климатом, почвой и т. д.; 3) на основе установления такой зависимости построения по данным сравнительно небольшого числа конкретных точек общей карты для данного признака, т. е. решения задач интерполяции.

Такой синтез следовало бы, например, получить в отношении сроков посева по данным отдельных опытных станций, по урожайности разных сортов злаков применительно к климатическим и почвенным условиям и т. д. Как правило, синтез не дается и отдельные опытные станции ограничиваются просто сообщением собственных эмпирических данных.

Я постараюсь показать пример такого объединения в отношении первых двух разделов указанной задачи (т. е. оставляя в стороне вопрос интерполяции) на материале по фумигации почвы парадихлорбензолом в керосине против личинок майских жуков. Данные мне любезно сообщены М. П. Войтенко. Опыты ставились в пяти пунктах Украины по совершенно одинаковой схеме. Именно испытывались: А) три дозировки ПДБ: 48,72 и 96 л/га; В) две сетки уколов, т. е. число уколов на 1 м^2 — 4 и 16; С) две глубины-нанесения яда: 10 и 20 см.

Всего таким образом получалось 3·2·2, или 12 вариантов опыта, не считая полного контроля. Эффективность определялась во всех случаях сравнением числа живых личинок на участке в 4 м^2 на контроле и обработанных участках. Во всех случаях эффективность оказалась вполне доказанной для всех вариантов в целом, хотя и была далеко от удовлетворительной (размеры ее выражались 53—68%). Нас в первую очередь интересует вопрос о различии между вариантами, которое в виду чрезвычайной изменчивости числа найденных личинок установить оказалось нелегко. Поэтому для исследования мы возьмем сначала только материал по 12 вариантам опыта. В каждом из пяти пунктов было проведено четыре повторности (четыре рандомизированных блока), и, следовательно, в каждом пункте мы имели по 48 дат, а всего 240 дат.

Приведем сначала оригинальные данные и анализ дисперсии (табл. 33—37) по каждому из пяти пунктов, а именно: 1) Черкассы, почва — легкий суглинок; 2) Носковцы Винницкой обл., около Жмеринки, — почва — тяжелый суглинок; 3) Чабаны — хозяйство Института плодоводства, около Киева, почва — средний суглинок; 4) совхоз им. Тельмана, Царычанского района,. Днепропетровской обл., почва — легкий суглинок и 5) совхоз Крыныца около Лохвицы Полтавской обл., почва — средний суглинок.

Во всех таблицах *А* — означает дозировку, *В* — сетку, *С* — глубину.

Рассмотрение всех этих данных приводит нас к следующим выводам:

1) только в Черкассах имеется некоторое указание на наличие существенных различий между вариантами, в остальных четырех пунктах тета для вариантов так мала, что никакого намека на значимость различий между вариантами по каждому пункту порознь мы усмотреть не можем;

2) для повторностей различия более выражены, правда, только в одном пункте (Чабанах) намечается очень существенное различие между повторностями, но в остальных четырех местах тета для повторностей, не достигая минимального уровня значимости, все таки много больше теты для вариантов.

Разложим теперь сумму квадратов для вариантов по отдельным степеням свободы и проделаем это подробно на черкасском материале, а потом и для остальных четырех пунктов. Это следует сделать, во-первых, для того, чтобы выяснить, какие же контрасты являются значущими для Черкасс; во-вторых, для того, чтобы выявить и в других пунктах те или иные значащие контрасты, которые могли затеряться среди большого числа совершенно несущественных; в-третьих, для того, чтобы подготовить синтетический вывод по сообщенному материалу пяти пунктов, и, наконец, для проверки вычислений: сумма квадратов должна совпасть с вычисленной ранее.

Таблица 33

Черкассы

A		48				72				96				Сумма
B		4		16		4		16				16		
C		10	20	10	20	10	20	10	20	10	20	10	20	
Повторности	1	4	6	2	3	1	5	0	0	1	3	2	4	31
	2	3	1	4	4	0	4	1	2	3	0	2	1	25
	3	1	3	2	3	3	2	0	0	2	2	0	1	19
	4	4	5	1	0	1	2	0	0	3	3	3	2	24
Всего		12	15	9	10	5	13	1	2	9	8	7	8	99

Анализ дисперсии

Категории изменчивости	Число степеней свободы	Сумма квадратов	Средний квадрат	θ	P
Варианты	11	47,5625	4,3233	2,26	0.05
Повторности	3	6,0625	2,0209		0,05
Ошибка	33	63,1875	1,9148		
Всего	47	116,8125			

Таблица 34

Носковцы

A		48				72				96				Сумма
B		4		16		4		16				16		
C		10	20	10	20	10	20	10	20	10	20	10	20	
Повторности	1	14	7	1	4	0	18	9	12	8	3	1	1	78
	2	5	4	6	5	0	6	2	5	1	0	4	0	38
	3	7	6	9	6	3	1	6	1	2	3	0	5	49
	4	5	2	4	2	8	5	3	4	6	5	4	2	50
Всего		31	19	20	17	11	30	20	22	17	11	9	8	215

Анализ дисперсии

Категории изменчивости	Число степеней свободы	Сумма квадратов	Средний квадрат	θ	P
Варианты	11	154,7292	14,0662	1,13	>0.05
Повторности	3	72,7292	24,2431	1,96	<0,05
Ошибка	33	410,5208	12,4400		
Всего	47	637,9792			

Таблица 35

Чабаны

A		48				72				96				Сумма
B		4		16		4		16				16		
C		10	20	10	20	10	20	10	20	10	20	10	20	
Повторности	1	4	2	3	7	4	1	3	0	5	1	1	0	31
	2	10	8	3	2	5	3	0	0	2	6	0	0	39
	3	10	13	5	7	4	6	11	7	4	4	9	5	85
	4	8	1	7	9	6	12	2	3	12	37	5	12	84

Всего	32	24	18	25	19	22	16	10	23	18	15	17	239

Анализ дисперсии

Категории изменчивости	Число степеней свободы	Сумма квадратов	Средний квадрат	θ	P
Варианты	11	89,2292	8,1117	0,77	>0.05
Повторности	3	206,8959	68,9653	6,52	ок. 0,001
Ошибка	33	348,8541	10,5713		
Всего	47	644,9792			

Таблица 36

Совхоз им. Тельмана

A		48				72				96				Сумма
B		4		16		4		16				16		
C		10	20	10	20	10	20	10	20	10	20	10	20	
Повторности	1	4	2	4	5	3	1	1	1	4	3	4	0	32
	2	3	1	4	1	4	3	10	3	1	5	0	9	44
	3	4	2	1	5	2	5	0	0	3	2	0	0	24
	4	0	4	3	1	2	2	0	0	5	1	3	0	21
Всего		11	9	12	12	11	11	11	4	13	11	7	9	121

Анализ дисперсии

Категории изменчивости	Число степеней свободы	Сумма квадратов	Средний квадрат	θ	P
Варианты	11	17,2292	1,5663	0,27	>0.05
Повторности	3	26,3959	8,7986	1,23	>0,05
Ошибка	33	190,3541	5,7683		
Всего	47	233,9792			

Таблица 37

Крыныца

A		48				72				96				Сумма
B		4		16		4		16				16		
C		10	20	10	20	10	20	10	20	10	20	10	20	
Повторности	1	5	7	4	4	4	1	4	0	4	1	4	1	39
	2	3	4	3	7	0	1	0	1	3	0	3	2	27
	3	14	6	8	2	1	0	4	6	4	3	4	2	54
	4	5	0	1	6	7	4	8	18	4	10	4	0	67
Всего		27	17	16	19	12	6	16	25	15	14	15	5	187

Анализ дисперсии

Категории изменчивости	Число степеней свободы	Сумма квадратов	Средний квадрат	θ	P
Варианты	11	113,2292	10,2936	0,84	>0.05
Повторности	3	76,0625	24,6875	2,00	>0,05
Ошибка	33	407,1875	12,339		
Всего	47	596,4792			

Разложение по степеням свободы (табл. 38), очевидно, не вызывает сомнений, так как по двум факторам (глубина и сетка) имеется только по одной степени свободы, а две степени свободы для дозировки, естественно, распределяются для решения двух вопросов: контраста между крайними дозировками и прямолинейности или непрямолинейности зависимости между дозировкой и количеством оставшихся личинок. Коэффициенты для взаимодействий, как всегда, получаются перемножением коэффициентов соответствующих прямых контрастов. Теты для 11 степеней свободы не рационально вычислять отдельно для каждой степени свободы, а целесообразно определить квадрат разности для данного контраста, умножая средний квадрат ошибки (в данном случае 1,9148) на теты для трех уровней значимости при 1 и 33 степенях свободы. Эти значения теты получаем путем интерполяции значений теты для 1 и 30 степеней свободы и для 1 и 60 степеней свободы: так как 60/33=1+9/11, то к значению теты для 60 степеней свободы надо прибавить 9/11 разницы значений теты для 30 и 60 степеней свободы.

Мы получаем для трех уровней значимости значения теты 4,14...7,47 и 13,05 и, умножая на 1,9148, получаем средний квадрат для изолированных степеней свободы, соответствующий уровню значимости 0,05...0,01 и 0,001 соответственно 7,93...14,38 и 25,00.

Таблица 38

	Разложение по степеням свободы для вариантов

A	48				72				96				Делитель Δ	Разность δ	$\frac{\delta^2}{\Delta}$
B	4		16		4		16		4		16				
C	10	20	10	20	10	20	10	20	10	20	10	20			
Сумма	12	15	9	10	5	13	1	2	9	8	7	8			
*A*2	1	1	1	1	0	0	0	0	-1	-1	-1	-1	32	+14	6,1250
*A*3	1	1	1	1	2	2	-2	-2	1	1	1	1	96	+36	13,5000
B	1	1	-1	-1	1	1	-1	-1	1	1	-1	-1	48	+25	13,0208
C	1	-1	1	-1	1	-1	1	-1	-1	-1	1	-1	48	-13	3,5208
BC	1	-1	-1	1	1	-1	-1	1	-1	-1	-1	1	48	-7	1,0208
*A*2*B*	1	1	-1	-1	0	0	0	0	-1	-1	1	1	32	+6	1,1250
*A*3*B*	1	1	-1	-1	-2	-2	2	2	1	1	-1	-1	96	-20	4,1667
*A*2*C*	1	-1	1	-1	0	0	0	0	1	1	-1	1	32	-4	0,5000
*A*3*C*	1	-1	1	-1	-2	2	-2	2	-1	-1	1	-1	96	+14	2,0417
*A*2*BC*	1	-1	-1	1	0	0	0	0	1	1	1	-1	32	-4	0,5000
*A*1*BC*	1	-1	-1	1	-2	2	2	-2	-1	-1	-1	1	96	+14	2,0417
													Всего	+61	47,5625

Сравнивая полученные квадраты разности для каждой степени-свободы с величинами, вычисленными для трех уровней значимости, видим, что только два контраста (именно *A*3 и *B*) дают различие некоторой значимости, а именно, лежащей между уровнями Р, равными 0,05 и 0,01. Из них один контраст (*B*) соответствует различию в сетке, т. е. частоте уколов: при 16 уколах на 1 м^2 число оставшихся личинок меньше, т. е. эффективность выше, чем при 4 уколах на 1 м^2. Другой контраст (*A*3) указывает на отсутствие прямолинейной связи между дозировкой и числом личинок. Но если мы посмотрим на цифры, то увидим, что нашим дозировкам (48,72 и 96 л/га) соответствуют численности личинок 46,21 и 32, такая зависимость, конечно, нелинейна, но вместе с тем признать ее как следствие нашей обработки трудно; если бы она подтвердилась на других пунктах или в суммарном анализе, то пришлось бы подумать, какое биологическое объяснение мы можем найти такому странному факту. Но ни на других пунктах, ни в суммарном анализе она не подтвердилась, и поэтому мы ,этот контраст не имеем оснований считать существенным: при наличии 11 сопоставлений по каждому из 5 пунктов мы имеем 55 независимых сопоставлений и, следовательно, вполне следует ожидать, что даже не имеющие реального значения контрасты приобретут величину, соответствующую уровню значимости 0,05 или даже 0,01 в единичных случаях.

Подобный же анализ для остальных четырех пунктов показал только в одном случае (в Носковцах) некоторую значимость для контраста *A*2 (т. е. для сопоставления крайних дозировок): именно квадрат разности оказался равным 55,1250 при квадрате для уровня 0,05, равном 51,626; все остальные 43 сопоставления не показали ни малейшего намека на значимость.

Таким образом, изолированное рассмотрение данных по пунктам приводит к таким ничтожным различиям, что получается впечатление, что никакого различия между собой варианты не дают, т. е. что эффективность одна и та же при всех испытанных дозировках, глубинах и сетках. Но очень часто ряд ненадежных .указаний в совокупности могут дать вполне надежное следствие. Прежде чем приступать к такому объединению, сопоставим данные по разложению по 11 степеням свободы вариантов всех пяти пунктов. Мы ограничиваемся только разностями соответствующего контраста (δ), так как знак этих разностей уже позволит судить, в какой степени перспективно является такое объединение и эти цифры могут быть в дальнейшем использованы (табл. 39).

Таблица 39

Контрасты	δ					Сумма
	Черкассы	Носковцы	Чабаны	Совхоз им. Тельмана	Крыныца	
A2	14	42	26	4	30	116
A3	36	—34	38	10	10	60
B	25	23	37	11	—5	91
C	—13	1	7	9	15	19
BC	—7	—3	13	—1	19	21
A2B	6	2	4	—12	0	0
A3B	—20	26	—8	—10	64	52
A2C	—4	8	—2	2	—4	0
A3C	14	64	—2	—12	24	88
A2BC	—4	4	8	—2	22	28
A3BC	14	48	40	20	—26	96
Сумма	61	181	161	19	149	571

Все суммирования мы имеем право делать потому, что весь опыт в целом является вполне уравновешенным. Знак плюс во всех случаях опущен. Мы видим, что первая степень свободы *(A2* — противопоставление, крайних дозировок) во всех пяти случаях положительна. Так как положительный знак ожидался до опыта (естественно, надо предполагать, что при более слабых дозировках останется больше личинок, чем при сильных), то уже такое чисто качественное совпадение знаков указывает на вероятность отсутствия разности, равную $1:2^5$, или 1/32, т. е. имеющую некоторую значимость. Поэтому этот один контраст интересно проверить более тщательно. Для всех остальных степеней свободы мы не имеем постоянства знаков для разности: оценить там значимость различий чисто наглядным путем трудно, поэтому целесообразно проверить, в особенности для В, а также выяснить наличие взаимодействия наших одиннадцати контрастов с особенностями наших пяти пунктов.

Составим план обобщения наших данных: этот план должен дать не только все требуемые величины, но и дать при вычислениях возможность проверки каждой цифры, кроме того, вычисление должно вестись возможно экономно в смысле затраты труд. Это достигается использованием всех предыдущих вычислений, и новые вычисления здесь сводятся к минимуму.

Объединяя весь материал, мы получаем 48·5, или 240 дат, что дает всего 239 степеней свободы, которые распределяются по следующим категориям изменчивости:

1) варианты — 11 степеней свободы, так как, очевидно, варианты повторяются в каждом пункте;
2) пункт (всего 5) —4 степени свободы;
3) взаимодействие вариантов с пунктами (действие вариантов в условиях разных пунктов может быть различным (4-11 степеней свободы, или 44 степени свободы);
4) повторности: в каждом пункте 3 степени свободы, так как повторности каждого пункта совершенно не связаны с повторностями других, то, очевидно, будем иметь 3·5, или 15 степеней свободы;
5) наконец, ошибка составляется из 33 степеней свободы для каждого пункта, также взаимно независимых, получаем: 33·5, или 165 степеней свободы, можем просуммировать:

варианты	11
пункты	4
взаимодействие вариантов с пунктами	44
повторности	15
ошибка	165
Всего:	239

Мы последовательно вычислим общую сумму квадратов от общего среднего, затем вычислим суммы квадратов для вариантов пунктов, взаимодействия и прочего, а затем уже перейдем к детальному рассмотрению контрастов в пределах вариантов. Используем наши прежние вычисления по каждому пункту отдельно для составления табл. 40, в которую войдут общие суммы дат, общая сумма квадратов для всех дат и общая сумма квадратов от среднего для каждого пункта — величины нам уже известные. Ясно, что нет надобности вновь суммировать 240 наших дат и квадраты этих 240 дат, когда все эти вычисления по частям уже были сделаны.

Таблица 40

Пункт	Сумма дат	Сумма квадратов дат от 0	Общая сумма квадратов от *M*	Сумма квадратов для ошибки от *M*
Черкассы	99	321	116,8125	63,1875
Носковцы	215	1601	637,9792	410,5208
Чабаны	239	1835	644,9792	347,8541
Совхоз им. Тельмана	121	539	233,9792	190,3541
Крыньща	187	1325	596,4792	407,1875
Сумма	861	5621	2230,2293	1420,1040

Очевидно, общая сумма квадратов от общего среднего получится, если мы из общей суммы квадратов единичных дат от нуля вычтем поправку, равную квадрату общей суммы, деленному на число дат, т. е. $\frac{861^2}{240}$, или 3088,8375, получаем 5621—3088,8375, или 2532,1625.

Эта величина сразу же может быть проверена. В самом деле, общая изменчивость всех дат вокруг общей средней слагается из двух величин: 1) изменчивости дат каждого пункта вокруг пунктовой средней 47·5=235 степеней свободы и 2) изменчивости средних величин по каждому пункту вокруг общей средней 4 степени свободы. Последняя величина вычисляется легко, она, очевидно, равна

$$\frac{99^2+215^2+239^2+121^2+187^2}{48}-3088{,}8375,$$

или 3390,7708—3088,8375 = 301,9333.
2230,2293+301,9333= (с погрешностью в последнем знаке) 2532,1626.

Перейдем теперь к вычислению сумм квадратов для вариантов и для взаимодействия между вариантами и пунктами. Имеется всего 12 вариантов в пяти пунктах. Принимая, следовательно, особенности пунктов как новый фактор, наблюдаемый в пяти модальностях, мы имеем 60 дат, представляющих собой все комбинации ранее исследованных трех факторов (дозировка, глубина и сетка) и нового — пунктовые отличия. Это соответствует 59 степеням свободы (60—1), которые, очевидно, распределяются следующим образом:

варианты по дозировке, глубине и сетке	11 степеней свободы
различия по пунктам	4 » »
взаимодействие тех и других	44 » »
Всего:	59 » »

Для того чтобы получить сумму квадратов, соответствующую всем 59 степеням свободы, мы можем использовать прежние вычисления. Приведем старые данные по пунктам (табл. 41).

Таблица 41

Пункты	Суммы квадратов для вариантов		Сумма квадратов для повторностей
	от 0	от *M*	
Черкассы	251,75	47,5625	6,0625
Носковцы	1117,75	154,7292	72,7292

Чабаны	1279,25	89,2292	206,8959
Совхоз им. Тельмана	322,25	17,2292	26,3959
Крыньща	841,75	113,2292	76,0625
Сумма	3812,75	421,9793	388,1460

Очевидно, для того, чтобы получить общую сумму квадратов отклонения для всех 59 степеней свободы, надо сначала получить сумму квадратов всех 60 дат, суммируя пять сумм по пунктам (каждая для 12 дат), получим 3812,75. Вычтя из этой суммы поправку на общую среднюю 3088,8375, получим 723,9125. Мы сейчас уже можем произвести проверку на этом этапе анализа. В самом деле, мы имеем:

1) сумма квадратов для 59 степеней свободы для вариантов, пунктов и их взаимодействия.......723,9125

2) сумма квадратов для 15 степеней свободы по повторностям (получаемая прямым суммированием 5 сумм по 3 степени свободы для каждого пункта, что уже сделано в таблице)....388,1460

3) сумма квадратов для 165 степеней свободы ошибки (получается также суммированием сумм по каждому пункту, что приведено раньше)...1420,1040

Всего 239 степеней свободы....2532,1625

Это в точности совпадает с вычисленным ранее.

Другая проверка: сумма 723,9125 для 59 степеней свободы пунктов, вариантов и взаимодействий, очевидно, складывается из суммы квадратов для пяти пунктов 301,9333 и суммы квадратов для 55 степеней свободы по вариантам и по взаимодействию пунктов и вариантов — 421,9793; эта последняя сумма в точности равна сумме пяти пунктовых сумм вариантов от пунктовых средних, так как эта последняя сумма, соответствующая, конечно, тоже 55 степеням свободы, отражает как изменчивость всех вариантов вокруг общего среднего, так и изменчивость их вокруг пунктовых средних.

Определим теперь сумму квадратов для вариантов (11 степеней свободы). Для этого получим суммы по каждому варианту по всем пяти пунктам, пользуясь итоговыми цифрами первых пяти таблиц: для 1-го варианта сумма будет 12+31+32+11+27=113, для второго — 84, для третьего — 75 и т. д. Возведя все эти цифры в квадрат, просуммировав и разделив эту сумму квадратов на 20 (каждая сумма основана на 20 датах) и вычтя общую поправку 3088,8375, получим 175,3125 — сумму квадратов для 11 вариантов от общей средней. Остаток от 421,9793 и даст 246,6667 — сумму квадратов для 44 степеней свободы взаимодействий вариантов и пунктов. Обе суммы 175,3125 и 246,6667 проверяются вычислением по каждой степени свободы или по группам степеней свободы, что будет показано дальше.

Сейчас уже мы можем дать таблицу анализа дисперсии для основных категорий изменчивости (табл. 42).

Мы видим, что средний квадрат для взаимодействий вариантов и пунктов даже меньше среднего квадрата ошибки, поэтому для оценки значимости контрастов в пределах вариантов, пунктов и повторностей мы можем объединить эти две категории изменчивости. Получим сумму квадратов для 209 степеней свободы, равную 1666,7707, в среднем 7,9750. Теты в предпоследней графе определялись путем деления на этот средний квадрат. Мы видим, что между вариантами уже намечаются значимые различия, но особенно резкие различия оказываются между пунктами и между повторностями в пределах каждого пункта. Так как последние два различия лишь указывают на различие в плотностях личинок в пределах повторностей пункта и между пунктами, то мы их сейчас подробнее исследовать не будем, подробное же исследование проведем в отношении вариантов и взаимодействий. Сначала определим размеры квадратов разностей, соответствующих определенным уровням значимости. Для вариантов средним квадратом ошибки будет 7,9750, основанный на 209 степенях свободы, для взаимодействий—8,6067, основанный на 165 степенях свободы. Заметим, что степень значимости различий между вариантами можно определять, сравнивая со средним квадратом взаимодействий, и этот прием даже предпочтительнее в тех случаях, когда средний квадрат взаимодействий больше такового для ошибки. Этим путем мы получаем более осторожное суждение о значимости различий, так как в основу суждений об изменчивости кладется изменчивость между пунктами, а не в пределах пунктов. Но в данном случае средний квадрат взаимодействий даже меньше такового для ошибки, и поэтому объединение их допустимо.

Таблица 42

Категории изменчивости	Число степеней свободы	Сумма квадратов	Средний квадрат	θ	*P*
I варианты	11	175,3125	15,9375	1,998	<0,05
II пункты	4	301,9333	75,4833	9,465	<0,001
III взаимодействие вариантов и пунктов	44	246,6667	5,6061		
IV повторности	15	388,1460	25,8764	3,245	<0,001
V ошибка	165	1420,1040	8,6067		
Всего	239	2532,1625			

Для определения размеров квадрата разности для изолированных степеней свободы, имеющего определенную значимость, сначала путем интерполяции получим тету, а затем умножим полученные теты на средний квадрат ошибки.

Мы имеем теты (при одной степени свободы большей вариан-сы) для 60 и бесконечного числа степеней свободы меньшей вариансы (табл. 43).

Таблица 43

P	1/60	1/∞	Разность	1/209	1/165
0,05	4,00	3,84	0,16	3,89	3,90
0,01	7,08	6,64	0,44	6,77	6,80
0,001	11,97	10,83	1,14	11,16	11,24

Интерполяция производится таким образом: к тете, соответствующей бесконечному числу степеней свободы меньшей варианса, прибавляется 60/209 или 60/165 разницы между тетами для 60 и бесконечного числа степеней свободы. Получаем квадраты, соответствующие определенным уровням значимости в двух последних графах табл. 43.

Для вычисления квадратов различий, соответствующих отдельным степеням свободы для вариантов, мы будем пользоваться теми же сопоставлениями, что и в табл. 38. Но новых вычислений нам придется делать немного, так как суммарные различия по пяти пунктам уже приведены в табл. 39 (это и есть суммарные δ). Их надо возвести в квадрат и разделить на Δ, которые во всех случаях (как основанный на впятеро большем количестве дат) будут в пять раз больше делителей для отдельных пунктов (табл. 44).

Таблица 44

№	Контраст	Разность δ	Делитель Δ	$\frac{\delta^2}{\Delta}$	P
1	*A*2	116	160	84,1000	близко 0,001
2	*A*3	60	480	7,5000	не существенно
3	*B*	91	240	34,5042	<0,05
4	*C*	19	240	1,5042	
5	*BC*	21	240	1,8375	
6	*A2B*	0	160	0,0000	
7	*A3B*	52	480	5,6333	не существенны
8	*A2C*	0	160	0,0000	
9	*A3C*	88	480	16,1333	
10	*A2BC*	28	160	4,9000	
11	*A3BC*	96	480	19,2000	
	Сумма	571		175,3125	

Сумма совпадает с вычисленной ранее.

Размеры значимых квадратов получаем умножением 7,975 на теты из табл. 43 для 1/209 степеней свободы получаем квадраты для уровней значимости:

для P, равной 0,05, —31,02
» » 0,01 —53,99
» » 0,001 —89,00

Сравнивая с этими значениями величины, полученные в табл. 44, видим, что первый контраст — сопоставление крайних дозировок дает очень существенный квадрат (близкий к уровню значимости 0,001); некоторую значимость имеет и квадрат, соответствующий сетке; все остальные сопоставления никакого намека на значимость не обнаруживают. Объединение материала позволило нам таким образом получить один очень надежный вывод (о влиянии дозировки) и один малонадежный (о сетке).

Теперь перейдем к изучению взаимодействий между вариантами и пунктами. Конечно, на наличие серьезных взаимодействий у нас нет указаний, поскольку, как уже было отмечено, средний квадрат для взаимодействий меньше среднего квадрата для ошибки, но такой анализ помимо методического значения имеет две стороны: 1) не исключена возможность, что среди 44 степеней свободы, соответствующих незначащим контрастам, могли затеряться одна или две, имеющие значение; 2) такое вычисление дает хорошую проверку прежним. Можно такое вычисление проделать обычными приемами, т. е., установив ортогональную серию коэффициентов для пунктов (четыре степени свободы), перемножить их последовательно на коэффициенты всех 11 степеней свободы табл. 38. Но это путь слишком громоздкий и ненужный, так как по величине нашего среднего квадрата для взаимодействий мы можем с уверенностью сказать, что огромное большинство степеней свободы не дадут значащего контраста. Проще поступать таким образом (берем данные табл. 39). Колебания пунктовых значений дельты для каждой степени свободы около среднего значения для данной степени свободы укажут размеры взаимодействия для четырех степеней свободы. Если взаимодействия нет, то эти колебания будут по величине близкими к случайным колебаниям, если же они значительно превосходят размеры среднего квадрата ошибки, то, значит, налицо выраженное взаимодействие. Таким образом, сумму квадратов для взаимодействия контраста *A2* и пунктов (4 степени свободы) получаем равную:

$$\frac{14^2+42^2+26^2+4^2+30^2}{32}-\frac{116^2}{160}=111{,}0000-84{,}1000=26{,}9000,$$

то же для взаимодействия контраста *A*3 с пунктами:

$$\frac{36^2+34^2+38^2+10^2+10^2}{96}-\frac{60^2}{480}=42{,}6667-7{,}5000=35{,}1667.$$

Таким же образом для всех 11 групп по четыре степени свободы получим следующие значения (табл. 45).

Таблица 45

Взаимодействие пунктов с контрастами	Сумма квадратов
1 *A*2	26,9000
2 *A*3	35,1667
3 *B*	21,1000
4 *C*	9,4333
5 *BC*	10,4333
6 *A2B*	6,2500

7 *А3В*	49,9500
8 *А2С*	3,2500
9 *А3С*	36,1167
10 *А2ВС*	13,3500
11 *А3ВС*	34,7167
Всего:	246,6667

Общая сумма совпадает, как и должно, с вычисленной ранее. Ни на одну группу не выпало такой суммы квадратов, чтобы появилась определенная значимость. Даже 7-я группа с наибольшим значением суммы квадратов дает средний квадрат $\frac{49{,}9500}{4}$, или 12б4875. Тета $\frac{(12{,}4865)}{(8{,}6067)}$ равна 1,45, тогда как минимальное значение теты для 4 и 165 степеней свободы равно 2,43. Использовав значение теты табл. 43 для 1 и 165 степеней свободы и умножив их на 8,6067, получаем значения квадратов-разности для изолированных степеней свободы, равные 33,55„ 58,53 и 96,74, для обычных трех уровней значимости. Отсюда мы видим, что в большинстве групп по четыре степени свободы заведомо значащие контрасты отсутствуют, так как даже если бы вся изменчивость какой-либо группы была сосредоточена на одной степени свободы, то и тогда только группы 2, 7, 9 и 11 дали бы значащие контрасты (не превосходящие низшего уровня значимости), так как суммы квадратов только для этих групп превосходят 33,57. На примере 7-й группы, дающей наибольшую сумму квадратов, покажем разложение ее на четыре степени свободы, которая производится обычным способом (табл. 46).

Таблица 46.

Разложение группы взаимодействии по четырем степеням, свободы

Степень свободы	Крыныца	Носковцы	Чабаны	Совхоз им. Тельмана	Черкассы	Делитель Δ	Разность δ	$\frac{\delta 2}{\Delta}$
	64	26	—8	—10	—20			
I	3	3	—2	—2	—2	2880	346	41,5681
II	1	--1	0	0	0	192	38	7,5208
III	0	0	1	0	—1	192	12	0,7500
IV	0	0	1	—2	1	576	—8	0,1111
							Сумма	49,9500

Делитель получался умножением сумм квадратов всех коэффициентов на 96, поскольку разность этого контраста всегда основана (см. табл. 39) на 96 датах. Разложение здесь велось чисто эмпирическое: суммы расположены в убывающем порядке и для первой степени свободы объединены две положительные против трех отрицательных. Этим путем удалось большую часть изменчивости сосредоточить на одной степени свободы. Получился как бы один контраст, имеющий некоторую значимость, но надо иметь в виду, что это контраст чисто эмпирический и ему не соответствует никакая заранее поставленная гипотеза, так как нельзя найти какого-либо -признака (по почве или чему-либо другому), общего Носковцам и Крыныце, который можно было бы противопоставить аналогичному признаку у трех других пунктов. Среди же 44 степеней свободы один контраст, удовлетворяющий низкому уровню значимости, может возникнуть совершенно случайно. Мы приходим, таким образом, к выводу, что никаких намеков на взаимодействие между вариантами и пунктами нет. Наша обработка позволила только выделить два контраста.

Первый — зависимость от дозировки; наличие вполне значащего различия между крайними дозировками и отсутствие отклонений от прямолинейности для трех дозировок, что позволяет считать эффективность пропорциональной дозировке. Объединяя все данные и беря также число личинок на контроле, получим такие цифры:

дозировка 0 (контр.)	48	72	96
число личинок 750	355	267	239
общая средняя эффективность (%)	53	64	68

Второй вывод — связь с сеткой уколов (много менее надежный) — позволяет принять для 4 уколов на 1 $м^2$ эффективность 58%, а для 16 уколов 66%, различие, .как видим, очень небольшое. Отсутствие влияния глубины позволяет отказаться от глубоких уколов и ограничиться только внесением на 10-сантиметровую глубину.

Произведенное исследование касалось лишь различий между .вариантами затравки почвы (дозировка, глубина и сетка), но не касалось вопроса о сравнительной эффективности затравки вообще в условиях разных пунктов, которые, как было уже указало, были расположены на разных почвах. Можно, конечно, подобным же способом произвести исследование и в отношении этого .контраста, введя контрольные площадки в общую обработку и определяя значение контраста *А*1 (противопоставление контроля всем остальным вариантам): этот контраст имеет высокую значимость *(Р* меньше 0,001) во всех пунктах, и при определении взаимодействия между этим контрастом и пунктами выясняется очень существенное отличие при противопоставлении Чабанов и Крыныцы, с одной стороны, к трем остальным пунктам — с другой. Но так как зараженность личинками пунктов очень различна (в частности, зараженность контроля Чабанов и Крыныцы в 1.5—2 раза выше зараженности трех остальных пунктов), то такой результат не дает нам возможности решить: определяется ли •существенность различия разницей эффективности в разных пунктах или просто разницей исходной зараженности (о которой мы .можем судить по зараженности контроля). Это можно достичь разными способами: более совершенным способом является анализ ковариансы, проще и распространеннее замена абсолютных значений числа личинок процентом эффективности, определяемым по соответствующему контролю.

Можно было бы, конечно, вычислить эффективность (разность .личинок по сравнению с контролем той же повторности, отнесенная к числу личинок на контроле) для всех вариантов опыта, но ввиду того, что кроме дозировки другие варианты показали ничтожное или отсутствие различия и отсутствие намеков на взаимодействие с пунктами, то такая обработка вряд ли внесла бы существенные изменения в выводы. Поэтому ограничимся определением эффективности, связанной с различными дозировками. Для этой цели, пользуясь данными табл. 33—37, объединяем все четыре варианта по глубине и сетке одной дозировки в одну сумму . ад эти суммы выписываем в табл. 47 и там же приводим данные по контролю, то же — для четырех площадок по 4 м^2 для каждой повторности каждого пункта; таким образом, каждая цифра левой части табл. 47 показывают количество личинок на 16 м^2. Из числа личинок на контроле вычитаем число личинок варианта по дозировке той же повторности и разность делим на число личинок в контроле, получаем цифру эффективности с точностью до одного процента. Эти данные приведены в правой части табл. 47.

Таблица 47

Количество личинок по вариантам дозировок, пунктам и повторностям, и эффективность в процентах

Пункт	Пов-торн.	Количество личинок по вариантам дозировок					Эффективность при дозировках			
		0	48	72	96	сумма	48	72	96	сумма
Черкассы	1	26	15	6	10	57	42	77	62	181
	2	52	12	7	6	77 33 50	77	87	89	253
	3	14	9	5	5 11		36	64	64	164
	4	26	10	3			62	89	58	208
	Всего	118	46	21	32	217	217	317	273	807
Носковцы	1	37	26	39	13	115	30	-5	65	90
	2	25	20	13	5	63	20	48	80	148
	3	17	28	11	10	66	—65	35	41	11
	4	62	13	20	17	112	79	68	73	220
	Всего	141	87	83	45	356	64	146	259	469
Чабаны	1	23	16	8	7	54	30	65	70	165
	2	58	23	8	8	97	60	86	86	232
	3	58	35	28	22	143	40	52	62	154
	4	67	25	23	36	151	63	66	46	175
	Всего	206	99	67	73	445	193	269	264	726
Совхоз им. Тельмана	1	16	15	6	11	48	6	63	31	100
	2	35	9	20	15	79	74	43	57	174
	3	21	12	7	5	45	43	67	76	186
	4	18	8	4	9	39	56	78	50	184
	Всего	90	44	37	40	211	179	251	214	644
Крыньща	1	42	20	9	10	81	52	79	76	207
	2	25	17	2	8	52	32	92	68	192
	3	52	30	11	13	106	42	79	75	196
	4	76	12	37	18	143	84	51	76	211
	Всего	195	79	59	48	382	210	301	295	806
Сумма по всем пунктам							863	1284	1305	3452

Мы видим, что полученные цифры очень колеблются: даже суммарные эффективности по пунктам не дают, что можно было бы ожидать, плавного возрастания эффективности при повышении дозировки. В четырех пунктах из пяти (все, кроме Носковцев) максимальная эффективность оказывается не при максимальной дозировке, а при средней. Единственный же пункт, Носковцы, где суммарные эффективности монотонно возрастают с дозировкой,. обнаруживает в двух случаях отрицательные эффективности (число личинок в двух площадках по 16 м^2 превышает число на соответствующем контроле), которые, конечно, указывают или на крайне неравномерное распределение личинок, или на какие-либо дефекты. в проведении опыта или регистрации его.

В литературе широко распространен прием в таких явно абсурдных случаях (так как обработка не может увеличить число личинок, и если бы такое явление повторялось часто, то пришлось бы заключить, что яд вместо убивающего действия привлекает личинок с других мест, для такого предположения в данном случае никаких оснований нет) отбрасывать эти цифры или вместо отрицательного значения эффективности ставить ноль. Такой прием основан на непонимании самого характера исследования.

Сравнивая наши данные, мы должны определить систематическое различие, вносимое нашим опытом (в данном случае отравливанием почвы), с различием случайным, основанным на влиянии всех сопутствующих факторов: величина этой разницы и покажет степень надежности наших выводов.

Эти сопутствующие факторы влияют на все решительно даты, но там, где результат не явно абсурден, они не бросаются в глаза. Выбраковывая некоторые цифры только по их размерам, мы искусственно изменяем среднее значение и уменьшаем размеры случайной ошибки, являющейся критерием надежности наших выводов. Выбраковка тех или иных дат без специального исследования допустима только в тех случаях, где нами твердо установлено, что на бракуемой нами делянке имели место такие события (например, потрава скотом, потеря части материала, перерыв в работе из-за непогоды и т. д.), которые на всех остальных делянках места не имели. В данном случае браковка «отрицательных эффективностей» имела бы еще один результат, также не могущий быть принятым: суммарная эффективность по всем пунктам оказалась бы максимальной для средней, а не для максимальной дозировки: 67,8% Для 72 кг/га и 65,0% для 96 кг/га, Поэтому будем вести обработку наших 60 дат по эффективности (три варианта и четыре повторности на пять пунктов), не выбраковывая дат, но имея в виду особенности пункта Носковцы.

Общий анализ варианcы должен дать, очевидно, такое разложение:

	Число степеней свободы
варианты	2
пункты	4
взаимодействие вариантов с пунктами	8
повторности	15
ошибка	30
Всего	59

Общая сумма квадратов вычисляется, как всегда, возведением в квадрат всех 60 дат и вычитанием поправки $\frac{3452^2}{60}$,

получаем _239428,00
198605,07
общая сумма — 40822.93

Сумму квадратов для вариантов получаем:

$$\frac{863^2+1284^2+1305^2}{20}=204822{,}50$$

минус поправка 198605,07
сумма квадратов от *M*—6217,43

Для пунктов получим аналогичным образом:

$$\frac{807^2+469^2+726^2+644^2+806^2}{12}=205221{,}50$$

минус поправка 198605,07
сумма квадратов от *M*—6616,43

Для того чтобы получить сумму квадратов для взаимодействий вариантов с пунктами (8 степеней свободы), вычисляем, подобно тому, как было сделано ранее, сумму квадратов для 15 комбинаций вариантов и пунктов (т. е. 217,-317, 273, 64 и т. д.), разделив эту сумму квадратов на четыре и вычтя поправку, получим значение суммы квадратов для 14 степеней свободы, включающее 2 степени для вариантов, 4 для пунктов и 8 для взаимодействий.

Получаем
15512,43—(6217,43+6616,43)—12833,86
разность-сумма квадратов для взаимодействия—2678,57

Все эти три полученные суммы проверяются, как увидим дальше, путем вычисления по частям.

Нам остается вычислить сумму квадратов по повторностям. Для этого возводим в квадраты значения сумм для всех повторностей, т. е. 181, 253, 164 и т. д. Вычисление суммы квадратов по повторностям можно вести от общего среднего. Тогда из суммы квадратов (181, 253 и т. д.), деленной на три, вычитается не .только общая поправка 198605,07, но и сумма квадратов для пунктов 6616,43, так как двадцать сумм по повторностям на всех пяти пунктах соответствуют не только 15 степеням свободы по повторностям, но и 4 суммам по пунктам. Мы получаем 12325,17. Или же можно вычислить сумму квадратов повторностей для каждого пункта отдельно, вычитая из суммы квадратов повторностей одного пункта свою пунктовую поправку. Например, для Черкасс имеем

$$\frac{181^2+253^2+164^2+209^2}{3}-\frac{807^2}{12}=1511{,}58$$

Суммируя все данные по пунктам, получаем ту же величину — 12325,16. Последний способ предпочтительнее, так как он уже включает в себя разложение общей суммы квадратов на группы по пунктам.

Теперь мы получаем общий анализ варианcы (табл. 48).

И здесь, как и при предыдущем общем анализе, средний квадрат взаимодействия вариантов и пунктов оказался меньше среднего квадрата ошибки, и потому эти две категории объединены и тета вычислена не в отношении к 432,84, а к 412,21.

Таблица 48

Анализ варианcы по эффективности

Категории изменчивости	Число степеней свободы	Сумма квадратов	Средний квадрат	θ	P
Варианты	2	6217,43	3108,72	7,54	<0,01
Пункты	4	6616,43	1654,11	3,82	ок. 0,01
Взаимодействие вариантов и пунктов	8	2678,57	334,82		
Повторности	15	12325,16	821,68	1,90	<0,05
Ошибка	30	12985,34	432,84		
Всего	59	40822,93	412,21		
Взаимодействие + ошибка	38	15663,91			

Мы видим, что разница вариантов показывает хорошую значимость и почти такую же показывают и пункты; повторности показывают значимость на низшем уровне.

Разложение сумм квадратов для вариантов по двум степеням свободы можно произвести, как уже в свое время указывалось, разными способами. Обычный прием — противоположение крайних дозировок и проверка степени отклонения их от прямой линии дает разложение:

$$A1=\frac{(1305-863^2)}{40}=4884{,}10$$

$$A2 = \frac{(863 + 1305 - 2 \cdot 1284)^2}{120} = 1333{,}33$$

Сумма—6217,43 (совпадает с ранее вычисленными)

Определяя размеры квадратов разности для трех уровней значимости, беря средний квадрат ошибки, равный 412,21, принимая во внимание, что 6038 равны 1+0,5789, и, следовательно, прибавляя к значению теты для 60 степеней свободы 0,5789 интервала между тетами для 60 и 30 степеней свободы, получим:

		θ	δ^2/Δ
для *P*, равной	0,05	4,10	1690
» »	0,01	7,35	3029
» »	0,001	12,73	5247

Мы видим, что значимость для первой степени свободы приближается к 0,001, а для второй не достигает, хотя и немного, даже минимального уровня значимости. На близость значений для дозировок в 72 и 96 л заставляет нас догадаться, что более контрастное распределение суммы квадратов получатся, если мы противопоставим низшую дозировку двум более высоким. Получим тогда такое разложение:

$$A1 = \frac{(1284 + 1305 - 2 \cdot 863)^2}{120} = 6206{,}11$$

$$A2 = \frac{(1305 - 1284)^2}{40} = 11{,}03$$

Всего: 6217,44

Здесь первая степень свободы дает значимость очень высокую (*P* меньше 0,001), а вторая практически равна нулю. Оба этих разложения соответствуют двум гипотезам: вторая степень свободы в первом разложении проверяет степень отклонения линии, соединяющей эффективности трех вариантов от прямой. Мы видим, что нельзя надежно говорить, что эта линия отклоняется от прямой и что, следовательно, нет пропорциональности между приростом эффективности и приростом дозировки. Вторая же степень свободы во втором разложении, наоборот, выявляет наличие различия между двумя высшими дозировками, и мы видим, что здесь нет и намека на такое различие, но отсутствие различия соответствует криволинейному характеру зависимости между дозировкой и эффективностью (при наличии высокой значимости первой степени свободы). Значит, наш опыт недостаточно точен, чтобы сделать надежный выбор между гипотезами прямолинейной и криволинейной зависимости, но более вероятна криволинейная зависимость, так как 11,03 много меньше 1333,33. Иначе говоря, в общем можно сказать, что переход от средней дозировки (72 литра) к высшей (96) не дает намеков на повышение эффективности.

Разложим теперь сумму квадратов для пунктов. Это уже позволит нам судить о том, существуют ли различия эффективности в разных пунктах, как известно, отличающихся по почве. На основании всего предыдущего опыта можно ожидать (поскольку парадихлорбензол всегда наиболее эффективным был на самых легких почвах с низкой абсорбцией) наименьшую эффективность в Носковцах — с тяжелыми почвами. И действительно, мы видим, что суммарная эффективность в Носковцах сильно уступает остальным. Первый контраст является и эмпирически наибольшим, и теоретически вполне обоснованным.

Дальше чисто теоретически надо противопоставить Черкассы и совхоз им. Тельмана с их легкими почвами Чабанам и Крыньще со средними почвами, но по цифрам можно видеть, что это не даст значимого контраста. Поэтому наряду с таким теоретически .обоснованным контрастом введем чисто эмпирическое противопоставление Черкасс и Крыныц двум остальным: это противопоставление, как нетрудно видеть, будет ортогонально с предыдущим.

Последний контраст будет дополнять серию и служить только для контроля общей суммы квадратов. Получаем разложение (табл.49).

Таблица 49

Разложение по степеням свободы для пунктов

Степень свободы	Носковцы	Черкассы	Чабаны	Совхоз им. Тельмана	Крыньща	Δ	δ	$\frac{\delta^2}{\Delta}$
	469	807	726	644	806			
I	—4	1	1	1	1	240	1107	5106,04
II	0	—1	1	—1	1	48	81	136,69
III	0	1	—1	—1	1	48	243	1230,19
IV	0	—1	—1	1	1	48	—83	143,52
							Всего:	6616,44

Мы видим, что кроме очень значимого контраста между Носковцами и остальными четырьмя пунктами мы не находим других существенных различий. Эффективность значительно хуже на тяжелых почвах, но различие для средних и легких почв уловить не можем. К сожалению, уже указанные ранее дефекты материала Носковцев не позволяют нам дать этому выводу ту убедительность, которая вытекает из чисто статистической оценки надежности различий, и это можно подтвердить еще и тем, что; анализируя данные по изменчивости повторностей, мы ясно приходим к выводу о недоброкачественности носковского материала. Мы уже видели, что средний квадрат по повторностям указывает на наличие довольно существенных различий. Теоретически это не имеет никакого обоснования в нашем материале, так как это могло бы быть только в том случае, если бы в каждом пункте повторности были заложены на различной почве. Но если мы выпишем суммы квадратов повторностей для пункта по каждому пункту отдельно,. то получим табл. 49а.

Таблица 49а

Пункт	Средняя	Средний квадрат	θ

Носковцы	7844,92	2614,97	6,3
Черкассы	1511,58		
Чабаны	1207,00		
Совхоз им. Тельмана	1681,33	4480,24….373,35	меньше 1
Крыница	80,33		
Сумма	12325,16		

Таким образом, большая часть изменчивости по повторностям, падает на Носковцы и значимость различий там близка к 0,001. В остальных же пунктах повторности не дают намека на значимые различия по эффективности (что и следует ожидать, так как сумма для трех степеней свободы не превышает 1681, а величина квадрата для низшего уровня значимости равна 1690. Поэтому разложение по степеням свободы для всех пунктов не имеет смысла. Для Носковцев мы можем проделать такое разложение.. которое будет носить, очевидно, чисто эмпирический характер, так как нам ничего не известно о реальных различиях между четырьмя повторностями в Носковцах. Возьмем для примера такое разложение, где минимальное значение (11) противопоставляется всем. остальным, а затем следующее по величине — оставшимся. Получим следующее разложение (табл. 49б).

Таблица 496

Повторность	90	148	11	220	Δ	δ	$\frac{\delta^2}{\Delta}$	P
I	1	1	-3	1	36	425	5017,36	ок. 0,001
II	—2	1	0	1	18	188	1963,56	<0,05
III	0	—1	0	1	6	72	864,00	

Всего: 7844,92

Один контраст — очень существенный, другой — на грани значимости. На основании этого одну повторнрсть следует вообще взять под сомнение (мы видим по исходным цифрам табл. 47, что там помимо невозможной отрицательной эффективности и другие цифры очень низки). Но вообще ни одна повторность Носковцев не дает цифр вполне удовлетворительных. В двух повторностях встречаются отрицательные эффективности, в четвертой — наивысшая эффективность при наименьшей дозировке (и, кроме того, в этой повторности эффективность не уступает эффективности остальных пунктов), во второй, напротив, слишком резкое различие между дозировками. Все это иллюстрирует тот вывод, что в Носковцах мы имеем чрезвычайно резкие (и совершенно непонятные теоретически) колебания эффективности по повторностям, что заставляет нас не придавать данным по Носковцам значения, одинакового со значениями других пунктов.

Нам остается теперь проделать разложение взаимодействия дозировок и пунктов. Здесь общая сумма квадратов такова, что на очень существенные контрасты мы рассчитывать не можем, но один контраст умеренной значимости попасть может, так как сумма квадратов восьми степеней свободы равна 2678,57. Очевидно, ни одного контраста со значимостью, высшей 0,01, мы получить не можем, так как при среднем квадрате ошибки, равном 432,84, мы получаем для 1 и 30 степеней свободы следующие квадраты для трех уровней значимости:

для Р, равной		0,05, — 1805
»	»	0,01 —3272
»	»	0,001—5752

Очевидно, поэтому нет надобности производить вычисления для каждой из восьми степеней свободы. Разложим всю сумму на две группы по четыре степени свободы и будем искать значущие контрасты в той из двух групп, где есть перспектива найти. Разложение проведем применительно к обоим способам разложения по вариантам.

Для этого вычислим значения обоих степеней для вариантов для каждого пункта отдельно. Характер вычисления дельты для каждой степени свободы покажем соответствующим символом — 1, 0, 1 и т. д. Тогда для первого способа получим следующее разложение (табл. 49в).

Таблица 49в

Степень свободы	Характер вычисления	Черкассы	Носковцы	Чабаны	Совхоз им. Тельмана	Крыница	Сумма
*А*1	1 0 —1	56	195	71	35	85	442
*А*2	1 —2 1	—144	31	—81	—109	—97	—400

Учитывая взаимодействие вариантов и пунктов (суммы по четырем), получим, очевидно:

$$A1 \text{ с пунктами} = \frac{56^2+195^2+71^2+35^2+85^2}{8} - \frac{442^2}{40} = 1947{,}40$$

$$A2 \text{ с пунктами} = \frac{144^2+31^2+81^2+109^2+97^2}{24} - \frac{400^2}{120} = 731{,}17$$

Всего: 2678,57

Сумма совпадает с ранее вычисленной другим путем.

Очевидно, во второй группе (*А*2) не надо искать значащих контрастов, так как сумма для четырех степеней свободы много меньше минимального значения для одной степени свободы (1805). В первой же группе такой контраст, может быть, найдется. Для этого разложим сумму 1947,40 по трем степеням свободы. Ясно, что главный контраст — одна группа (Носковцы) против четырех остальных пунктов. Получим разложение (табл. 49г).

Таблица 49г

Степень свободы	Черкассы	Носковцы	Чабаны	Совхоз им. Тельмана	Крыныца	Δ	δ	$\frac{\delta^2}{\Delta}$
	56	195	—71	35	85			
I	—1	4	—1	—1	—1	160	533	1775,56
II	—1	0	1	1	1	32	65	132,03
III	0	0	—1	0	1	16	14	12,25
IV	1	0	0	—1	0	16	21	27,56
							Всего:	1947,40

Как видим, противопоставление Носковцев всем остальным приближается к минимальному уровню значимости.

Более резкое различие мы получим при втором способе. Там разности для двух степеней свободы по вариантам будут иметь такое значение (табл. 49д).

Таблица 49д

Степень свободы	Характер вычисления	Черкассы	Носковцы	Чабаны	Совхоз им. Тельмана	Крыныца	Сумма
*А*1	—2 1 1	156	277	147	107	176	863
*А*2	0 —1 1	—44	113	—5	—37	—6	21

Суммы квадратов для двух групп по четырем степеням свободы равны:

*А*1 с пунктами— 662,72
*А*2 с пунктами—2005,85
Всего: 2678,57

Здесь первая группа не дает никакой перспективы для нахождения значащих контрастов. Ограничимся поэтому разложением второй группы (табл. 49е).

Таблица 49е

Степень свободы	Черкассы	Носковцы	Чабаны	Совхоз им. Тельмана	Крыныца	Δ	δ	$\frac{\delta^2}{\Delta}$
I	—1	4	—1	—1	—1	160	544	1849,60
II	—1	0	1	—1	1	32	70	153,13
III	0	0	—1	0	1	16	—1	0,06
IV	1	0	0	—1	0	16	—7	3,06
							Всего	2005,85.

Здесь первая степень свободы дает контраст минимального уровня значимости: Носковцы всюду выделяются против всех остальных пунктов, которые между собой не показывают никаких различий. Все говорит за то, что данные по Носковцам не могут быть приняты безоговорочно и что эти данные представляют резкий контраст по сравнению с остальными четырьмя пунктами, не обнаруживающими никаких неувязок и могущими поэтому служить основанием для суждения об эффективности в 1939 г.

4.10. ОБРАБОТКА НЕУРАВНОВЕШЕННЫХ ДАННЫХ

Все изложенные методы (рандомизированных блоков, латинского квадрата и факториальной схемы) построены на предположении, что каждый вариант опыта представлен одинаковым числом дат. Поэтому в каждом сравнении участвуют всегда вполне сбалансированные данные опыта, но соблюсти требование такой уравновешенности не всегда удается (та или иная делянка может быть уничтожена благодаря случайным обстоятельствам: забежала на участок лошадь или корова, на дереве сняли урожай неожиданно, опытное животное погибло от случайных обстоятельств и т.д.)

Конечно, было бы очень неэкономно ликвидировать опыт из-за такого непредвиденного обстоятельства. Беда еще не столь велика при методе рандомизированных блоков, если блоков много, а вариантов опыта мало: тогда, выкинув весь тот блок, в котором произошла подобная неприятность, мы снова получаем вполне уравновешенный материал.

При методе латинского квадрата этого сделать невозможно, а тем более при сложных факториальных схемах, где большей частью вариантов много, а повторностей мало или даже вовсе нет. Поэтому разработаны методы, помогающие выправить этот недостаток.

4.10.1. Вычисление недостающих дат

Этот метод наиболее распространен и его всего удобнее применять там, где число выпадающих дат невелико. В нашей литературе он подробно изложен у Н. Ф. Деревицкого (1933) и основан на применении метода наименьших квадратов. Не давая математического обоснования, постараюсь показать смысл такого вычисления, а затем на конкретном примере приведу все вычисления.

Предположим, что в нашем материале из общего числа *N* дат часть дат отсутствует. Для простоты возьмем, что отсутствует одна дата. Каким будет ее наиболее вероятное значение? Если бы все даты соответствовали одному варианту опыта, то, очевидно, что наиболее вероятным значением выпавшей даты было бы среднее арифметическое из всех дат. Это общее среднее арифметическое, назовем его *М,* равно

$$\frac{S+a}{N}$$

где S обозначает сумму всех имеющихся дат, *a* — значение недостающей даты. Значит,

$$a\frac{S+a}{N}=\frac{S}{N-1}$$

Но наш материал состоит не из однородных дат, а даты распределяются по *m* вариантам в *n* повторениях. Наша недостающая дата относится к какому-то одному варианту *i*, и поэтому ее обозначим a_i. Поэтому мы вправе принять, что ее наиболее вероятное значение будет отклоняться от общего среднего на ту разницу, на которую средняя варианта, к которому относится недостающая дата, отклоняется от общего среднего.

Но этого мало. Наша дата относится также к одной из повторностей, и поэтому ее следует обозначить не a_i, а a_{ij}, поскольку между повторностями могут быть существенные различия. Следовательно, влияние той повторности, к которой относится наша недостающая дата, тоже скажется в смысле прибавки разницы между средней данной повторности и общей средней. Мы получаем общую формулу:

$$a_{ij}=M+(M_i-M)+(M_j-M)=M_i+M_j-M,$$

$$a_{ij}=\frac{1}{n}(S_i+a_{ij})+\frac{1}{m}(S_j+a_{ij})-\frac{(S+a_{ij)}}{mn}=\frac{mS_i+nS_j\_S}{mn}$$

где *Si* — сумма всех имеющихся дат того варианта, куда относится наша выпавшая дата a_{ij}; S_j — сумма имеющихся дат той повторности, где находится наша выпавшая дата.

При латинском квадрате мы имеем уже группировку по трем признакам (строки, столбцы и варианты опыта), при греко-латинском квадрате и факториальном опыте мы можем иметь очень большое число группировок. Но изложенный принцип применим и там: сколько бы не было способов группировки, мы суммируем для каждой группировки средние по модальности, в которую входит наша выпавшая дата, и затем вычитаем число общих средних» на единицу меньшее, чем число наших способов группировки. Этот способ применим и в том случае, когда отсутствует не одна дата, а несколько. Конечно, в таких случаях нецелесообразно выводить общую формулу для вычисления выпавших дат, а составлять соответствующее число уравнений, которые затем и решаются. Покажем на примере, как это делается. В качестве материала возьму данные М. Д. Таранухи о влиянии породы кормового растения и освещения на плодовитость непарного шелкопряда. Материал приведен в табл. 50 и дает количество отложенных яиц на одну самку.

Таблица 50

Количество отложенных яиц на самку у непарного шелкопряда (данные М. Д. Таранухи)

Повторность	Дуб		Клен		Яблоня		Всего
	освещ.	затенен.	освещ.	затенен.	освещ.	затенен.	
I	155	137	194	25	117	128	756
II	128	134	194	58	140	184	838
III	194	73	154	63	*a*1	*a*2	484+al+a2
Всего	477	344	542	146	257+*a*l	312+*a*2	2078+*a*l+*a*2

Здесь мы имеем две недостающие даты. Так как по дубу и клену все даты налицо, то суммы дат по этим модальностям фактора пищи вычисляем только для контроля суммы, получаем:

Сумма для дуба — 821
» » клена — 688
» » яблони — 569+*a*1+*a*2
Всего: 2078 + *a*1 + *a*2

По фактору освещения имеем такие суммы:

Сумма для освещенных — 1276+*a*1
» » затененных — 802 + *a*2
Всего: 2078+*a*1 + *a*2

Отсюда получаем (принимая во внимание положение каждой недостающей даты, а также то, что суммы по каждой модальности кормового растения основаны на шести датах, то же и для повторностей, а суммы для модальностей освещения на девяти датах):

$$a1=\frac{569+a1+a2}{6}+\frac{1276+a1}{9}+\frac{484+a1+a2}{6}-2\cdot\frac{2078+a1+a2}{18},$$

$$a2=\frac{569+a1+a2}{6}+\frac{802+a2}{9}+\frac{484+a1+a2}{6}-2\cdot\frac{2078+a1+a2}{18}$$

Приводя к одному знаменателю и соединив подобные члены, получим два уравнения:

$$12a1—4a2=1555,$$
$$12a2—4a1=807.$$

Решив эти уравнения, получим: *a*1=164,75 и *a*2=105,5 или, округляя их до той же степени точности, с которой приведены все наши остальные даты, получим: *a*1=165 и *a*2=106.

Подобный метод, как мы видим, очень легок при недостатке одной-двух дат. Чем больше дат, тем больше и тем сложнее уравнения. У Деревицкого разбирается случай вычисления четырех недостающих дат. При большом числе дат процесс вычисления становится очень сложным.

Наши вычисленные даты вставляются в табл. 50, и затем все операции по анализу варивансы и распределению по степеням свободы ведутся, как и обычно. Надо только иметь в виду: 1) что сравниваются варианты с полным числом дат; 2) при определении числа степеней свободы для ошибки необходимо принять во внимание, что две даты вычислены и число степеней свободы надо уменьшить на две.

В данном случае вариантов у нас 6, число степеней свободы — 5. Повторности не выделяются в особую группу, так как они не образуют блоков, и потому число степеней свободы для ошибки будет (16—1)—5=10, а не 12, как было бы, если бы все даты были получены в опыте. Ясно, что и сумма квадратов, соответствующая этим двум датам, должна быть приведена к десяти степеням свободы. В литературе нет достаточной ясности и единообразия по этому вопросу. Мне представляется, как будто одинаково допустимо применять один из следующих трех методов:

1) вычисленная со включением восстановленных дат сумма квадратов для ошибки делится на полное число степеней свободы (т. е. как будто бы наши восстановленные даты были данными в опыте), но этот средний квадрат считается основанным на числе степеней свободы при исключении восстановленных дат;

2) из общей суммы квадратов и из суммы квадратов для ошибки вычитается сумма квадратов отклонений восстановленных дат от новой арифметической средней;

3) наконец, просто пользуются средним квадратом ошибки, полученным без вычисления недостающих дат (метод, предложенный Снедекором, о котором речь будет дальше).

Разберу все эти приемы на нашем примере.

Новая общая сумма дат равна: 2078+165+106, или 2349. Общая средняя — 130,5.

Сумма квадратов для 16 дат опыта — 310574,000
» » » 2 вычисленных дат — 38461,000
для всех 18 дат — 349035,000
поправка — $\frac{2349^2}{18}$ =306544,500
общая сумма квадратов от *М* — 42490,500.

Сумма квадратов для вариантов получается:

$$\frac{477^2+344^2+542^2+146^2+422^2+418^2}{3}-306544{,}500=31373{,}13$$

Получаем анализ варианcы (табл. 50а).

Сумму квадратов для ошибки 11117,33 делим на 12, но при определении теты считаем, что она основана только на десяти степенях свободы.

При другом способе мы из суммы квадратов для ошибки и из общей суммы вычитаем квадраты разностей вычисленных дат от общей средней, получаем:

$$(165—130{,}5)^2+(106—130{,}5)^2=1790{,}50.$$

Таблица 50а

	Число степеней свободы	Сумма квадратов	Средний квадрат	θ	*p*
Варианты	5	31373,17	6274,63	6,773	<0.01
Ошибка	10(12)	11117,33	926,44		
Всего	15	42490,50			

Вычитая эту величину из 11117,33 и 42490,50, получим новый анализ варианcы (табл. 50б).

Таблица 50б

	Число степеней свободы	Сумма квадратов	Средний квадрат	θ	*p*
Варианты	5	31373,17	6274,63	6,728	<0.01
Ошибка	10	9326,83	932,68		
Всего	15	40700,00			

Вычисление без определения недостающих дат будет дано ниже. Наименьшая тета получается в данном случае по второму методу. Произведем разложение суммы квадратов для вариантов по степеням свободы (табл. 50в).

Таблица 50

A	Дуб		Клен		Яблоня		Δ	δ	$\frac{\delta^2}{\Delta}$	*p*
B	освещ.	затенен.	освещ.	затенен.	освещ.	затенен.				
Сумма	477	344	542	146	422	418				
*A*1	—1	—1	2	2	—1	—1	36	—285	2256,25	<0,001
*A*2	1	1	0	0	—1	—1	12	—19	30,08	<0,01
B	1	—1	1	—1	1	—1	18	533	15782,72	
*A*1*B*	—1	1	2	—2	—1	1	36	655	11917,36	
*A*2*B*	1	—1	0	0	—1	1	12	129	1386,75	

Сумма 31373,16

Для определения значимости наших сравнений умножаем наш средний квадрат 932,68 на теты, соответствующие трем уровням значимости, получим:

для Р, равной 0,05, надо квадрат, равный 4626
» » 0,01 » » » 9364
» » 0,001 » » » 19624

Мы видим, что из пяти степеней свободы для двух (влияние освещения и взаимодействие освещения с породой) имеется очень существенное различие: для остальных трех — не достигается и низший уровень значимости.

Влияние освещения на яблони сказывается в минимальных размерах (весьма вероятно, как полагает и автор М. Д. Тарануха, что благодаря конкретным условиям опыта его влияние не сказывается вовсе на яблони, бывшей в опыте), сильнее на дубе и особенно сильно на клене. Это различие влияния освещения на разных породах и выражено значимостью степени свободы *А*1*В*.

4.10.2. Обработка без вычисления недостающих дат

Для проведения обычного анализа вариансы можно использовать прием, предложенный, например, Снедекором (Snedecor, 1939), который мы применим к тому же примеру.

Общая сумма квадратов берется для наличных 16 дат от их общей средней.

Грубая сумма квадратов равна 310574,000

поправка $-\frac{2078^2}{16}=269880{,}250$

сумма квадратов от средней — 40693,750

Сумма квадратов для вариантов вычисляется аналогично обычному методу, но квадрат суммы по каждому варианту делится на число дат, послуживших для определения той же суммы (или, что одно и то же, каждая сумма умножается на соответствующее среднее). В данном случае четыре варианта основаны каждый на трех датах, а два — на двух, и мы получаем следующую грубую сумму квадратов для вариантов:

$$\frac{477^2+344^2+542^2+146^2}{3}+\frac{257^2+312^2}{2}=302011{,}50$$

поправка — 269880,25

сумма квадратов для вариантов от М — 32131,25

Подобным же образом и для повторностей можно было бы проделать вычисление:

$$\frac{756^2+838^2}{6}+\frac{484^2}{4}-269880{,}25$$

Но так как в данном случае повторности не носят характера блоков, то нецелесообразно вычислять изменчивость для повторностей, а следует присоединить ее к ошибке. Мы получаем анализ вариансы (табл. 50г).

Таблица 50г

	Число степеней свободы	Сумма квадратов	Средний квадрат	θ	P
Варианты	5	32131,25	6426,25	7,505	<0,01
Ошибка	10	8562,50	856,25		
Всего	15	40693,75			

Для суждения о наличии существенных различий между вариантами это вычисление достаточно, но для определения значимости по каждой степени свободы, очевидно, нельзя уже будет пользоваться без какого-то преобразования материала той таблицей ортогональных коэффициентов, которая была использована выше, так как разные варианты основаны на разном числе дат. Чтобы справиться с этим затруднением, можно применить два метода, каждый из которых неодинаков. Поэтому применение обоих способов может служить для проверки допустимости подобной обработки материалов.

4.10.3. Введение уравнительных коэффициентов

В данном случае различные варианты уравновешиваются путем умножения на уравнительные коэффициенты. Для этого берется общее наименьшее кратное числа дат всех вариантов и сумма для каждого варианта умножается на число (уравнительный коэффициент), приводящее число дат к этому общему наименьшему кратному, Так как в данном случае число дат в одних случаях 3, в других — 2, то общее наименьшее кратное — 6 и уравнительные коэффициенты будут соответственно 2 и 3. Перемножив суммы вариантов на эти коэффициенты, мы уже получим уравновешенные суммы, с которыми оперируем; используя нашу старую таблицу ортогональных коэффициентов, получаем новую — табл. 50д.

Таблица 50д

A	Дуб		Клен		Яблоня		Δ	δ	$\frac{\delta^2}{\Delta}$	P
B	освещ.	затенен.	освещ.	затенен.	освещ.	затенен.				
Число дат	3	3	3	3	2	2				
Уравнительные коэффициенты	2	2	2	2	3	3				
Выравненные суммы	954	688	1084	292	771	936				
*A*1	—1	—1	2	2	—1	—1	156	—597	2284,67	
A2	1	1	0	0	—1	—1	60	—65	70,42	
B	1	—1	1	—1	1	—1	84	893	9493,44	<0,01
*A*1B	—1	1	2	—2	—1	1	156	1483	14098,01	<0,0l
A2B	1	—1	0	0	1	1	60	431	3096,02	
							Сумма	2145	29042,56	

Вычисление делителей производится аналогично тому, как делается при уравновешенных опытах, здесь только надо принимать во внимание наличие уравнительных коэффициентов. Так, по первой степени свободы мы имеем (принимая во внимание уравнительные коэффициенты) не такую серию коэффициентов (не обращая внимания на знаки): 1, 1, 2, 2, 1 и 1, а другую: 2, 2, 4, 4, 3 и 3. Возведя каждый такой коэффициент в квадрат и помножая на соответственное число дат, получаем $4\cdot3+4\cdot3+16\cdot3+16\cdot3+9\cdot2+9\cdot2$, или 156. Подобным же образом вычисляем и все остальные делители.

Полученная сумма квадратов вариантов (29042,56) нетождественна с ранее вычисленными (31373,17 и 32131,25); здесь тождества и не может быть, так как путем приведения сумм к одинаковому числу дат мы тоже вычислили недостающие даты, но иным способом, чем ранее. Здесь мы, в сущности, вычисляли недостающие даты на основе наличных дат того варианта, где имеются недостающие даты, игнорируя все остальные варианты. Таким образом, при вычислении была положена иная гипотеза, чем при :прежнем вычислении недостающих дат. Раньше предполагалось, что, например, действие фактора освещения одинаково для всех пород и так как у дуба и клена (где нет недостающих дат) освещение вызвало большую плодовитость, то и для недостающих дат на яблоне недостающая «освещенная» дата оказалась больше, чем «затененная» (165 и 106); мало того, «освещенная» недостающая дата (165) оказалась больше любой из наличных освещенных дат (117 и 140) и, наоборот, недостающая «затененная» дата оказалась меньше обеих наличных дат (сравни 106 с 128 и 184).

Когда же мы работаем с уравнительными коэффициентами, то результат получается такой же, как если бы мы для недостающих дат взяли просто среднюю арифметическую из наличных дат соответствующего варианта. Это не делается просто потому, что оперировать с суммами как с целыми числами удобнее: вместо 165 мы получим 128,5 и вместо 106 получим 156. Какой же способ правильнее? Первый способ является более общим и в тех случаях, когда нет резко выраженного различия во взаимодействии факторов, более правильным. В данном случае, однако, приходится .отдавать предпочтение другому способу, так как ясно видно резкое различие влияния фактора света на фоне различных пород. Для клена действие освещения выражено очень резко, для дуба гораздо слабее, а для яблони по существующим датам мы не имеем никакого намека на положительное влияние. И мы видели, что, даже положив в основу вычисления недостающих дат гипотезу сходного влияния фактора света на всех трех породах, мы все же получили не только значимость факторов света, но и различие взаимодействия при сравнении клена с остальными двумя породами. При нашем втором вычислении уровень значимости для обоих этих степеней свободы *(B* и *A*1*B)* остался того же порядка ($p<0{,}01$), но квадрат разности уменьшился для *B* и увеличился для *AB*1 (используя значение среднего квадрата ошибки, равное 856,25, получаем квадраты разности и для трех уровней значимости соответственно: 4247, 8597 и 18016).

4.10.4. Выравнивание в пределах каждой степени свободы

При предыдущем способе путем умножения на уравнительные коэффициенты все варианты были приведены к одному и тому же числу дат, и в дальнейшем мы уже пользовались старой таблицей ортогональных коэффициентов. Но можно поступать и иначе: производить приведение к одному числу дат для каждой степени свободы отдельно. Этот способ имеет то преимущество, что для некоторых степеней свободы никакого уравнивания проводить не придется, так как для них число дат сравниваемых контрастов и так уравнено, и, следовательно, введение недостающих дат будет проводиться лишь там, где нет такого равенства. Недостатком его является то, что не получается ортогональности системы коэффициентов. Но опять и здесь степень отклонения полученной суммы квадратов от ранее вычисленной покажет нам размеры отклонения от ортогональности. Составим опять таблицу, использовав первоначальные даты, и покажем весь ход вычисления (табл. 50ж).

Таблица 50е

A	Дуб		Клен		Яблоня		Δ	δ	$\frac{\delta^2}{\Delta}$	*p*
B	освещ.	затенен.	освещ.	затенен.	освещ.	затенен.				
Число дат	3	3	3	3	2	2				
Суммы	477	344	542	146	257	312				
A1	—3	—3	5	5	—3	—3	240	—730	2220,42	
A2	2	2	0	0	—3	—3	60	—65	70,42	
B	1	—1	1	—1	1	—1	16	474	14042,25	<0,01
A1B	—1	1	2	—2	—1	1	34	714	14994,00	<0,01
A2B	1	—1	0	0	—1	1	16	188	2209,00	
Сумма	0	—2	8	2	—7	—5		—581	33536,09	

Для первой степени свободы мы противопоставляем данные по клену данным по двум другим породам. Поскольку набор дат по клену равен шести, а сумма дат для двух других пород — 10, то для уравновешения в пределах первой степени свободы мы положительным вариантам (клен) придаем коэффициент 5, отрицательным — 3. Получается равенство $(5\cdot(3+3)=3(3+3+2+2)$. Точно так же для второй степени свободы, где мы противопоставляем дуб яблоне, так как один комплекс вариантов, принятый нами за положительный (дуб), основан на шести датах, а другой (яблоня), принятый за отрицательный — на четырех, то надо взять для положительных сумм коэффициенты 2, для отрицательных — 3.

Остальные три степени свободы не требуют уравновешения, так как они, как нетрудно видеть, и без того уравнены, поэтому там остаются прежние коэффициенты.

Для вычисления делителей применяем обычный прием: только ввиду неодинакового числа дат надо квадраты коэффициентов умножать на соответствующие даты. Так, для первой степени свободы мы имеем: $9(3+3+2+2)+25(3+3)$, или 240; для второй $4\cdot6+9\cdot4=60$ и т. д.

Так как данная система не является ортогональной, то второе требование ортогональности (сумма попарных произведений коэффициентов для любых двух степеней свободы равна нулю) не соблюдается и сумма квадратов не равна ранее вычисленной, но отличается от нее немного: 33536,09 и 32131,25. Опять, как и раньше, мы имеем значимые контрасты (P меньше 0,01) для тех же двух степеней свободы *(B* и *A*1B), причем для данного примера это последнее разложение следует признать наиболее удовлетворительным, так как оба значущих контраста не требовали никакого искусственного уравновешивания.

Проверка вычислений в данном случае не может быть произведена тем способом, как раньше, так как система коэффициентов неортогональна. Проверку разностей можно вести таким образом, что сначала проверяются суммы коэффициентов: сумма произведений коэффициентов на соответственное число дат должна быть равна нулю.

$$0{\cdot}3—2{\cdot}3+8{\cdot}3+2{\cdot}3—7{\cdot}2— 5{\cdot}2==30—30=0.$$

Затем умножаются полученные суммы коэффициентов (0, —2, 8,2, —7 и —5) на соответствующие суммы (177,344 и т. д.) и полученная сумма должна быть равна сумме дельт —581. Этот способ приложим, конечно, и для систем ортогональных коэффициентов, и его там можно рекомендовать в тех случаях, где не получилось точного совпадения двух сумм и мы никак не можем найти ошибки.

Сравнивая все примененные нами три метода, видим, что они вполне согласованные результаты: суммы квадратов для вариантов хотя и отличаются, но незначительно, и, что особенно важно, все три способа дают одинаковые выводы в отношении значимости или незначимости тех или иных контрастов. Всегда получаем существенное различие для фактора света и для его взаимодействия с контрастом клен — другие две породы.

Работа с неуравновешенным материалом, конечно, всегда кропотливее и не столь надежна, как работа с уравновешенным. Но если мы получаем, применяя два способа обработки, сходные выводы, то мы можем считать опыт вполне удавшимся. Расхождение заставляет нас забраковать опыт.

4.11. О ПРЕОБРАЗОВАНИИ ИСХОДНЫХ ДАТ

Дисперсионный анализ приводит к определению среднего квадрата ошибки, основанного на объединении всех индивидуальных дат. При этом предполагается, что каждая индивидуальная дата подвержена одинаковой случайной изменчивости, иначе говоря, что размер случайной погрешности не зависит от значения дат. Между тем это последнее положение далеко не всегда справедливо, и в этом случае размеры среднего квадрата ошибки могут оказаться слишком большими для одних сравнений (отчего выводы теряют в своей значимости) или, что гораздо хуже, слишком малыми для других (отчего может получиться кажущаяся надежность случайных различий). В этих случаях большую пользу могут принести преобразования исходных дат, предложенные Бартлеттом и Блиссом. На примере покажу их применение.

Использованные материалы собраны сотрудником Украинского НИИ плодоводства М. П. Войтенко и касаются эффективности .полихлоридов на личинок майских жуков. Применялись: А) три дозировки (48, 72 и 96 кг/га и, кроме того, был контроль — 0 кг); В) две сетки нанесения уколов в почву (4 и 16 на 1 м^2) и С) две глубины (10 и 20 см). Эффективность измерялась по количеству оставшихся в живых личинок на площадках в 4 м^2, все данные приведены в табл. 51.

Таблица 51

Эффективность полихлоридов против личинок майских жуков, Лохвица, 1939 г. (данные М. П. Войтенко)

A — дозировка		0	48				72				96				Всего
B — сетка			4		16		4		16		4		16		
C — глубина			10	20	10	20	10	20	10	20	10	20	10	20	
Повторности	1	3 4 1 3	3	3		0	1	3	0	0	1	2	0	0	24
	2	13 6 1 2	3	0	0 0 0	0	0	1	0	0	0	0	0	1	27
	3	8 3 5 4	2	2	1	0	0	0	0	0	1	0	0	0	25
	4	14 5 3 3	1	1		0	0	3	0	0	0	3	0	0	34
Всего		38 18 10 12	9	6	1	0	1	7	0	0	2	5	0	1	1100

Работа проведена по той же схеме, как и с парадихлорбензолом (см. главу об обработке результатов сложных опытов), но в той главе с целью избежать влияния крайнего разнообразия цифр на размеры вариансы ошибки при обработке сравнивались только различные варианты опыта. Здесь взят пример особенно резко выраженной неравномерности числа личинок, и для начала мы проделаем обычный анализ вариансы. К прежним 11 степеням свободы здесь прибавят еще 4: *A*1 — противопоставление контроля всем обработкам в целом (коэффициенты для сумм четырех контролей будут, очевидно, 3, а для всех обработок—1, делитель равен 192) и три степени свободы для различий между контролями, которые проведем по схеме:

1	1	—1	—1
1	—1	1	—1
1	—1	—1	1

Так как контроли совершенно одинаковы, то контрасты в пределах контроля не могут иметь реального значения и не должны значительно превосходить средний квадрат для ошибки, поэтому вычисление этих контрастов служит контролем исследования. Общий анализ вариансы приводит к следующему (табл. 51а).

Таблица 51а

Категории изменчивости	Число степеней свободы	Сумма квадратов	Средний квадрат	*t*	p
Варианты	15	363,4375	24,2292	9,42	<0,001
Повторности	3	3,8125	1,2708	0,4943	
Ошибка	45	115,6875	2,5708		
Всего	63	482,9375			

Разлагая сумму квадратов для вариантов по 15 степеням свободы, получаем табл. 51б.

Таблица 51б

Степень свободы	Делитель Δ	Разность δ	$\frac{\delta^2}{\Delta}$	*P*
*A*1	192	202	212,5208	исчез. мала
*A*2	32	8	2,0000	
*A*3	96	8	0,6667	
B	48	28	16,3333	<0,05
C	48	—6	0,7500	
BC	48	—6	0,7500	
A2B	32	8	2,0000	
A3B	96	4	0,1667	
A2C	32	8	2,0000	
A3C	96	12	1,5000	
A2BC	32	4	0,5000	
A3BC	96	12	1,5000	
1	16	34	72,2500	<0,001
2 между контролями	16	18	20,2500	<0,01
3	16	22	30,2500	<0,01
		Сумма	363,4375	

Так как средний квадрат ошибки основан на 45 степенях свободы, то для определения размеров квадратов, соответствующих различным уровням значимости, необходимо к значениям для 60 степеней свободы прибавить одну треть (Так как 60/45 равно 1+1/3) разницы значений теты для 30 и 60 степеней свободы. Мы получаем для трех уровней значимости:

$$\left.\begin{array}{ll} \text{для Р, равной 0,05, —} & 4,06 \\ \text{» 0,01 —} & 7,24 \\ \text{» » 0,001—} & 12,41 \end{array}\right\} \text{x2,5708=} \left\{\begin{array}{l} 10,44 \\ 12,61 \\ 31,90 \end{array}\right.$$

Мы видим, что результат разложения по степеням свободы несколько неожидан. С исключительной надежностью доказана эффективность (для всех вариантов вообще): между вариантами намечается только одно различие по сетке, отнюдь не резко выраженное, но зато очень отчетливы все три сравнения для контролей, где, вообще говоря, мы различий встретить не ожидаем. Само собой разумеется, если при очень многочисленных обработках мы встретим единичный изолированный контраст с необоснованно высокой значимостью, то это особого удивления не вызывает (так как возможность редких случайных значительных отклонений лежит в природе вещей), в особенности если, как в данном случае, мы нарочно выбрали самый парадоксальный случай. Но одновременное появление трех необоснованных контрастов, конечно, наводит на мысль о какой-то ошибке в методике исследования. Это сразу становится ясным, если (так же, как в изложенном уже примере с парадихлорбензолом) при обработке исключить все контроли и сравнивать между собой только 12 обработок. Получаем следующий анализ варианса (табл.51в).

Таблица 51в

Категории изменчивости	Число степеней свободы	Сумма квадратов	Средний квадрат	θ	*p*
Варианты	1	28,1667	2,5606	4,056	близко к
Повторности	3	3,6667	1,2222		0,001
Ошибка	33	20,8333	0,6313		
Всего	37	52,6667			

Как видим, средний квадрат ошибки уменьшился в четыре раза. Разложение по степеням свободы, начиная от *А*2 и кончая *А3ВС,* сохранилось старое, и сумма для этих 11 степеней свободы, как можно проверить, должна в точности равняться 28,1667. Для оценки значимости квадратов ошибки по отдельным степеням свободы надо новый средний квадрат ошибки (0,6318) умножить на тету для 1-й и 33-й степеней свободы, которую получаем из теты для 60 степеней свободы, прибавляя 9/11 интервала между тетами для 30-й и 60-й степеней свободы (так как 60/33 равно 1+9/11).

Получаем для трех уровней значимости:

$$\left.\begin{array}{ll} \text{для Р, равной 0,05, —} & 4,14 \\ \text{» » 0,01 —} & 7,47 \\ \text{» « 0,001—} & 13,05 \end{array}\right\} \text{x 0,6313=} \left\{\begin{array}{l} 2,614 \\ 4,716 \\ 8,238 \end{array}\right.$$

Мы видим, что сравнение по сетке *(В)* приобрело гораздо более высокую значимость (значительно меньше 0,001), чем раньше: остальные сравнения даже при исключении контроля никакой значимости не приобрели.

В данном случае благодаря объединению всего материала полученная средняя ошибка оказалась преувеличенной для суждения различий между отдельными обработками и преуменьшенной для суждения о различиях между контролями. Это объясняется тем, что в данном случае изменчивость числа личинок не следует нормальной кривой изменчивости, а выражается рядом Пуассона в случае нормальной дисперсии. При соответствии же этому ряду среднее квадратическое отклонение зависит от среднего арифметического и эта зависимость выражается простой формулой:

$$\sigma = \sqrt{M}$$ (где M — среднее арифметическое).

Для того чтобы уменьшить зависимость среднего квадратического от абсолютного значения исследуемого признака, Бартлетт (Bartlett) предложил заменять исходные числовые значения преобразованными, именно вместо x брать $\sqrt{x+0,5}$; назовем эту преобразованную величину Z.

Вместо наших исходных значений получим следующие преобразованные:

исходные значения	преобразованные значения	
x	Z	Z^2
0	0,71	0,5
1	1,22	1,5
2	1,58	2,5
3	1,87	3,5
4	2,12	4,5
5	2,35	5,5
6	2,55	6,5
8	2,92	8,5
13	3,67	13,5
14	3,81	14,5

Поставим эти значения вместо исходных значений, с этими новыми значениями проделываем ту же работу по анализу вариансы, как и с исходными данными. Получаем общую сумму — 82,99 и по отдельным вариантам опыта получаем следующие суммы (табл.51г).

Таблица 51г

A	0				48			
в					4		16	
с					10	20	10	20
	12,27	8,89	6,66	7,44	6,54	5,38	3,35	2,84
A	72				96			
в	4		16		4		16	
с	10	20	10	20	10	20	10	20
	3,35	5,67	2,84	2,84	3,86	4,87	2,84	3,35

Общая сумма квадратов равна 142,0000

поправка равна $\frac{82,99^2}{64}$,или 107,6147

Разность — 34,3853

Анализ вариансы дает следующие результаты (табл.51д).

Мы видим, что тета для вариантов оказалась не меньше полученной при первом анализе (там было 9,427), но, как увидим ниже, распределение по отдельным степеням свободы носит иной характер. Мы получаем разложение, принимая точно те же контрасты, что и ранее (табл.52).

Таблица 51д

Категории изменчивости	Число степеней свободы	Сумма квадратов	Средний квадрат	θ
Варианты	15	26,8117	1,78745	10,98
Повторности	3	0,2461	0,08203	
Ошибка	45	7,3275	0,16283	
Сумма	63	34,3853		

Таблица 52

Степень свободы	Делитель Δ	Разность δ	$\frac{\delta^2}{\Delta}$	p
*A*1	192	58,05	17,5111	исчез, мала
A2	32	3,19	0,3180	
*A*3	96	3,63	2,1373	
B	48	11,61	2,8082	<0,001
C	48	—2,17	0,0981	
BC	48	—2,17	0,0981	
A2B	32	3,19	0,3180	
A3B	96	1,59	0,0263	
A2C	32	3,19	0,3180	
A3C	96	4,79	0,2390	
A2BC	32	1,15	0,0413	
A3BC	96	4,79	0,2390	
1	16	7,06	3,1152	<0,001
2 между контролями	16	2,60	0,4225	
3	16	4,16	1,0816	<0,05

Сумма 1,0816

Для суждения о значимости контрастов умножаем средний квадрат ошибки на теты для ответственных уровней значимости: получаем квадраты для трех уровней значимости:

$$\left.\begin{array}{l}\text{для Р, равной 0,05,— 4,06}\\ \text{» » 0,01 — 7,24}\\ \text{» » 0,001— 12,41}\end{array}\right\}\text{x}0{,}16283=\left\{\begin{array}{l}0{,}661\\ 1,\ 179\\ 12{,}021\end{array}\right.$$

Мы видим, что контраст по сетке приобрел значительно большую значимость, чем при первом разложении. Из вызывавших недоумение сравнений между контролями потеряло всякую значимость одно сравнение и стало малосущественным другое (осталось высокозначимым лишь одно из трех), конечно, как изолированный случай такие случаи возможны. Преобразование Бартлетта, таким образом, значительно улучшило разложение. Можно проделать и вычисление, аналогичное тому, которое было сделано с исходными данными, именно исключение контролей для уточнения сравнения между вариантами обработок. Получаем новые анализы варианcы (табл.52а).

Таблица 52а

Категории изменчивости	Число степеней свободы	Сумма квадратов	Средний квадрат	θ	*P*
Варианты	11	4,641306	0,42194	4,19	около
Повторности	3	0,574473	0,19149		0,001
Ошибка	33	3,322702	0,10069		
Всего		8,538481			

Квадраты, соответствующие трем уровням значимости, оказываются равными:

$$\left.\begin{array}{l}\text{для Р, равной 0,05, — 4,14}\\ \text{» » 0,01 - 7,47}\\ \text{» » 0,001— 13,05}\end{array}\right\}\text{x } 0{,}10069=\left\{\begin{array}{l}0{,}417\\ 0{,}752\\ 11{,}314\end{array}\right.$$

Мы видим, что средний квадрат ошибки несколько уменьшился (примерно в полтора раза), так что, видимо, преобразование Бартлетта не устраняет полностью зависимости вариансы от абсолютных размеров. Несмотря на большую точность, ни один новый контраст не приобрел значимости, но значимость сравнения по сетке еще повысилась (само собой разумеется, что квадраты для 11 степеней свободы от *А*2 до *А*3*ВС* можно взять из последнего разложения).

То обстоятельство, что преобразование Бартлетта не устранило, а только смягчило зависимость вариансы от средней, объясняется в значительной степени тем, что средние величины числа индивидов для многих вариантов очень малы. Бартлетт, вообще, указывает (цитирую по сообщению Ч. Блисса), что если среднее число индивидов для любой повторности варианта больше десяти, то целесообразнее употреблять преобразование $\sqrt{x}$, а не $\sqrt{x}$+0,5, если же среднее число варьирует от 2 до 10, то более точным будет преобразование $\sqrt{x}$+0,5. При x меньшем 2 получаются неправильности. Поэтому целесообразно увеличивать размеры площадок так, чтобы среднее число индивидов достигало указанного размера (не менее двух).

Другое преобразование касается случая, когда исходные данные выражены в процентах. Задача была поставлена Ч. Блиссом, который указал, что и в случае процентов мы имеем ясную зависимость вариансы от средней величины. Р.А. Фишер показал преобразование процентов путем введения эквивалентного угла θ по формуле $p=\sin^2\theta$, где p — исходный процент. Эта зависимость устраняется Ч. Блиссом (1937) который составил таблицы для преобразованных 9 и дал пример их применения (аналог современного метода φ-Фишера, 1958).

Блисс показал, что такое преобразование выгодно только тогда, когда в нашем материале имеются проценты, близкие или к нулю (0—15%), или к 100% (85—100%), в остальных же случаях оно не дает улучшения результатов по сравнению с простым использованием процентов.

Несомненно, что преобразования Бартлетта и Блисса не исчерпывают случаев зависимости вариансы от средней арифметической и здесь предстоит еще большая работа. Если у нас возникают сомнения в том, что в каком-либо конкретном случае имеет место такая зависимость, то наиболее простым способом убедиться в этом будет проведение анализа вариансы двояким путем: по всему материалу и с исключением какого-либо одного крайнего варианта или повторности. Если средний квадрат ошибки при обоих способах вычисления покажет заметную разницу (как это получилось у нас), то это ясное указание на наличие зависимости вариансы от среднего значения и тогда выводы, основанные на анализе материала в целом, будут менее надежны, чем при частичной обработке его.

4.12. АНАЛИЗ КОВАРИАНСЫ

Все изложенные выше методы (рандомизированных блоков, латинского и греко-латинского квадратов, факториальная схема) имеют один общий признак: каждая единица исследования, будь то единичный растительный или животный организм, делянка поля, цветок и т. д., давала только одну дату, непосредственно нас интересующую. Само собой разумеется, что все эти индивиды отличались и многими другими признаками, как-то: возраст, сорт, пол, расположение на более плодородных или менее плодородных участках и т. д. Все эти различия рассматривались как факторы исследования и входили в общую схему, или же, как в методе рандомизированных блоков, исследованные организмы собирались в группы, по возможности однородные, и путем анализа вариансы различие между этими группами не смешивалось с ошибкой опыта и потому не вызывало ее увеличения.

Во всех этих методах широко проводился принцип эквализации (рандомизации), т. е. уравнения возможных источников ошибки при сравнении отдельных вариантов. Этот принцип эквализации известен давно: при опытах с деревьями рекомендуется выбирать деревья одного сорта, возраста и т. д., и иногда такое стремление к эквализации заходит даже слишком далеко в смысле того, что опыты, поставленные на не чисто сортном материале, огульно берутся под сомнение.

Само собой разумеется, что по мере возможности надо этот принцип проводить, но полное уравнение индивидуумов исследования фактически невозможно. Если мы работаем в поле, то, разделив поле на более или менее сходные участки, выделенные под отдельные блоки, мы все же всех различий не уравняем. Метод латинского квадрата представляет дальнейший шаг в этом направлении» но и при нем различия остаются. Как бы мы ни подбирали однородные деревья, различия между ними будут, и иногда эти различия сказываются чрезвычайно резко. Во многих случаях эти особенности если не могут быть уравнены, то могут быть тем или иным путем измерены, и тогда наряду с признаком, наиболее нас интересующим, появится один или несколько сопутствующих признаков, служащих для характеристики каждого из индивидуумов исследования.

Например, при работе на поле наиболее интересующим нас признаком будет, конечно, урожай разных вариантов опыта. Но если мы за год до опыта организуем так называемый разведочный посев, т. е. засеем намеченное для опыта поле какой-либо совершенно однородной культурой, то урожай этой культуры в год, предшествующий опыту, является характеристикой почвенного плодородия и может быть принят для внесения поправки в урожай того года, когда ставился опыт. Для деревьев опять-таки наиболее интересующим нас признаком является урожай дерева. Однако урожай зависит не только от той или иной обработки, но и от количества завязавшихся плодов, и этот последний признак (хотя он нас непосредственно и не интересует) может быть использован как сопутствующий для внесения поправок в результаты исследования урожая, и прежде всего в уточнении результатов опыта. Такой подход в методике полевого опыта известен был давно, и практика производства разведочных посевов показывает, насколько остро чувствовалась необходимость в таком подходе. И здесь, как и в других случаях, часто вдавались в крайность, например в утверждение, что никакие полевые опыты без предварительных разведочных посевов невозможны. Однако часто бывает рационально и выгодно обходиться без разведочных посевов, в других же случаях они оказываются очень полезными.

Обычное использование данных разведочных посевов страдает шаблонностью, которая часто сильно отражается на точности. Например, очень часто рекомендуется принимать плодородие отдельных участков поля в данном году пропорциональным плодородию прошлого года. При этом упускается из виду, что пропорциональность может быть не только прямая, но и обратная. Пониженные участки поля в засушливые годы дадут максимальный урожай, в дождливые же, напротив, дадут минимальный. Поэтому необходимо применять такой метод, которой в самом ходе исследования позволял бы выяснить, каково соотношение плодородия отдельных участков поля в разные годы. К таким методам относятся использование линий регрессии и анализ ковариансы, являющийся обобщением метода использования линий регрессии. Анализ ковариансы отличается исключительной гибкостью: он может использовать один или несколько сопутствующих признаков, регрессия основного интересующего нас признака по сопутствующему может быть прямолинейной или криволинейной, все эти особенности изучаются на нашем же материале и используются для максимального уменьшения случайной ошибки в случае рандомизированного исследования или для устранения систематических ошибок при нерандомизированном исследовании.

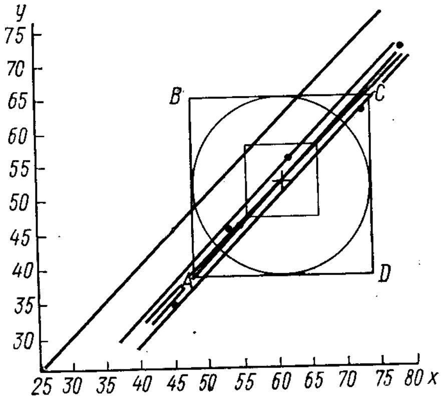

Рис. 6. Анализ вариансы методом редуцированной суммы квадратов и линии регрессии средних. (Объяснения см. в тексте)

Основным моментом анализа ковариансы является получение редуцированной суммы квадратов для зависимой переменной на основе полученного коэффициента регрессии. Такая редукция проводится на основании некоторых простых формул, но чтобы понять чисто физический смысл такой редукции, полезно подробно разобрать простейший пример, который нами уже частично был использован. Именно в главе об основах дисперсионного анализа был приведен пример сравнения урожайности двух сортов и показано, как путем привлечения коэффициента корреляции средняя ошибка разности была сильно сокращена (редуцирована). Изобразим все наши данные на рисунке (рис.6). По оси абсцисс мы возьмем значения одного сорта — *x,* а по оси ординат другого — *y*. Для первого участка возьмем на оси абсцисс значение 62 и на оси ординат — 51 и на пересечении двух перпендикуляров поставим точку. Таким же образом поставим и все остальные пять точек. Средние арифметические, как показано на с. 75, соответственно равны 61 и 52; пометим в виде креста положение обеих средних. Ошибки средних арифметических, как известно, равны 5,29 и 5,45; отложив эти значения в обе стороны от центра креста по обоим направлениям и соединив их овальной линией, получим графическое обозначение средней ошибки средних арифметических. Средние квадратические отклонения (являющиеся, как известно, средними ошибками единичных наблюдений) будут для наших переменных соответственно:

$$\text{для } x = \sqrt{\frac{838}{5}}, \text{ или } \sqrt{167{,}6} = 12{,}9$$

и для $y = \sqrt{\frac{892}{5}} = \sqrt{178{,}4} = 13{,}4.$

Отложим средние квадратические отклонения для *x* вправо и влево по обе стороны от 61 (среднего арифметического для *x),* а таковые для *y* — кверху и книзу от 52; проведя овал, получим изображенный на рисунке эллипс, характеризующий размеры изменчивости обеих переменных около своих арифметических средних. Мы видим, что в этих двух случаях из шести точек четыре лежат между противоположными сторонами прямоугольника, а две — за их пределами. Это отвечает тому, что среднее квадратическое отсекает примерно две трети площади общей кривой. Теперь зададим себе вопрос: когда мы сравниваем два сорта и наметили их на чертеже, нельзя ли графически изобразить ту нулевую гипотезу, которую мы проверяем? Так как мы хотим узнать, указывает ли наш материал на отсутствие или наличие существенных отличий между сортами, то, очевидно, испытываемой нами нулевой гипотезой является гипотеза отсутствия какого бы то ни было различия между сортами. Опровержение этой гипотезы явится вместе с тем доказательством наличия существенного различия между сортами.

Но если сорта не имеют существенных отличий, то каждому значению одного сорта соответствует совершенно такое же значение другого сорта (вернее, отличающееся от него лишь в рамках ошибки опыта), и, следовательно, графически нулевая гипотеза отсутствия различий между сортами выразится прямой линией, проходящей через начало координат и делящей пополам угол между осями, иначе говоря, уравнением *y=x.* Эту линию мы и изобразим на рис.6. Теперь, если пока не обращать внимание на расположение отдельных точек на рисунке, а обратить внимание только на положение средних арифметических и эллипс, отсекающий средние квадратические обоих сортов, то мы придем к выводу, что наших данных недостаточно, чтобы опровергнуть нулевую гипотезу. Мы видим, что эллипс лежит по обе стороны от диагонали, среднее арифметическое лежит так близко к диагонали, что эллипс, отграничивающий зону средней ошибки, почти касается диагонали — это и есть графическая иллюстрация того результата, что сравнение разницы средних с ошибкой разности не дает существенного отличия между сортами.

Но теперь обратим внимание на распределение точек на рис.6. Мы видим, что расположены они с ясно выраженной закономерностью и все вместе если и не образуют прямой линии, то лежат в узкой прямой полосе — это ясное выражение корреляционной связи, выражение факта, что с увеличением урожая сорта *x* увеличивается и урожай сорта *y.* Мы уже использовали эту связь для вычисления коэффициента корреляции и для внесения этим путем поправки в размеры средней ошибки. В данном случае удобнее использовать коэффициент регрессии одного сорта по другому, это и нагляднее, и послужит хорошим подспорьем для понимания сущности анализа ковариансы. Коэффициент регрессии, как известно, определяется по формуле:

$$R\frac{y}{x} = \frac{\sum xy}{\sum x^2}, \text{ или } \frac{\sum \alpha_x \alpha_y}{\sum a_x^2}$$

Коэффициентов прямолинейной регрессии всегда бывает, конечно два смотря по тому, которую из переменных *x* или *y* примем за независимую, а знаки сумм даны двояко, смотря по тому, обозначаем ли мы знаками *x* и *y* значения переменных в их расстоянии от соответствующих средних арифметических, или же *x* и *y* дают просто значения переменных и тогда отклонения от средних арифметических выражаются буквами α_x и α_y.

Подставляя в формулы для коэффициентов регресии полученные нами раньше значения, получим:

$$R(\frac{y}{x}) = \frac{858}{838} = 1{,}024 \text{ (точнее, 1,023866),}$$

$$R(\frac{y}{x}) = \frac{858}{892} = 0{,}962.$$

Коэффициент корреляции является, как известно, средним геометрическим обоих коэффициентов регрессии и равен 0,992. Получаем два уравнения регрессии:

$$y = 52 + 1{,}024(x — 61),$$

$$x = 61 + 0{,}962(y — 52).$$

Подставляя в первое уравнение регрессии, положим, два значения *x,* равные 41 и 81 (эти два наиболее удобны для вычисления и, кроме того, охватывают всю амплитуду изменчивости нашего материала), получим *y,* соответственно, равный 31,52 и 72,48; для второго уравнения, подставляя значение *y,* равные 32 и 72, получим *x,* равный 41,76 и 0,24. Проведя соответственно две прямые (обе проходят, конечно, через точку, соответствующую обоим средним арифметическим), мы увидим, что они почти совпадают, и это вполне естественно, так как коэффициент корреляции очень высокий.

Взяв одну из этих линий, линию регрессии *y* по *x* (в данном случае выбор линии регрессии – произвольный, так как обе переменные могут быть приняты за независимую переменную), определим расстояние около линии регрессии. Оно выражается формулой

$$\sum y = \sigma y \sqrt{1 - r^2}$$

где Σу (это не знак суммы) обозначает среднее квадратическое отклонение около линии регресии; σ^2 – среднее квадратическое отклонение у около среднего арифметического, *r* – коэффициент корреляции.

Мы получаем

$$13{,}4\sqrt{0{,}0159}=13{,}4\cdot 0{,}126=1{,}69$$

Мы видим, что как только мы приняли во внимание расположение точек, так тотчас среднее квадратическое отклонение y уменьшилось примерно в восемь раз. Отложим от нашей линии регрессии кверху и книзу величину 1,69, проведем параллельные линии и мы получим сравнительно узкую полоску, заключающую в себе (поскольку расстояние от линии регрессии равно средней квадратической) примерно две трети общего числа вариантов. И действительно, мы видим, что из шести точек четыре лежат в этой полоске, одна на границе и одна за пределами полоски, т. е. по отношению к полоске расположение их ничуть не хуже расположения тех же точек относительно прямоугольника *ABCD*. Но площадь полоски значительно меньше площади прямоугольника, в этом и заключается редуцирование изменчивости (основной процесс в анализе ковариансы) путем использования линии регрессии. Вместе с тем уже из рис. 6 ясно видно, что полоса около линии регрессии значительно отстоит от прямой линии, биссектрисы прямого угла, графически изображающей нулевую гипотезу — отсутствие различия сортов, и, следовательно, наличие сортовой разницы делается совершенно наглядным.

Для того чтобы арифметически показать редукцию изменчивости y при использовании коэффициента регрессии, приведем опять уравнение регрессии y по x:

$$y=52+1{,}023866\,(x—61)=1{,}023866x—10{,}455826$$

(коэффициент регрессии вычислен с большим числом знаков только для получения близкого совпадения вычисляемых величин) .

y в дальнейшем будет обозначать вычисленные по данному уравнению регрессии величины для разных x, y — наблюденные значения для тех же значений x, y — как это принято сейчас, среднюю арифметическую для y.

Расстояние каждой нашей точки от арифметической средней $y—\bar{y}$ может быть разложено на две части:

$$y-\bar{y}=(y-Y)+(Y-\bar{y}).$$

Первая величина $y—Y$ обозначает расстояние наших наблюденных величин от вычисленных значений линии регрессии, а $Y—y$ — расстояние вычисленных величин от среднего арифметического. Как доказывается в курсах математической статистики, справедливо также и равенство

$$\Sigma(y-y)^2=(y-Y)^2+\Sigma(Y-y)^2,$$

т. е. общая сумма квадратов отклонений от общего арифметического среднего равна сумме квадратов отклонений наблюденных величин от прямой регрессии (это есть рассеяние около линии регрессии) плюс сумма квадратов расстояний тех же точек линий регрессии от общего среднего (это будет рассеяние по линии регрессии). Таким образом, общее рассеяние равно сумме рассеянии около линии рассеяния и по линии рассеяния. Вместо общего математического доказательства покажем на примере, что это равенство вполне справедливо. Приведем все вычисления (табл.53).-

Таблица 53

x	y	Y	$y-Y$	$Y-\bar{y}$
62	55	53,023866	1,976134	1,023866
45	34	35,618144	—1,618144	—16,381856
53	45	43,809072	1,190928	—8,190928
54	45	44,832938	0,167062	—7,167062
73	62	64,286392	—2,286392	12,286392
79	71	70,429588	0,570412	18,429588
Сумма				
366	312	312,000000	3,904536	31,739846
			—3,904536	—31,739846

Правильность вычислений корректируется совпадением сумм y и У и равенством нулю обоих сумм $y—Y$ и $Y—\bar{y}$.

Возведя в квадрат все значения $y—Y$ и $Y—\bar{y}$, получим следующие суммы (табл.53а).

Точное значение общей суммы квадратов около y равно, как мы знаем, 892 и совпадает с вновь вычисленным.

Таблица 53а

Рассеяние	Число степеней свободы	Сумма квадратов
Около линии регрессии $\Sigma(y-\text{У})^2$	4	13,522673
По линии регрессии $\Sigma(Y-\bar{y})^2$	1	878,476729
Общее $\Sigma(Y-\bar{y})^2$	5	891,999402

Эта сумма (обозначающая просто Σy^2, если понимать под y только отклонение от арифметического среднего), равная 892, оказывается редуцированной приблизительно в 66 раз, но при этом одна степень свободы утрачена. Совершенно ясно, почему именно только одна степень свободы. При первоначальных вычислениях изменчивости около арифметического среднего y мы имели 6—1 (или 5) степеней свободы (естественно, что когда пять величин уже изменились, то шестая вычисляется на основании суммы первых пяти и общей суммы, послужившей для вычисления средней арифметической). Но, используя прямую линию регрессии, мы вводим новый параметр. Вычислив на основе старой константы (среднеарифметической) и новой (коэффициента регрессии) теоретические значения y, мы на это используем одну степень свободы (поскольку сам новый параметр, вообще говоря, изменчив). Но так как все величины $Y—y$ получаются вычислением, то все шесть цифр этой категории соответствуют одной степени свободы.

Примененный прием разложения общей суммы квадратов на компоненты дан исключительно для лучшего понимания всего процесса и значения анализа ковариансы. Он, конечно, слишком громоздок и при большом числе дат применяются несравненно более сжатые и удобные приемы вычислений, в основе которых лежит тот же прием: разложение общей изменчивости около арифметической средней на изменчивость около линии регрессии и на изменчивость по линии регрессии (последняя обычно и не вычисляется).

В качестве примера применения анализа ковариансы разберем материал по эффективности борьбы с плодожоркой, собранный сотрудником Украинского НИИ плодоводства В. П. Роде в 1939г. в совхозе «Зеленый Яр». В опыте находилось 25 деревьев сорта «бойкен»: имелось пять вариантов (4 различные обработки и контроль — отсутствие обработки), расположенных в пяти рандомизированных блоках. Обработки заключались в опыливании тремя инсектицидами: меритолем, арсенатом кальция и парижской зеленью. Приведем сначала данные по каждому дереву по наиболее интересующему нас признаку — количеству яблок в урожае в десятках плодов. Обозначим этот признак буквой у (табл.53б).

Таблица 53б

Вариант	Урожай — *у в* десятках плодов					
	1	2	3	4	5	Сумма-
Меритоль, опыление	106	171	80	44	128	529
Арсенат кальция, опыление	125	108	73	38	47	391
Парижская зелень, опыление	87	127	118	29	33	394
Парижская зелень, опрыскивание	72	32	38	40	19	201
Контроль	64	111	20	48	4	247
Сумма	454	549	329	199	231	1762

Разберем сначала этот материал, не привлекая никакого сопутствующего признака. Мы видим, что по урожайности варианты дают значительное отличие: суммы по пяти деревьям колеблются от 201 до 529, по повторностям колебания урожайности ничуть не меньше (от 199 до 549). Следовательно, несомненно, в саду очень значительные колебания урожайности, независимые от обработки.

Простое сравнение показывает, что опрыскивание парижской зеленью в этом году на данном сорте не эффективно, но не ясно, в какой мере доказаны эффективность опылений и в какой мере существенно различие между меритолем и двумя другими опыливаниями: можем ли мы утверждать, что меритоль в данных условиях является наиболее эффективным инсектицидом. Анализ варианса, приложенный к данному признаку, дает следующий результат (табл. 53в).

Различие между повторностями, как видим, более существенно, чем различие между вариантами. Разлагая сумму квадратов для вариантов по степеням свободы, удобнее всего будет противопоставить три варианта опыления остальным двум и в пределах этих двух групп уже проводить дальнейшее сопоставление порознь, получаем следующие данные (табл.53г).

Таблица 53в

Категории изменчивости	Число степеней свободы	Сумма квадратов	Средний квадрат	θ	*p*
Варианты	4	13687,84	3421,96	3,819	<0,05
Блоки	4	17558,24	4389,56	4,898	<0,01
Ошибка	16	14338,16	896,135		
Сумма	24	45584,24			

Таблица 53г

Опыление			Опрыскивание парижской зеленью	Контроль	Δ	δ	$\frac{\delta 2}{\Delta}$
меритоль	арсенат кальция	парижская зелень					
529	391	394	201	247			
2	2	2	—3	—3	150	1284	10991,040
0	0	0	1	—1	10	—46	211,600
—1	2	—1	0	0	30	—141	662,700
1	0	—1	0	0	10	135	1822,500

Сумма 13687,840

Единственный значащий контраст дает первая степень свободы: противопоставление вариантов с опылением двум остальным.

$$\theta = \frac{10991{,}040}{893{,}135} = 12{,}26$$

соответствует вероятности отсутствия существенной разницы, меньшей 0,01 (для P=0,01 достаточна θ=8,53). В пределах вариантов с опылением нет никакого намека на существенное различие. Конечно, взятое нами разложение (третья степень свободы) не дает максимально возможного контраста: в этом случае лучше было бы противопоставить меритоль двум остальным опыливаниям, но даже если бы сумма квадратов, соответствующая этим двум степеням свободы (2485,2), была сосредоточена на одной степени свободы, это дало бы θ=2485,2/896,115,или 2,77, далеко не достаточную даже для минимального уровня значимости (для чего требуется 4,41).

Посмотрим теперь, какое улучшение результатов даст нам использование сопутствующего признака, в данном случае суммы плодов в падалице и урожае, выраженной тоже в десятках плодов и обозначаемой как *х,* данные приведены в табл.53д.

Таблица 53д

Вариант	Общий запас в десятках					
	блоки					
	1	2	3	4	5	Сумма
Меритоль, опыление	180	252	139	112	191	874

Арсенит кальция, опыление	159	182	129	72	86	628
Парижская зелень, опыление	131	184	175	73	109	672
Парижская зелень, опрыскивание	180	118	115	137	70	620
Контроль	153	217	128	117	103	718
Сумма	803	953	686	511	559	3512

Данные по общему запасу плодов касаются, конечно, тех самых деревьев, как и данные по урожаю. Если мы нанесем все данные на скаттер-диаграмму (рис.7), беря значения *x* по оси -абсцисс, а значения *y* по оси ординат, то получим ясную картину зависимости урожая от общего запаса плодов. Наличие такой ясной зависимости и позволяет надеяться, что применение анализа ковариансы внесет улучшение в обработку, так как при отсутствии связи анализ ковариансы никакого улучшения не дает.

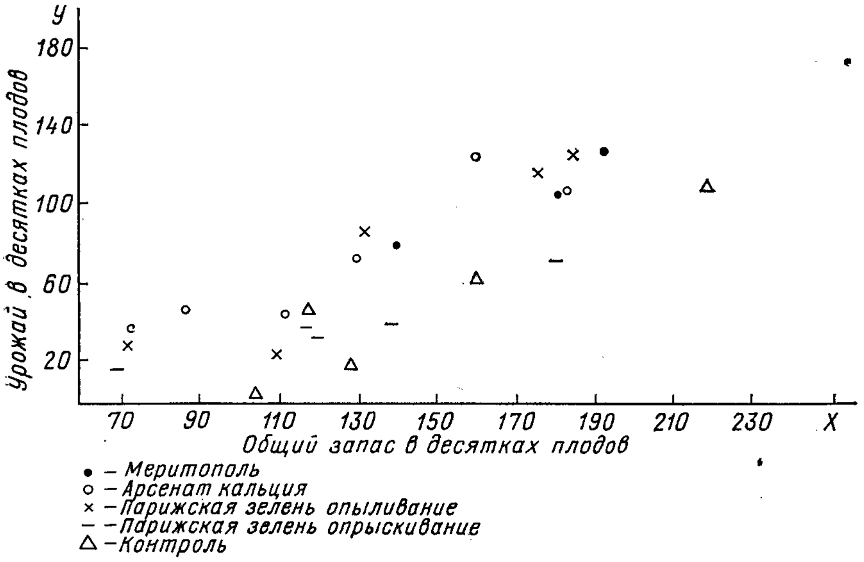

Рис. 7. Сравнительная эффективность препаратов в опытах по борьбе с плодожоркой

Сначала проделаем анализ варианcы по признаку *x*. Это необходимо как один из этапов анализа ковариансы, но кроме того такая обработка дает нам возможность произвести некоторый контроль за правильностью организации опыта. Анализ варианcы дает результат, приведенный в табл. 54.

Таблица 54

Анализ варианcы по запасу плодов *(x)*

Категории изменчивости	Число степеней свободы	Сумма квадратов	Средний квадрат	θ	*P*
Варианты	4	8587,84	2146,26	2,01	>0,05
Блоки	4	26077,44	6519,36	6,10	<0,01
Ошибка	16	17094,96	1068,44		
Всего	24	51760,24			

Анализ дает картину, отвечающую нашим ожиданиям: по повторностям (блоки) существенные различия имеются, следовательно, блоки с самого начала отличались неодинаковым запасом плодов, по вариантам же нет ни намека на существенное различие по запасу плодов, следовательно, нет никаких указаний на какую-либо ошибку в организации опыта.

Для анализа ковариансы необходимо к вычисленным уже двум. суммам (Σx^2 и Σy^2) вычислить ковариансу Σxy, так как весь анализ основан на использовании линии регрессии, коэффициент же регрессии вычисляется по формуле $B = \frac{\sum xy}{\sum x^2}$,т. е. равен сумме произведений обеих переменных от соответствующих средних. арифметических, деленной на сумму квадратов независимой переменной от ее средней арифметической. Ковариансa, как и варианса, вычисляется для всех категорий изменчивости отдельно. И для. вычисления ковариансы можно идти различными путями. При наличии счетной машины удобнее взять нуль за условное среднее, подобно тому, как это делалось и при вычислении варианcы, и затем вычитать из этой первичной суммы соответствующую поправку.

Для вычисления общей ковариансы перемножаем попарно наши исходные даты по *x* и *y*, т. е. 106·180, 171·252, 80·139 и т. д. (всего 25 произведений соответственно 25 деревьям). Поправка же будет равна произведению сумм, деленному на число дат, т. е. $\frac{17623512}{25}$ (она же равна произведению сумм по одному признаку на среднее арифметическое по другому).

Мы получаем:

сумма произведений — 289906,00
поправка — 247525,76
Общая сумма *xy* — 42380,24

В данном случае сумма *xy* положительна, так как оба признака связаны положительной зависимостью, но она может быть и отрицательна (т. е. поправка может быть больше суммы произведений исходных дат).

Аналогичным образом получаем для вариантов:

сумма *xy* от нуля —

$$\frac{529\cdot 874+391\cdot 628+394\cdot 672+201\cdot 62+274\cdot 718}{5}=254925{,}60$$

поправка — 247525,76

сумма *xy* для вариантов — 7399,84

Точно так же для повторностей (блоков):

$$\frac{454\cdot 803+549\cdot 953+329\cdot 686+199\cdot 511+231\cdot 559}{5}=268854{,}20$$

поправка — 247525,76

сумма *xy* для блоков 21328,44

Сумма *xy* для ошибки вычисляется по разности от общей ковариансы (за вычетом ковариансы для вариантов и блоков).

Сводим результаты анализа варианcы по обоим признакам и .анализ ковариансы (табл.54а). Первый и левый столбцы взяты из предыдущих анализов.

Таблица 54а

Категории изменчивости	Число степеней свободы	Σx^2	Σxy	Σy^2
Варианты	4	8587,84	7399,84	13687,84
Блоки	4	26077,44	21328,44	17558,24
Ошибка	16	17094,96	13651,96	14338,16
Сумма	24	51760,24	42380,24	45584,24

Весь смысл анализа ковариансы заключается в том, чтобы изменчивость около среднего арифметического уменьшить, заменив ее изменчивостью около линии регрессии. Так как в таблице у нас все отклонения обоих признаков даны от соответствующих арифметических средних, то линия регрессии (проходящая, конечно, через точку, соответствующую обоим арифметическим средним), будет выражаться уравнением

$$Y=bx,$$

где b — коэффициент регрессии; Y — значение переменной на прямой регрессии, а x — фактически наблюденные величины зависимой переменной. Очевидно, отклонения наблюденных величин от вычисленных (по прямой регрессии) будут равняться $y—Y$, или $y—bx$, и сумма квадратов таких отклонений (то, что мы обозначим редуцированная сумма квадратов *y)* будет равняться $\Sigma(y—bx)^2$; разложив ее, получим $\Sigma y^2—2b\Sigma xy+b^2\Sigma x^2$ (коэффициент регрессии, как величину постоянную, мы можем вынести за знак суммы).

Следовательно, для того, чтобы получить редуцированную сумму квадратов *y,* мы должны к цифре последнего столбца прибавить цифру первого столбца, умноженную на квадрат коэффициента регрессии, и вычесть цифру среднего столбца, умноженную на удвоенный коэффициент регрессии. Так как при вычислении коэффициента регрессии мы должны получить величину, независимую как от вариантов опыта, так и от влияния гетерогенности доля (блоки), то для его вычисления мы пользуемся данными строки «ошибка», получаем:

$$b=\frac{13651{,}96}{17094{,}96}=0{,}798596$$

$$B^2=0{,}637756.$$

Использовав все наши данные (для всех строк мы пользуемся одним и тем же коэффициентом регрессии, и поэтому вычисление последней строки, соответствующей суммарной изменчивости, дает хорошую проверку вычислений первых трех строк), получаем редуцированную сумму квадратов *y* (табл. 54б).

Таблица 54б

Редуцированные суммы квадратов *y*

Категории изменчивости	Число степеней свободы	Сумма квадратов	Средний квадрат	θ	*p*
Варианты	4	7345,821	1836,455	8,018	ок. 0,001
Блоки	4	123,670	30,918		
Ошибка	15	3435,772	229,051		
Всего	23	10905,263			

По сравнению с первоначальным анализом варианcы *y* мы видим два существенных изменения:

1) средний квадрат ошибки почти в четыре раза уменьшился, поэтому, хотя средний квадрат по вариантам также уменьшился, тета выросла более чем в два раза, отчего значимость различий для вариантов почти достигла значимости 0,001;

2) средний квадрат для повторностей, напротив, чрезвычайно уменьшился и стал значительно меньше среднего квадрата ошибки. Поэтому эту сумму квадратов мы можем присоединить к сумме квадратов ошибки. Последняя соответствует не 16 степеням свободы, а лишь 15, так как одна степень свободы использована при вычислении коэффициента регрессии. Прибавив к ним четыре степени свободы повторностей, получаем сумму квадратов для 19 степеней свободы 3559,442, средний квадрат 187,339, тета 9,804 и вероятность отсутствия существенных различий значительно меньшую, чем 0,001 (для этого достаточна тета 7,26).

Редуцированную сумму квадратов для урожая *(y)* можно также разложить по степеням свободы. Для этого наиболее удобно произвести вычисление значений y для отдельных вариантов, соответствующих средним значениям x для тех же вариантов. Так как мы все время оперировали суммами для пяти деревьев каждого варианта (суммами оперировать удобнее, так как мы избегаем дробей и вместе с тем вычисление ведем с максимальной точностью), то и здесь будем вычислять значения y, соответствующие сумме пяти деревьев.

Средний размер урожая пяти деревьев равен: 1762/5, или 352,4, а для запаса плодов 3512/5 или 702,4. Поэтому получаем уравнение регрессии: $5y—352,4=0,798596(5x—702,4)$, или раскрывая скобки и приводя подобные члены, получим

$$5y = 0,798596 \cdot 5x—208,533830$$

(большое число знаков вычисляем только потому, чтобы при проверке суммы квадратов по вариантам получить наилучшее совпадение).

Подставляя в это уравнение последовательно значения $5x$: 874, 628, 672, 620 и 718, получим вычисленные значения y, которые обозначим y''.

Сравнивая их с наблюденными и найдем превышение (или снижение) значений для каждого из вариантов по сравнению с уровнем, даваемым прямой регрессии (табл.54в).

Таблица 54в

Варианты	Наблюденные y	Вычисленные y″	$y—y''$
Меритоль	529	489,43907	39,56093
Арсенат кальция	391	292,98446	98,01554
Парижская зелень, опыление	394	328,12268	65,87732
Парижская зелень, опрыскивание	201	286,59569	—85,59659
Контроль	247	364,85810	—117,85810
		Сумма	0,00000

Сумма отклонений равна, как и должно быть, нулю. Эти значения $y—y''$ и служат уже для характеристики эффективности: на первое место выходит арсенат кальция, на последнее — контроль. Пользуясь этими значениями $y—y''$ вместо прежних сумм у, мы можем вычислить величину квадрата отклонения для каждой степени свободы в отдельности, применив ту же ортогональную систему коэффициентов, как и раньше, получаем табл. 54г. Так как средний квадрат ошибки равен 187,339, то ясно, что единственным значащим контрастом остается лишь контраст по первой степени свободы. Здесь тета равна $\frac{6898,908}{187,339}$ или 36,826, т. е. в три раза больше против полученной ранее без использования сопутствующего признака (12,26). Применение анализа ковариансы значительно повысило надежность нашего вывода, но в данное случае не позволило извлечь никакого нового вывода.

Таблица 54г

Вариант	Разность δ	$\frac{\delta^2}{\Delta}$
1. Опыление — другие варианты	1017,2690	6898,908
2. Опрыскивание парижской зеленью — контроль	32,2624	104,086
3. Арсенат кальция — два других опыления	90,5928	273,569
4. Опыление: меритоль — парижская зелень	—26,3169	69,258
	Сумма	7345,821

Указанное разложение по степеням свободы является наиболее наглядным, так как мы непосредственно оперируем с корректированными значениями зависимой переменной, именно с ее отклонениями от линии регрессии. Но при этом, для проверки правильности вычислений (точного совпадения полученной суммы квадратов с квадратом, ранее вычисленным) нам приходится вести вычисления с большой точностью, т. е. оперировать с многозначными числами. Удобнее поэтому, хотя и не так нагляден, другой метод, который нам пригодится и при изложении некоторого уточнения анализа ковариансы. Для этого надо произвести разложение по степеням свободы суммы квадратов для вариантов независимой переменной x по той же схеме, как было произведено для зависимой переменной y (см. табл. 54г), т. е. умножать суммы для вариантов x (иначе говоря, суммы запаса плодов на деревьях, предназначенных для каждого варианта обработки) на соответствующие коэффициенты. Получим серию разностей для каждой степени свободы, которые и приведем в табл.55, наряду с переписанными разностями для тех же степеней свободы для y. В следующих графах показаны квадраты для соответствующих степеней свободы для x и y и ковarianса xy. Все они получаются аналогичным способом: для x — возведением в квадрат разности (для первой степени свободы 334) и делением квадрата на делитель 150. Для ковариансы же умножаем разности x и y (например, .для первой степени свободы 334 и 1284) и делим на соответствующий делитель (для первой степени свободы 150). Мы получаем для вариантов ковариансу 8587,84, уже полученную ранее в виде общей суммы (см. табл. 54а), но здесь она разложена по степеням свободы. Чтобы получить значения редуцированных квадратов y для тех же степеней свободы, проделаем совершенно те же вычисления, которые мы проделали, чтобы получить суммы редуцированных квадратов в табл. 546, т. е. используем формулу

$$y^2—2b \cdot xy+b^2x:^2.$$

К столбцу y^2 табл.55 прибавляем значение x^2, умноженное на квадрат коэффициента регрессии, т. е. на 0,637756, и вычитаем -значение xy, умноженное на $2b$. Получаем цифры последнего столбца, которые в сумме и дадут 7345,8212, т. е. число, совпадающее как с суммарным значением, данным в табл. 546, так и с суммой квадратов по отдельным степеням свободы (табл.54г).

Конечно, и значения y^2 для отдельных степеней свободы (табл. 54г) совпадают с соответствующими значениями табл. 55, отличаясь только в некоторых случаях в третьем десятичном знаке (это следствие меньшей точности вычислений табл. 54г).

Таблица 55

Разложение редуцированной вариансы по степеням свободы

Делитель	Разности		Вариансы и ковarианса			Редуцированная варианса
	x	y	x^2	xy	y^2	
150	334	1284	743,707	2959,040	10991,840	6898,9076
10	—98	—46	960,400	450,800	211,600	104,0867
30	—290	—141	2803,333	1363,000	662,700	273,5699
10	202	135	4080,400	2727,000	1822,500	69,2570

Всего
8587,840 7399,840 13687,840 7345,8212

Весь этот процесс вычисления редуцированных средних квадратов различных категорий изменчивости оперирует с одним коэффициентом регрессии, вычисленным по данным «ошибки». Но этот коэффициент регрессии и сам заключает в себе некоторую погрешность, которая особенно сильна в том случае, если опыт не был проведен с соблюдением тщательной рандомизации. Предложен ряд методов для получения таких результатов, которые принимали бы в расчет эту погрешность, и наиболее простым,, предложенным самим Р. А. Фишером (1937в), является следующий метод. Как было указано выше, при редукции вариансы (для получения цифр табл.546) был использован один коэффициент регрессии, полученный по строке из табл.54а, соответствующей ошибке. Средний квадрат ошибки сохраняется, но для суждения о размерах среднего квадрата для вариантов производится новое вычисление следующим образом. По данным табл.54а суммируются значения Σx^2, Σxy и Σy^2, соответствующие вариантам и ошибке. Мы получаем три значения: для x^2= 25682,80, для xy=21051,80 и для y^2=28026,00.

По этим трем данным и вычисляется новая редуцированная сумма квадратов для y. Можно, конечно, вычислить коэффициент регрессии, равный $\frac{21051{,}80}{25682{,}80}$, и использовать его, как это уже мы делали раньше, для редукции. Но в данном случае, так как нам приходится делать редукцию только одной величины, целесообразно обойтись без вычисления коэффициента регрессии и просто из Σy^2 вычесть $\frac{(\sum xy)^2}{\sum x^2}$. Допустимость такого приема ясна из того, что коэффициент регрессии $b=\frac{\sum xy}{\sum x^2}$ и, следовательно, $\sum y^2-2b\sum xy+b^2\sum x^2 = y^2-\frac{2(\sum xy)^2}{\sum x^2}+\frac{(\sum xy)^2\sum x^2}{(\sum x^2)^2}=\sum y^2-\frac{(\sum xy)^2}{\sum x^2}$.

В данном случае мы из 28026,00 вычитаем $\frac{21051{,}80^2}{25682{,}80}$ или 17255,84, и получаем 10770,16 — редуцированную сумму квадратов для вариантов и ошибки вместе. Так как сумма квадратов для ошибки уже нам известна (см. табл. 37) — 3435,77, то, вычитая эту величину, получаем редуцированную сумму квадратов для вариантов, равную 7334,40. Мы видим, что эта величина практически совпадает с ранее полученной величиной для вариантов (7345,82, см. табл. 54б). Если бы мы нашли значительное расхождение результатов, то это являлось бы серьезным указанием на то, что сопутствующий признак (в данном случае — общий запас плодов на дереве) не был рандомизирован как следует; второй метод (давший меньшую надежность различий) был бы правильнее, но при значительном отклонении от рандомизации точная оценка значимости результатов является всегда сомнительной.

Этот метод может быть применен и для наиболее точной оценки по изолированным степеням свободы. Приложим, например, этот метод к первой степени свободы табл.55. Проводим опять попарное суммирование:

1 степень свободы	x^2	xy	y^2
вариант	743,71	2859,04	10991,04
ошибка	17094,96	13651,96	14338,16
Сумма	17838,67	16511,00	25329,20

Редуцированное y^2 (для первой степени свободы вариантов плюс ошибка); равно

$$25329{,}20=\frac{16511{,}00^2}{17838{,}67}=10047{,}06$$

Вычитая из 10047,06 сумму редуцированных квадратов для ошибки 3435,76, получим окончательно для 1 степени свободы — 6611,30, величина очень немногим меньшая прежде вычисленной 6898,91. Такое близкое схождение более точного и менее точного способов свидетельствует о доброкачественности опыта.

Теперь остается сравнить метод анализа ковариансы с другие методом использования сопутствующих признаков, именно применением индексов. Вместо того чтобы использовать линию регрессии урожаев по запасу плодов, можно выразить урожай в виде процента от запаса плодов и эту величину использовать для анализа вариансы. Проделаем такое вычисление с нашим материалом (табл.56 и 56а).

Проделав анализ вариансы для всех этих новых дат, получаем табл.56а.

Таблица 56

Вариант	x/y в процентах					
	блоки					
	1	2	3	4	5	сумма

1. Меритоль, опыление	59	68	58	39	67	291
2. Арсенат кальция, опыление	78	59	57	53	55	302
3. Парижская зелень, опыление	66	69	68	40	30	273
4. Парижская зелень, опрыскивание	40	27	33	29	27	156
5. Контроль	42	51	16	41	4	154
Сумма	285	274	232	202	183	1176

Таблица 56а

Категории изменчивости	Число степеней свободы	Сумма квадратов	Средний квадрат	О	*P*
Варианьы	4	4374,16	1093,54	7,241	<0,01
Блоки	4	1564,56	391,14		
Ошибка	16	2416,24	151,015		
Сумма	24	8354,96			

Результат — значительно лучший по сравнению с анализом исходных дат *y,* но уступающий результату, полученному при помощи анализа ковариансы. Разлагая сумму квадратов 4374,16 по тем же четырем степеням свободы, получаем соответственно:

1 степень свободы - 4288,0267
2 « « - 0,4000
3 « « - 53,3333
4 « « - 32,4000
Сумма 4374,1600

Только один квадрат оказывается большим среднего квадрата ошибки (первой степени свободы) и тета равна 28.39, т. е. результат лишь немногим уступает результату, полученному путем анализа ковариансы (где тета была равна 36.826).

Таким образом, в данном случае применение этого более простого способа практически дает то же, что и применение анализа ковариансы. Не трудно показать, в каких случаях удобно применение индексов и в каких нет. Когда выводим индекс $\frac{y}{x}=c$, мы получаем то же уравнение регрессии $y=cx$, отличающееся от применяемого в анализе ковариансы $y=bx+a$ отсутствием одного параметра *a*.

Поэтому анализ ковариансы является более общим методом и не может дать худших результатов, чем использование индексов но если величина *a* близка к нулю, то никакого улучшения по сравнению с индексами метод анализа ковариансы не вносит преимущества его выступают лишь тогда, когда *a* велико. Само собой разумеется, все эти рассуждения относятся к случаю прямолинейной регрессии. При наличии криволинейной зависимости в анализе ковариансы должны быть внесены соответствующие изменения, о которых уже здесь распространяться не будем.

ЗАКЛЮЧЕНИЕ

Изложенные примеры дисперсионного анализа далеко, конечно, не исчерпывают ни областей применения, ни разнообразия методов этой непрерывно развивающейся области знания. Укажу только на некоторые особенно интересные моменты,

Прежде всего, как преодолеть затруднение, возникающее при очень большом числе вариантов, все равно будет ли речь идти о качественном различии большого числа вариантов, исследуемых по методу рандомизированных блоков, или о сложном факториальном комплексе? В обоих случаях надо стремиться к тому, чтобы в пределах одного блока была достигнута возможно большая однородность участка. А при очень большом числе вариантов найти подходящие участки такой величины, чтобы охватить все разнообразие вариантов, затруднительно. Выбранные крупные участки будут обнаруживать большую гетерогенность поля, а так как внутриблоковая изменчивость входит в изменчивость, определяемую как «ошибка», то ясно, что точность опыта может быть сильно снижена.

На помощь приходят два метода: один, разработанный Иетсом (Yates) — метод неполных рандомизированных блоков (Кендалл, Стьюарт, 1973), другой, принадлежащий самому Р. А. Фишеру (1937*а, в*), — метод смешения. Этот последний приложим только к факториальной схеме, метод же неполных рандомизированных блоков одинаково приложим и при простом сопоставлении изолированных вариантов. При обоих методах число блоков увеличивается, так как каждый блок содержит лишь часть вариантов. Площадь блоков уменьшается, и гетерогенность поля снижается. Варианты же размещаются по блокам так, что сохраняется строгая уравновешенность распределения. Обработка приобретает, конечно, несколько большую сложность. О деталях методов изложено в книгах Р. А. Фишера (1937 *а,в,* 1958) статьях Йетса и в руководстве М. Кендалла и А. Стьюарта (1966, 1973, 1976) и монографии Н. А. Плохинского (1980).

Что касается областей приложения дисперсионного анализа, то они чрезвычайно разнообразны.

Плодотворному применению дисперсионного анализа мешает еще обильное количество недоразумений, связанных с преувеличенным представлением о трудностях самого метода. Конечно, его нельзя усвоить на ходу, необходимо основательно поработать, чтобы чувствовать полную уверенность при его применении, но ведь это же справедливо в отношении любого другого метода научного исследования. По сравнению с многими другими методами прикладной математики дисперсионный анализ обладает одним огромным преимуществом: лежащая в основе его теорема аддитивности варианты, несмотря на трудность ее чисто математического доказательства, чрезвычайно проста для понимания, а. главное, доступна для постоянной ее проверки. Вот эта-то возможность постоянно проверять себя, приспособляя метод к конкретным задачам, и делает возможным то, что разработка этого-» метода для решения задач новых типов может производиться и лицами, не имеющими основательной математической' подготовки. Поэтому эта ветвь математической статистики помимо своей плодотворности и более прост в своем применении, чем многие классические статистические методы. Задачей настоящего руководства. и являлась популяризация дисперсионного метода.

СПИСОК ОСНОВНЫХ БИОМЕТРИЧЕСКИХ РАБОТ А. А. ЛЮБИЩЕВА

1. О вредоносности хлебного пильщика и узловой толстоножки. — Тр. Ср.-Волжск. станции защиты растений. Самара, 1930, с. 24—37.
2. К методике учета экономического эффекта вредителей. — Тр. Всесоюз. ин-та защиты растений ВАСХНИЛ. Л., 1931, № 1, вып. 2, с. 353—505.
3. К вопросу об установлении размеров потерь, причиненных вредными насекомыми. — Защита растений, 1931, т. 8, вып. 5—6, с. 472—488.
4. Эффективность мероприятий и учет потерь. — Тр. Всесоюз. ин-та защиты растений ВАСХНИЛ. Л., 1933а, № 5, с. 123—133.
5. Роль энтомологии в продвижении пшеницы. — Тр. Всесоюз. ин-та защиты растений ВАСХНИЛ. Л., 19336, № 7, с. 16—21.
6. Основы методики учета потерь от вредителей. — Защита растений,,, 1935, т. 12, вып. 4, с. 12—20.
7. О методике количественного учета вредителей. И. Н. Степанцев, М. И. Кособуцкий, А. А. Любищев. — В кн.: Методика энтомофитопатологического учета. Ташкент, 1936, с. 7—35.
8. Критический разбор книги акад. П. Н. Константинова «Методика полевых опытов». — Вестн. с.-х. литературы, 1940, 8 5, с. 24—26.
9. К методике оценки эффективности мероприятий по борьбе с вредителями и болезнями сада. — Тр. Украинского НИИ плодоводства. Киев—Харьков, 1940, с. 3—23.
10. Об определении вредоносности методом искусственного повреждение (критический обзор). — Бот. журн. АН УССР, 1940а, т. 1, № 1, с. 159—188.
11. О построении системы мероприятий по борьбе с сельскохозяйственными вредителями. — Изв. Кирг. ФАН СССР, 1947, вып. 6, с. 121—133.
12. К методике полевого учета сельскохозяйственных вредителей и эффективности мероприятий по борьбе с ними. — Учен. зап. Ульяновского пед. ин-та,, 1955, вып. 6, с. 3—55.
13. Биометрические методы в систематике. — В сб.: Материалы II совещ. по применению математики в биологии. Л., Изд-во Ленингр. ун-та, 1958, с. 12—17.
14. Проблематика и методика количественного учета. — В сб.: Материалы II совещ. по применению математики в биологии. Л., Изд-во Ленингр. ун-та,, 1959, с. 24—26.
15. Статистические методы в энтомологии. — Тез. IV Всесоюз. энтомоло-гич. о-ва. М., Изд-во АН СССР, 1959, с. 34—36.
16. О применении биометрии в систематике. — Вестн. Ленингр. ун-та, 1959,,. .№ 9, с. 128—136.
17. Об использовании дискриминантных функций в таксономии. — Байометрикс, 1962, т. 18, № 4, с. 455—477.
18. О количественной оценке сходства. — В сб.: Применение математических методов в биологии, т. 1. Л., Изд-во Ленингр. ун-та, 1962, с. 152—160.
19. Систематика и эволюция. — В сб.: Внутривидовая изменчивость наземных позвоночных животных и микроэволюция. Свердловск, 1966, с. 45—47.
20. О некоторых новых направлениях в математической таксономии. — .Журн. общ. биол., 1966, т. 27, № 6, с. 688—696.
21. Рецензия на книгу Ч. Блисса «Статистика в биологии». — Журн. общ. биол., 1968, т. 29, № 2, с. 252—253.
22. К методике установления связи между температурой и длительностью развития. — Тр. Новосибирской станции ВИЗР. Новосибирск, 1969, с. 5—22.
23. Об ошибках в применении математики в биологии. I. Ошибки от недостатка осведомленности. — Журн. общ. биол., 1969, т. 30, № 5, с. 572—584.
24. Об ошибках в применении математики в биологии. II. Ошибки, связанные с избытком энтузиазма. — Журн. общ.^биол., 1969, т. 30, № 6, с. 715—723.
25. Рецензии на книгу Е. С. Смирнова «Таксономический анализ». 1969. — Энтомол. обозрение, 1971, т. 50, № 2, с. 493—496.
26. О приложении математической статистики к практической систематике. — В кн.: Прикладная математика в биологии. М., Изд-во Моск. ун-та, 1979, -с. 12—28.

СПИСОК ЛИТЕРАТУРЫ, ИСПОЛЬЗОВАННОЙ В ДАННОЙ МОНОГРАФИИ

1. Блисс Ч. Анализ данных полевого опыта, выраженный в процентах. — Защита растений, 1937, № 12, с. 67—77.
2. Деревицкий Н. Ф. Новейшие данные из области применения вариационной статистики. — В кн.: Иогансен В. Элементы точного учения об изменчивости и наследственности. М., 1933, с. 307—340.
3. Кендалл М., Стьюарт А. Теория распределения. М., Наука, 1966, 587 с.
4. Кендалл М., Стьюарт А. Статистические выводы и связи. (Таблицы Ф. Йетса, Стьюдента, М. Бартлетта и др.). М., Наука, 1973, 900 с.
5. Кендалл М., Стьюарт А. Многомерный статистический анализ и временные ряды. М., Наука, 1976, 736 с.
6. Кокрен У. Методы выборочного исследования. М., Статистика, 1976, 439 с.
7. Константинов П. Н. Методика полевого опыта. М., 1939, 150 с.
8. Леонтович А. В. Элементарное пособие к применению методов Г. Гаусса и К. Пирсона при оценке ошибок в статистике и биологии. Киев, 1911, 101 с.
9. Леонтович А. В. Вариационная статистика. М., 1935, 211 с.
10. Лобашев П. Г. Полевой опыт. М„ 1935, 167 с.
11. Любименко В. Н. (ред.). Физиология больного и поврежденного растения. — Тр. ВИЗР, сер. III. Л., 1933, вып. 3, 296 с. *I*
12. Мизес Р. Вероятность и статистика. М.—Л., 1930, 140 с.
13. Плохинский Н. А. Алгоритмы биометрии. Изд. 2-е. М., Изд-во Моск. ун-та, 1980, 150 с.
14. Поморский Ю. Л. Методы биометрических исследований. М.—Л., 1935, 174 с.

15. Поморский Ю. Л. Новейшие методы вариационной статистики. (Гл. 3. Статистический анализ комплексных признаков). М.—Л., 1939, 139 с.
16. Романовский В. И. О новейших методах математической статистики, применяемой в полевом опыте. — Соц. наука и техника, 1934, № 3—4, с. 75—86.
17. Романовский В. И. Применение математической статистики в опытном деле. М.—Л., 1947, 290 с.
18. Слуцкий Е. Е. Теория корреляций и элементы учения о кривых распределения. Киев, 1912, 111 с.
19. Снедекор Дж. Статистические методы в применении к исследованиям в сельском хозяйстве и биологии. М., Сельхозгиз, 1961„ 503 с.
20. Старовский В. Н. Выборочный метод. Гл. IV. — В кн.: Боярский А. Я., Старовский В. Н., Хотимский Д. И., Ястремский Б. С. Теория математической статистики. М., 1931, с. 87—128.
21. Терентьев П. В. Метод корреляционных плеяд. — Вести. Ленингр.» ун-та, 1959, т. 9, вып. 2, с. 137—144.
22. Фишер Р. А. Статистические методы для исследователей. Изд. 11-е.-М., Статистика, 1958, 268 с.
23. Чугунин Я. В. Результаты испытания пиретрума против яблоневой плодожорки. — Защита растений, 1937, № 13, с. 73—75.
24. Ястремский Б. С. Можно ли пользоваться непосредственными данными переписи? — В кн.: Избранные труды по статистике. М., 1937, с. 132—147.
25. Bliss Q. A. Calculus of Variation. N. Y., 1925.
26. Fisher R. A. Statistical Methods for Research Workers, 6th. ed. Edinburg. London, 1937a.
27. Fisher R. A. The Desing of Experiments, 2th. ed. Edinburg—London, 1937b.
28. Fisher R. A. Statistical Tables, llth. ed. Edinburg—London, 1938.
29. Fisher R. A. Contribution to Mathematical Statistics. N. Y„ 1950.
30. Gosset A. Student's Collected Papers. — In: Biometrice Office, Univer. College Press, London, 1942.
31. Hervey C. E. R„ Hartzell F. Z. Influence of planting dates of sweet corn on european corn borer infestitation. — J. of Europ. EnthomoL, 1931,. vol. 24, p. 183—188.
32. Jennings H. S., Lashley K. C. Biparental Inheritance and the-Question of Sexuality in Paramecium caudatum. — J. of experimental Zoology,. 1903, vol. 14, p. 393—466.
33. Pearson K. The Grammer of Sciences, 2th. ed. Paris—Brussel, 1938.
34. Snedecor G. W. Calculation and Interpretation of Analysis of Variance and Covariance. Jowa, 1934.

ОГЛАВЛЕНИЕ