К. А. БРАУНЛИ

СТАТИСТИЧЕСКИЕ ИССЛЕДОВАНИЯ В ПРОИЗВОДСТВЕ

INDUSTRIAL
EXPERIMENTATION

by

K. A. BROWNLEE

1947

К. А. БРАУНЛИ

СТАТИСТИЧЕСКИЕ ИССЛЕДОВАНИЯ В ПРОИЗВОДСТВЕ

ПЕРЕВОД С АНГЛИЙСКОГО
В. А. ГОВОРКОВА

ПОД РЕДАКЦИЕЙ АКАДЕМИКА
А. Н. КОЛМОГОРОВА

1949
ИЗДАТЕЛЬСТВО
ИНОСТРАННОЙ ЛИТЕРАТУРЫ
Москва

ПРЕДИСЛОВИЕ РЕДАКТОРА ПЕРЕВОДА

Небольшая книга Браунли дает хорошее представление о технике применения методов математической статистики при промышленных испытаниях в Англии. С нашей точки зрения книга эта чрезмерно рецептурна и догматична. Ценность ее по преимуществу в хорошо подобранных примерах, большинство которых вполне конкретны и реальны. На постепенно усложняющихся примерах автор дает возможность читателю, не привыкшему к языку общих формул, усвоить технику расчетов. Многие примеры удачны еще и в том отношении, что показывают пограничные случаи, когда значимость тех или иных различий действительно не очевидна с первого взгляда и тем не менее устанавливается вычислением, или, наоборот, когда какой-либо выбор кажется неискушенному читателю вполне очевидным, подсчет же дает для него лишь весьма скромный „уровень значимости“.

Наиболее полное теоретическое освещение всех приемов, рекомендуемых Браунли, содержится в книге Г. Крамера „Математические методы статистики“, русский перевод которой издан Государственным издательством иностранной литературы. Однако книга Крамера далеко не элементарна. Менее подготовленному читателю можно рекомендовать прочесть соответствующие разделы из „Элементарного курса математической статистики“ В. И. Романовского (Госпланиздат, 1938). Для того чтобы помочь читателю связать чисто рецептурное изложение Браунли с общими теоретическими курсами, к переводу приложены краткие комментарии редактора перевода.

Особенно большое внимание Браунли уделяет *дисперсионному анализу*. К сожалению, на русском языке нет достаточно подробного изложения этой важной теории, а достаточно полного *критического* ее изложения вообще не существует. Как указано в комментариях редактора перевода, автор несколько увлекается этим методом, и из его изложения можно получить преувеличенное впечатление о его силе. Тем не менее, популярное изложение приемов дисперсионного анализа с большим числом факторов и учетом их взаимодействия, данное Браунли, весьма ценно, так как именно в таких сложных случаях для предварительной ориентировки и выделения наиболее существенных факторов и их взаимодействий дисперсионный анализ особенно полезен.

А. Колмогоров

ПРЕДИСЛОВИЕ

Настоящая монография основана на инструкции, выпущенной ранее Директоратом артиллерийских заводов взрывчатых веществ для лиц, связанных с испытаниями в опытном и общезаводском масштабе по химическим производственным процессам на этих заводах.

Опыт проведения многих испытаний доказал желательность проверки результатов с помощью критериев значимости. Однако не было описания соответствующих методов, предназначенного для тех, кто должен применять критерии на практике, не обладая достаточными познаниями в их теоретических обоснованиях. Настоящая книга и представляет собой попытку дать такого рода описание.

Очевидно, что для удобного и экономичного проведения проверок с помощью критериев значимости испытания должны планироваться надлежащим образом. Мы считаем, что излагаемые методы проверок должны стать таким же стандартным инструментом заводского экспериментатора, каким химические весы служат экспериментатору в лаборатории.

При проведении заводского испытания не ставится вопрос о выборе между применением статистического расчета с приложением надлежащих критериев значимости и обычными методами; речь идет о выборе правильного метода. Даже простейшее испытание требует проверки значимости его результатов.

Нельзя не осудить строго случаи, когда консультируются со статистиком только после окончания испытания и спрашивают его: „Что можно получить из результатов?". Важно, чтобы испытание проводилось в удобной для анализа форме; этого можно достигнуть, вообще говоря, только путем проектирования испытания при консультации статистика или с должным учетом привлекаемых статистических принципов.

Настоящая монография выпускается поэтому как руководство по планированию испытаний в заводском масштабе и по объяснению их результатов; мы надеемся, что описанные здесь методы станут частью повседневной работы тех, кто проводит испытания.

Январь, 1946 г.

Р. Боуден,
директор заводов взрывчатых веществ
Министерства снабжения

ИЗ ПРЕДИСЛОВИЯ АВТОРА

Настоящая монография предназначена для лиц, ведущих исследовательскую работу, в качестве руководства по современным статистическим методам как в отношении применения критериев значимости для получения надежных выводов из данных испытаний, так и в отношении применения статистических расчетов в целях получения наибольшей точности с наименьшими затратами.

Предмет рассматривается исключительно с практической точки зрения. Теория сведена к минимуму, а математический аппарат — к простой арифметике. Каждый рассмотренный статистический метод подробно иллюстрирован практическими примерами, чтобы показать точно, в чем его суть. Некоторые более сложные методы могут показаться довольно трудно выполнимыми: к ним настоятельно рекомендуется подходить путем овладения более простыми приемами. К последним выработается тем самым привычка и доверие, и читатель увидит, что более сложные методы являются сравнительно простым развитием знакомых ему простых приемов.

Применение статистических методов требует определенного количества вычислений. В основном это — суммирование чисел и их квадратов, легко выполнимое с помощью таблиц квадратов (см., например, Б а р л о у „Таблицы квадратов и др.“) и простого арифмометра. Для проведения большого количества корреляционных расчетов желательна вычислительная машина, производящая умножение и деление, однако эти машины дороги, и при недостаточной загрузке их приобретение не оправдается. Как бы то ни было, та или иная форма механизации вычислений существенно важна.

Теоретическим основанием этой книги служили методы, разработанные в основном профессором Р. А. Фишером в его трудах: „Статистические методы для исследовательских работников“ и „Проектирование испытаний“*.

Таблицы III—VI перепечатаны из книги проф. Р. А. Фишера и д-ра Фрэнка А. Иэйтс „Статистические таблицы для исследований в биологии, сельском хозяйстве и медицине“**.

Некоторые факторы для диаграмм контроля качества перепечатаны из Британского стандарта 600 R „Диаграммы контроля качества“.

Все приведенные примеры являются результатами экспериментальных и исследовательских работ на заводах взрывчатых веществ.

---

* *R. A. Fisher*, „Statistical Methods for Research Workers“ и „The Design of Experiments“.

** *R. A. Fisher & F. A. Yates*, „Statistical Tables for Biological, Agricultural and Medical Research“.

ГЛАВА I

ВВЕДЕНИЕ

(А) ОШИБКИ ПРИ НАБЛЮДЕНИЯХ

Результаты экспериментальной исследовательской работы всегда содержат ошибки наблюдений. Ни один исследователь не будет всерьез это оспаривать. Тем более замечательно, что лишь в последнее время становится общепризнанной необходимость оценивать ошибки результатов каждого исследования для суждения об их достоверности; соответствующая операция известна в статистике под названием „проверка значимости“. Можно прямо сказать, что отсутствие этих проверок значимости является основной причиной неудовлетворительного качества многих исследований, произведенных в промышленности; их ошибки велики и притом больше, чем предполагал исследователь, который относил ряд обнаруженных эффектов за счет влияния варьированных им факторов, тогда как на самом деле они возникли исключительно из-за случайных колебаний, вызванных ошибками наблюдений. Если бы были произведены строгие проверки значимости, многие из сделанных утверждений оказались бы совершенно недостаточно обоснованными; некоторые оказались бы просто неверными. (Заметим, что из недостаточной обоснованности того или иного утверждения не вытекает, конечно, что само утверждение неверно; может оказаться, что автор был прав в своем рискованном утверждении. Однако директор предприятия, рассматривая рекомендацию своего консультанта, должен знать, основана ли она на научно доказанном факте или на субъективной догадке.)

(В) КЛАССИЧЕСКИЕ И ПРОИЗВОДСТВЕННЫЕ ИСПЫТАНИЯ

В вопросе оценки ошибки возникает одно из принципиальных различий между исследованиями в лаборатории и в производстве, т. е. в заводском масштабе. Исследователь, работающий в лаборатории, находится обычно в более благоприятном положении, будучи в состоянии осуществлять точный контроль над всеми независимыми переменными; его материалы могут быть наивысшего качества, а измерительная аппаратура — высшей точности; его ошибки обычно малы.

Напротив, экспериментатор на производстве часто не может добиться полного контроля над всеми переменными, так как масштаб испытания зачастую столь обширен, что для такого контроля понадобилась бы армия инспекторов. Достижение хорошего контроля может оказаться также и технически трудным; так, например, в хорошо оборудованной лаборатории легко поддерживать температуры в весьма узких пределах с помощью обычной термостатической аппаратуры. При испытаниях же в заводском масштабе это обычно совершенно неосуществимо. В любых лабораторных условиях достаточно легко задержать процесс; с другой стороны, задержка работы поглотительной башни высотой 15 м (в производстве серной кислоты) обошлась бы абсурдно дорого.

Кроме того, для экспериментатора на производстве важно проводить испытания так, чтобы возможно меньше нарушать нормальный ход производства. Может оказаться, что для осуществления контроля над независимыми переменными потребовалось бы столько времени и внимания к каждой партии продукции, что производительность завода заметно сократилась бы. Или же для выяснения влияния какой-нибудь независимой переменной может оказаться желательным менять эту переменную в широких пределах, чтобы получить достаточно далекие друг от друга показания; это могло бы повести к выпуску большого количества продукции, не удовлетворяющей стандарту, т. е. брака. В таком случае приходится работать в узком диапазоне изменения независимой переменной; это значит, что эффект, подлежащий наблюдению, будет намного уменьшен и может оказаться размытым погрешностями наблюдений.

Поэтому результаты испытаний в производстве по сравнению с лабораторными обычно содержат гораздо бо́льшие погрешности. При этих условиях бывает трудно решить, является ли результат данного индивидуального испытания истинным или ошибочным: возникает необходимость применять статистические критерии значимости. Но для получения значимого результата (т. е. такого, в котором наблюдаемый эффект значительно больше величины ошибки в его определении), если не применять специальных статистических приемов, требуется большое количество испытаний, а это в заводских условиях крайне нежелательно из-за дороговизны.

Наконец, быть может, уместно заметить, что, по опыту автора, даже в лабораторных условиях после объективного подсчета с помощью точных статистических методов часто оказывается, что результаты содержат гораздо бо́льшие ошибки, чем думает и утверждает экспериментатор.

Из сказанного выше следует, что экспериментатору на производстве приходится преодолевать ряд дополнительных трудностей, с которыми не сталкивается экспериментатор в лаборатории; поэтому экспериментатор на производстве нуждается в особых дополнительных приемах исследования. Эти приемы разработаны в последние годы применительно к сельскому хозяйству рядом исследователей (в частности, Р. Фишером и его коллегами на Ротэмстэдской экспериментальной станции), которым приходилось считаться с погодой и изменениями плодородности почв среди целого ряда факторов, находившихся вне их контроля.

(С) ПОВТОРЕНИЯ ИСПЫТАНИЯ

Вообще, существуют два альтернативных метода уменьшения слишком большой ошибки в оценке какого-либо эффекта. Первый метод заключается в усовершенствовании техники эксперимента, например, в применении более прецизионного термостатического оборудования, во взвешивании на лучших весах, в принятии более строгих мер предосторожности против испарения и т. п. Другой метод, который в производственных условиях часто является единственно возможным, состоит в многократном повторении эксперимента и осреднении результатов. К сожалению, этот метод не очень продуктивен, так как ошибка среднего значения обратно пропорциональна квадратному корню из числа наблюдений. Беря среднее четырех наблюдений, получаем ошибку среднего в два раза меньше ошибки одиночного наблюдения; среднее шестнадцати наблюдений имеет ошибку в четыре раза меньшую, чем ошибка одиночного наблюдения.

Эксперименты в лабораторных условиях обходятся сравнительно дешево; поэтому бывает возможным получить гораздо большее число наблюдений, чем строго необходимо для вывода заключения с заданным уровнем значимости. Благодаря этому, точной проверки значимости не требуется. Работа же в производственном масштабе обходится столь дорого и требует столь больших затрат сырья, энергии и т. п., что проведение числа наблюдений, бо́льшего чем требуется для поставленной цели — вывода того или иного заключения с заданным уровнем значимости, было бы непростительной расточительностью.

(D) ПРОЕКТИРОВАНИЕ ЭКСПЕРИМЕНТОВ. СЛУЧАЙНЫЕ БЛОКИ

До сих пор теория статистики была упомянута лишь в связи с теорией ошибок, для которой она, конечно, является основой. Однако теория статистики может оказать большую помощь и в проектировании порядка проведения экспериментов, имеющих целью получить неискаженную оценку ошибки и свести ошибку к минимуму.

Начнем с одного из простейших понятий, с так называемого „случайного блока.“

Пусть желают сравнить каких-нибудь четыре способа „обработки". Слово „обработка" имеет здесь обобщенный смысл; например, могут рассматриваться четыре рода удобрений для посева пшеницы, или четыре различных значения температуры в процессе нитрации, или четыре значения концентрации пластификатора в пластмассе и т. п. Предположим, затем, что материал поступает на обработку партиями такого размера, что возможно провести только четыре эксперимента по каждой партии. Предположим, что поступающие друг за другом партии материала могут заметно различаться по качеству. Назовем отдельные партии „блоками".

В упомянутом примере с посевом пшеницы можно рассматривать как „блок" четыре соседних опытных участка земли; они, вероятно, более сходны друг с другом по урожайности, чем другие участки, принадлежащие к другим более удаленным блокам, находящимся, возможно, в другом поле. В примере процесса нитрации понятие „блока" можно отнести к баку, наполненному хорошо размешанной кислотой, и партии материала, подлежащего нитрации, тщательно размешанного до достижения однородности. Затем предположим, что каждое наблюдение будет производиться четыре раза.

Обозначим способы обработки буквами A, B, C и D и блоки цифрами 1, 2, 3 и 4. Эксперимент можно провести, например, в порядке, приведенном в таблице 1.1.

ТАБЛИЦА 1.1.

Блок	Обработка			
1	A	A	A	A
2	B	B	B	B
3	C	C	C	C
4	D	D	D	D

Достаточно взглянуть на таблицу 1.1, чтобы убедиться, что такой порядок проведения эксперимента был бы нелепым, так как любые различия в результатах могут быть отнесены в равной степени как за счет различий между блоками, так и за счет различий между способами обработки.

Несколько лучше было бы распределить способы обработки в чисто случайном порядке. Можно заготовить 16 карточек, написав на четырех из них A, на других четырех B и т. д., положить их в шляпу и затем вынимать по одной и раскладывать по порядку, слева направо по блокам, идя сверху вниз. Результат одной из таких операций показан в таблице 1.2.

ТАБЛИЦА 1.2.

Блок	Обработка			
1	D	B	B	D
2	C	D	A	B
3	B	A	C	A
4	C	C	D	A

Если карточки были перемешаны как следует, средние четырех результатов по каждому способу обработки будут правильны со статистической точки зрения, но величина ошибки в каждом среднем будет преувеличена, поскольку она будет охватывать и различия между блоками. Поэтому эксперимент будет не так точен, как он мог бы быть.

Очевидно, будет лучше распределить способы обработки по одному на каждый блок, как показано в таблице 1.3.

ТАБЛИЦА 1.3

Блок	Обработка			
1	A	B	C	D
2	A	B	C	D
3	A	B	C	D
4	A	B	C	D

Различие между средними по четырем наблюдениям A и по четырем наблюдениям B будет теперь совершенно независимо от различий между блоками.

Более удовлетворительно, однако, было бы наметить порядок применения способов обработки по каждому блоку различным, в разбивку, а не систематическим, в порядке A, B, C, D, как указано в таблице 1.3. В самом деле, если, например, существует тенденция получения высокого результата в первом эксперименте каждого блока, то получится высокий итоговый результат по способу обработки A, что неверно.

Этого можно избежать, распределив порядок проведения экспериментов, как показано в таблице 1.4.

ТАБЛИЦА 1.4.

Блок	Обработка			
1	*D*	*B*	*A*	*C*
2	*B*	*C*	*A*	*D*
3	*C*	*B*	*A*	*D*
4	*A*	*C*	*D*	*B*

Мы пришли теперь к понятию „случайного блока". Все подлежащие сравнению способы обработки распределены по блокам, по равному числу в каждом, причем блок достаточно велик, чтобы охватить по крайней мере по одному разу каждый способ обработки. Повторения исследований осуществляются по различным блокам, порядок проведения везде сделан в разбивку. Разумеется, если объем блока таков, что в пределах отдельного блока можно сделать несколько повторений, весь эксперимент и следует провести именно так; при повторениях же в различных блоках ничего не выигрывается*.

Анализ результатов такого эксперимента, с проверкой уровня значимости результатов, рассмотрен в главе VII, (E).

(E) ЛАТИНСКИЙ КВАДРАТ

Рассмотренный в предыдущем разделе гипотетический эксперимент спроектирован достаточно строго в том отношении, что он не приведет к ложным заключениям; однако в некоторых случаях его можно еще усовершенствовать.

* См., впрочем, гл. XII, (E).

В качестве частного примера предположим, что в некотором многоступенном процессе партии материала по одной тонне делятся на четыре подпартии, которые обрабатываются приблизительно в одно и то же время в четырех стабилизаторах; требуется сравнить эффекты четырех различных способов обработки на этой последней ступени. При четырех способах обработки и четырех подпартиях в каждой партии применим прием „случайных блоков", с распределением четырех способов обработки по четырем стабилизаторам в случайном порядке. Можно, однако, сделать так, что каждый из способов обработки придется на каждый стабилизатор один и только один раз, так что средние результаты будут совершенно независимы от возможных различий между стабилизаторами.

Такой порядок показан в таблице 1.5, где буквенные обозначения относятся к способам обработки.

ТАБЛИЦА 1.5

Партия	Стабилизатор 1	Стабилизатор 2	Стабилизатор 3	Стабилизатор 4
1	*A*	*B*	*C*	*D*
2	*C*	*D*	*A*	*B*
3	*B*	*A*	*D*	*C*
4	*D*	*C*	*B*	*A*

Из таблицы 1.5 видно, что четыре подпартии, относящиеся к одной и той же партии, подвергаются всем четырем способам обработки. В то же время каждый из четырех способов обработки и каждая из партий приходятся на каждый из стабилизаторов только по одному разу. Возможные ошибки из-за различий между партиями и между стабилизаторами, таким образом, исключаются из средних и из величины ошибки; этим достигается получение неискаженных оценок и минимальной величины ошибки.

Подсчет результатов латинского квадрата рассматривается в главе XII, (G).

(F) СБАЛАНСИРОВАННЫЕ НЕПОЛНЫЕ БЛОКИ

Объем обычного случайного блока должен быть достаточен для охвата одного полного комплекта сравниваемых способов обработки. Однако может случиться, что число сравниваемых способов обработки больше размера блока.

Допустим, что сырье поступает партиями такой величины, что по каждой партии можно сделать 4 эксперимента; между партиями ожидается значительная дисперсия. Если требуется сравнить 2, 3 или 4 способа обработки, эта дисперсия между партиями не вызовет беспокойства у исследователя, поскольку каждый комплект из 2, 3 или 4 способов обработки можно будет провести по одной и той же партии. Если же требуется сравнить 5 или более способов обработки, то один или более из них придется провести по другой партии, а не по той, где проводилось испытание первых четырех способов; результат сравнения будет обесценен из-за влияния дисперсии между партиями.

Обычный метод обхода этого затруднения заключается в том, что один из способов обработки выбирается в качестве „стандартного“ и включается в каждую партию. Все способы обработки сравниваются затем непосредственно со стандартным; остальные сравнения способов между собой проводятся либо прямо — для способов, попавших в один и тот же блок, либо косвенным путем, например, из сравнения способа M со стандартным в его блоке, и из сравнения способа N со стандартным в блоке, в который входит N.

Хотя этот порядок испытаний и пригоден, но ни в какой мере не является достаточно удовлетворительным, так как здесь возникают три различные ошибки в трех различных типах сравнений. Кроме того, он и не наиболее продуктивен.

Рассмотрим альтернативный порядок проведения испытаний, показанный в таблице 1.6; здесь сравниваются пять способов обработки с четырехкратными повторениями, по пяти партиям.

ТАБЛИЦА 1.6

Партия №				
1	2	3	4	5
A B C D	A B C E	A B D E	A C D E	B C D E

Можно сделать ряд замечаний относительно этого порядка проведения исследований:

(a) Каждый способ обработки попадает один и только один раз в четыре партии из пяти.

(b) Любая пара способов обработки, например A и B или C и E, попадается в трех из пяти партий. Например, непосредственное сравнение между B и D возможно по партиям 1, 3 и 5.

(c) Непосредственные сравнения данной пары способов обработки невозможны по другим двум партиям. Так, партия 2 не содержит D, а партия 4 не содержит B. Тем не менее, можно сделать весьма удовлетворительное сравнение, поскольку A, C и E содержатся в партиях 2 и 4; можно использовать среднее этих трех как „стандарт“. Заметим, что использование в качестве стандарта средней по трем способам обработки более удовлетворительно, чем использование одиночной величины.

Такого рода порядок проведения испытаний известен под названием „сбалансированных неполных блоков“. Блоки сбалансированы, поскольку каждый способ обработки применяется одинаковое число раз; блоки являются неполными, поскольку ни один из них не охватывает всех способов обработки. Недостатком является то обстоятельство, что такая сбалансировка не всегда возможна. Вообще, при фиксированном числе подлежащих сравнению способов обработки и при фиксированном числе экспериментов по каждой партии (блоку), число необходимых повторений каждого способа обработки оказывается точно определенным. В этом основной недостаток такого порядка: число требующихся повторений может быть больше, чем предполагается необходимым для получения достаточной точности.

С этой точки зрения одни проекты порядка проведения испытаний оказываются лучше других. Вообще, при заданном числе способов обработки, требующееся для получения сбалансирования число повторений будет тем меньше, чем бо́льше объем блока. Так, для сравнения 21 способа обработки при объеме блока 3, потребуется 10 повторений. Работая с блоками объема 4, можно получить таблицу для 25 способов обработки при 8 повторениях (дублируя 4 наиболее важных способа в целях получения итога 25). Блоки объема 5 являются весьма удовлетворительными, так как при этом требуется только 5 повторений. Любопытно, что блоки объемом 6 и 7 менее удовлетворительны, поскольку потребуется соответственно 8 и 10 повторений. Точные требования в отношении числа повторений в каждом частном случае определяются комбинаторными свойствами чисел.

Расчеты, связанные с изложенным порядком проведения исследований, рассмотрены в главе XIII.

(G) РЕШЕТЧАТЫЕ КВАДРАТЫ

Наиболее существенный недостаток метода сбалансированных неполных блоков заключается в том, что необходимое для достижения сбалансирования число повторений становится тем больше, чем больше число способов обработки. В общем случае p^2 способов обработки потребуют $p+1$ повторений; например, 36 способов обработки потребуют 7 повторений; такое число повторений может не оправдываться практическими надобностями.

В подобных случаях может оказаться наиболее подходящим прием, называемый „решетчатым квадратом".

Предположим, что требуется сравнить 25 способов обработки. Выпишем их в форме квадрата так, как показано в таблице 1.7.

ТАБЛИЦА 1.7

11	12	13	14	15
21	22	23	24	25
31	32	33	34	35
41	42	43	44	45
51	52	53	54	55

Эти 25 способов можно подвергнуть испытанию блоками, по 5 в каждом; сперва рассматривают в качестве блоков строки, а затем столбцы. Одно повторение каждого комплекта испытаний даст в итоге 4 повторения по каждому способу обработки. Заметим, что эти блоки не являются сбалансированными.

Например, способ 11 попадает в те же блоки, что и 12, 13, 14 и 15, с одной стороны, и 21, 31, 41 и 51, с другой стороны; но способ 11 никогда не попадает в блоки, содержащие прочие 16 способов обработки. Удовлетворительные сравнения, тем не менее, вполне возможны путем использования промежуточных сравнений; при надлежащей алгебраической обработке этот прием проведения испытаний является весьма продуктивным *.

Основным достоинством решетчатого квадрата является то, что число требуемых повторений не зависит от числа способов

* Этот прием здесь не рассматривается детально, так как автор не встречался в производственных испытаниях со случаями, где требовались сравнения столь больших количеств способов обработки. Краткое описание приведено лишь потому, что этот прием является простейшим в ряде других аналогичных приемов. Более подробное описание со ссылками на литературные источники опубликовано Йэйтсом (F. Yates, *Empire Journal of Experimental Agriculture*, VIII, 223 (1940)). Описания и методы расчета ряда приемов даны Гульденом (C. H. Goulden, Methods of Statistical Analysis, 1939).

обработки. Так, например, для испытания 49 способов обработки потребуется только 4 повторения, в то время как метод сбалансированных неполных блоков потребовал бы 8 повторений.

(H) ПРИРОДА „БЛОКОВ"

В предыдущих разделах нами рассматривались приемы обхода затруднений, возникающих из-за ограниченности объема блоков. Понятие „блока" является чрезвычайно общим. Приведем ряд примеров:

(a) В сельскохозяйственных исследованиях блок составляется из участков земли, непосредственно прилегающих друг к другу, поскольку соседние участки более сходны по урожайности, чем участки, значительно удаленные друг от друга.

(b) В экспериментах по откорму животных было обнаружено, что животные одного и того же приплода дают более сходные показатели результатов заданного режима питания, нежели животные разных приплодов; в опытах такого рода целый приплод может рассматриваться как блок.

(c) В экспериментах по очистке продуктов ферментации каждый процесс ферментации может лишь в небольшой степени отличаться от других и поэтому может рассматриваться как отдельный блок.

(d) В любых экспериментах, где индивидуальная идиосинкразия оператора может повлиять на результаты, причем общее число подлежащих сравнению способов обработки связано с работой нескольких операторов, ясно, что каждый оператор представляет собою блок.

(e) Могут рассматриваться как блоки многие другие возможные источники неоднородности, лежащие в основе экспериментов. Например, при сравнении способов обработки, связанной с использованием термостата или инкубатора, очевидно, предпочтительно поместить по возможности все объекты, испытываемые по всем способам обработки, в один и тот же термостат или инкубатор, и притом в одно и то же время. Ошибки из-за несовершенства оборудования будут тогда сведены к минимуму. Однако может оказаться, что термостат недостаточно велик, чтобы сравнить одновременно все способы обработки. В таких случаях либо пользуются несколькими термостатами, каждый из которых будет представлять собой блок, либо используют один и тот же термостат в течение нескольких следующих друг за другом промежутков времени; в последнем случае блоком будет служить время.

(f) В экспериментах, где сравнивается ряд способов обработки, но материал, подвергаемый обработке, может слегка изменяться с течением времени, причем лишь часть способов обработки может быть испытана в течение одного дня, каждый день будет представлять собой блок. Примерами могут служить дифференциальная утечка из газгольдера, содержащего смесь газов, или какие-либо закваски (в частности, в производстве пенициллина), которые могут быть заметно нестабильными.

(g) Во многих случаях нет четких признаков, вроде описанных выше; тем не менее, разбивка эксперимента на блоки имеет известные преимущества. Вообще, эксперименты, близкие друг к другу во времени и в пространстве, будут более сходны между собой, чем отдаленные. Температуре воздуха, атмосферному давлению, влажности, отрегулированности станков, настроению и поведению рабочих и многим другим менее ощутимым факторам свойственно меньше изменяться в течение короткого промежутка времени, нежели в более длинные периоды.

Поэтому, если эксперимент поддается разбивке на блоки, эффективная величина ошибки будет заметно уменьшена.

(I) ИСПЫТАНИЯ СО МНОГИМИ ФАКТОРАМИ

Эксперименты, рассматривавшиеся до сих пор, содержали одно „независимое переменное“ или один „фактор“. Во многих экспериментальных исследованиях, однако, встречается сравнительно большое число независимых переменных, подлежащих изучению; для этого требуются специальные методы.

Классическим идеалом испытания считается возможность иметь все переменные, кроме одной, постоянными. Часто не представляют себе, что такое положение может оказаться далеко не идеальным, так как для получения полного представления об эффектах изменения данной переменной следует допускать также изменения всех остальных переменных по всем их диапазонам. Если же эти последние фиксировать, то выбор их фиксированных значений в конечном счете неизбежно окажется случайным. В то время как изменение какого-нибудь фактора A от его нормального значения A_1 до некоторого другого значения A_2 произведет определенное изменение в качестве изделия в случае, когда фактор B имеет значение B_1, то же изменение фактора A при другом значении B_2 фактора B может вызвать совсем иное изменение в качестве изделия. „Факторное“ планирование испытания предназначено для определения такого рода эффектов; при этом оно дает максимум производительности, т. е. при заданной затрате труда — максимальный запас нужных сведений о системе, подвергнутой испытанию.

Техника „факторных“ испытаний была разработана в сельскохозяйственной науке в двадцатых и тридцатых годах главным образом Р. Фишером и его коллегами на Ротэмстэдской экспериментальной станции.

Подход Фишера к испытаниям отличается от классической идеи испытаний „каждый раз по одной переменной“ по двум основным направлениям. Во-первых, он подчеркивает бóльшую важность получения точной оценки остаточной дисперсии, чем ее уменьшения. Точное определение остаточной дисперсии необходимо для применения точного критерия значимости.

Во-вторых, Фишер подчеркивает преимущества, получаемые при включении в одно и то же испытание как можно бóльшего количества факторов, влияние которых надлежит определить. Преимуществами такого рода являются:

(а) Гораздо бóльшая производительность испытания; оценка такого рода влияний с заданной степенью точности может быть получена из гораздо меньшего общего числа наблюдений; если испытания проводятся в заводском масштабе, это может весьма значительно сократить их стоимость.

(b) Получение информации о взаимодействии факторов, т. е. о том, как влияние одного фактора подвержено воздействию других; испытания поэтому дадут более широкий базис для заключений. В качестве иллюстрации представим себе, что один экспериментатор производит сравнение двух катализаторов для производства серной кислоты; исследуемой зависимой переменной является производительность процесса. Экспериментатор будет поддерживать постоянными все прочие факторы, как-то: температуру впуска газа и скорость подачи. Допустим, что в результате он придет к заключению, что катализатор A лучше, чем B. В то же время другой экспериментатор на основании опытов с изменением температуры впуска газа придет к выводу, что температура 420° при работе с катализатором B является оптимальной; если эксперименты по сравнению катализаторов A и B проводились при температуре 450°, то их результаты с точки зрения второго экспериментатора неверны. Первый же экспериментатор может заявить, что работа его коллеги производилась с катализатором B, и, следовательно, заключение об оптимальной температуре неправильно. Аналогично, третья группа испытаний с переменной скоростью подачи может обнаружить, что эти скорости в первых двух группах испытаний были выбраны неправильно и выведенные заключения не имеют силы. Однако испытания с подачей могли производиться при неправильно выбранной температуре и не на лучшем катализаторе; значит, и эти результаты окажутся неверными. Конечно, вполне возможно, что температура в 420° является наилучшей, независимо от того, какой бы катализатор ни применялся и какова бы ни была скорость подачи; возможно также, что и катализатор A является лучшим при любой температуре и скорости подачи, и что наилучшее значение скорости подачи остается одним и тем же при любом катализаторе и любой температуре.

Однако без новой проверки такие заключения были бы совершенно неоправданы. Рассмотренный тип влияния фактора, когда оно зависит от величины какого-нибудь другого фактора, известен под названием „взаимодействие“.

(J) ИСПЫТАНИЕ С ТРЕМЯ ФАКТОРАМИ

(I) *Классический прием.* Рассмотрим гипотетическое испытание, в котором требуется исследовать влияние трех независимых переменных P, Q и R (которые могут быть значениями температуры, давления, скорости течения, концентраций и т. п.) на зависимую переменную x (которая может быть количеством выпускаемого продукта, его чистотой и т. д.).

Предположим, что для начала достаточно провести исследование только при двух значениях каждой независимой переменной; пусть их нормальные значения равны P_1, Q_1 и R_1 и требуется найти эффект их изменения до значений P_2, Q_2 и R_2.

Как следует провести это испытание в классическом стиле?

Прежде всего проведем испытание при значениях независимых переменных P_1, Q_1 и R_1. Получим значение x, которое обозначим через $(P_1Q_1R_1)_x$.

Для получения изменения x при изменении значения P от P_1 до P_2 [что обозначим символом $(P_1-P_2)_x$] производим другое испытание при значениях P_2, Q_1, R_1 и получим новое значение x; обозначим его через $(P_2Q_1R_1)_x$. Очевидно, что $(P_1-P_2)_x=(P_1Q_1R_1)_x-(P_2Q_1R_1)_x$.

Аналогично, в результате испытаний при P_1, Q_2, R_1 и P_1, Q_1, R_2 получим:

$$(Q_1-Q_2)_x=(P_1Q_1R_1)_x-(P_1Q_2R_1)_x,$$
$$(R_1-R_2)_x=(P_1Q_1R_1)_x-(P_1Q_1R_2)_x.$$

Весьма важно, чтобы каждое из этих испытаний было произведено не менее двух раз, поскольку иначе было бы совершенно невозможно сделать какую-либо оценку погрешности испытаний. Без такой оценки нельзя сказать, является ли разница в значениях $(P_1Q_1R_1)_x$ и $(P_2Q_1R_1)_x$ реальной или же она получилась из-за ошибок в измерениях, в выборе образцов испытания и т. п. Таким образом, каждое из четырех испытаний необходимо произвести дважды; общее количество испытаний будет восемь. Влияние изменения каждой независимой переменной будет определяться двумя наблюдениями; это значит, что мы будем сравнивать среднее значение одной пары наблюдений со средним значением другой пары.

Испытания этого рода не дают никакого представления о возможном взаимодействии факторов. Так, например, $(R_1-R_2)_x$ при $P=P_1$ может значительно отличаться от $(R_1-R_2)_x$ при $P=P_2$; однако на

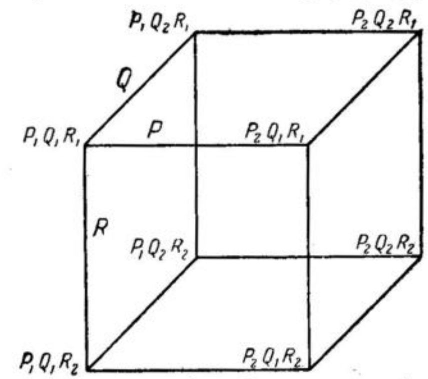
основании проведенных испытаний подобный эффект обнаружить нельзя.

(II) *Факторный прием испытаний.* Испытания должны проводиться при всех возможных комбинациях переменных P, Q и R, а именно при $(P_1Q_1R_1)$, $(P_2Q_1R_1)$, $(P_1Q_2R_1)$, $(P_1Q_1R_2)$ $(P_2Q_2R_1)$, $(P_2Q_1R_2)$, $(P_1Q_2R_2)$ и $(P_2Q_2R_2)$. Общее количество наблюдений — восемь, то же, что и при классическом приеме.

Указанные комбинации легче воспринять, если представить их себе как координаты восьми вершин куба, ребра которого параллельны осям P, Q и R (см. фиг. 1).

Оценку трех основных эффектов можно представить следующим образом. Величину $(P_1-P_2)_x$ мы получаем, сравнивая среднюю в плоскости P_1 [т. е. среднее значение $(P_1Q_1R_1)_x$, $(P_1Q_2R_1)_x$, $(P_1Q_2R_2)_x$ и $(P_1Q_1R_2)_x$] со средней в плоскости P_2 [т. е. со средним значением $(P_2Q_1R_1)_x$, $(P_2Q_2R_1)_x$, $(P_2Q_2R_2)_x$ и $(P_2Q_1R_2)_x$]. Хотя другие две переменные при этом и изменяются, однако их значения соответствуют друг другу в двух сравниваемых „плоскостях“, так что в первом приближении можно считать, что вносимые этим искажения взаимно уничтожаются.

Подобным же образом определяются и другие основные эффекты $(Q_1-Q_2)_x$ и $(R_1-R_2)_x$.



Фиг. 1

Первым преимуществом факторного приема является то, что оценки эффектов получаются каждый раз в виде разности средних для *четырех* наблюдений, а не двух, как это имело место в классическом приеме. Таким образом, при том же числе испытаний удваивается степень точности. Уже это обстоятельство может иметь величайшую важность в отношении экономичности, если испытания проводятся в заводском масштабе при значительных затратах и помехах для хода производства.

Вторым преимуществом факторного приема является возможность оценки взаимодействия между основными эффектами.

Обратимся опять к нашему кубу. Для выяснения взаимодействия между P и R вычисляем среднюю передней и задней плоскостей, т. е. находим среднюю от Q. Результат записываем в таблицу, в которой верхний индекс $\overline{Q}$ указывает, что значения осреднены по Q:

$$(P_1R_1)_x^{\overline{Q}}, \qquad (P_2R_1)_x^{\overline{Q}},$$
$$(P_1R_2)_x^{\overline{Q}}, \qquad (P_2R_2)_x^{\overline{Q}},$$

Верхняя строка дает значение $(P_1-P_2)_x$ при $R=R_1$, а нижняя строка — значение $(P_1-P_2)_x$ при $R=R_2$. Поскольку каждое число в таблице является средней из двух испытаний, получается действительная оценка ошибки; мы можем определить, значимо ли $(P_1-P_2)_x^{R_1}$ отличается по величине от $(P_1-P_2)_x^{R_2}$. Верхний индекс R_1 или R_2 указывает, какому значению R соответствует данное число. Если величины $(P_1-P_2)_x^{R_1}$ и $(P_1-P_2)_x^{R_2}$ значимо отличаются друг от друга, следует заключить, что между P и R существует взаимодействие. Аналогичным путем можно определить наличие других взаимодействий.

Вычисления для экспериментов с тремя факторами приведены в гл. XI, (D)

(K) ФАКТОРНЫЕ ИСПЫТАНИЯ ВЫСШЕГО ПОРЯДКА

Мы рассмотрели эксперимент с тремя факторами, причем каждый фактор имел два уровня; это может быть представлено символом $2\times2\times2$ или 2^3. Расширение эксперимента до четырех факторов легко себе представить, хотя его геометрическое изображение и невозможно. Изображая четвертый фактор символом S, сделанные ранее 8 испытаний рассматриваем как произведенные при низшем значении S_1 фактора S; для введения четвертого фактора S проводим подобные 8 испытаний при высшем значении S_2 фактора S. Оценка основного эффекта фактора S дается теперь в виде разности между средними результатов первой (при S_1) и второй (при S_2) групп испытаний. Аналогичным путем получаются оценки трех других факторов. Теперь возможно наличие шести взаимодействий первого порядка, соответствующих попарным комбинациям факторов — PQ, QR, PR, PS, QS и RS. Для оценки каждого из них берутся средние по двум факторам, не участвующим в данном взаимодействии; например, для оценки взаимодействия PQ берутся средние по R и S.

Очевидно, что введение четвертого фактора существенно улучшает результаты эксперимента. Основные эффекты получаются как результаты сравнения средних по восьми испытаниям, а не по четырем, как было ранее; взаимодействия первого порядка получаются как результаты сравнения средних по четырем наблюдениям, а не по двум, как было ранее.

Кроме того, теперь можно проверить наличие так называемых взаимодействий второго порядка. В качестве примера рассмотрим взаимодействие PQ; возможно, что его величина зависит от значения R, принимая различные значения при $R=R_1$ и при $R=R_2$. Такого рода эффект называется взаимодействием второго порядка и обозначается символом PQR или $P\times Q\times R$. Было найдено, что в действительности такие взаимодействия симметричны относительно соответствующих трех факторов. Взаимодействие PQR можно рассматривать как взаимодействие PQ с R или QR с P или же RP с Q. При четырех факторах возможно наличие четырех взаимодействий второго порядка: PQR, QRS, RSP и PQS. Эксперимент с четырьмя факторами позволит оценить наличие всех этих возможных эффектов [порядок расчета приведен в главе XI, (E)].

Заметим, что испытания с четырьмя факторами еще более продуктивны по сравнению с классическим приемом проведения испытаний, чем испытания с тремя факторами, которые сами по себе давали удвоение точности оценки основных эффектов.

Испытание с четырьмя факторами требует $2^4=16$ опытов; его основные эффекты оцениваются путем сравнения средних результатов одной группы в 8 опытов со средними результатов другой группы в 8 опытов. Эквивалентное, наиболее близкое, испытание по классическому методу заключается в осуществлении пяти комбинаций опытов, с тремя повторениями каждый:

$$\begin{matrix} P_1 & Q_1 & R_1 & S_1 \\ P_2 & Q_1 & R_1 & S_1 \\ P_1 & Q_2 & R_1 & S_1 \\ P_1 & Q_1 & R_2 & S_1 \\ P_1 & Q_1 & R_1 & S_2. \end{matrix}$$

В итоге получается 15 опытов. Основные эффекты при этом будут оцениваться, исходя из средних, только по трем опытам, что значительно уступает по качеству результатам факторного проведения испытаний. Последнее, к тому же, позволяет проверить наличие шести возможных взаимодействий первого порядка и четырех возможных взаимодействий второго порядка.

Теперь ясно, в чем заключается огромное превосходство факторного метода испытаний. В случае классического проведения испытаний, 12 опытов, относящихся к четырем последним комбинациям, были использованы лишь по одному разу каждый, для получения информации по одному только эффекту. Так, три повторения комбинации $P_2\ Q_1\ R_1\ S_1$ дают информацию только относительно основного эффекта фактора P и совершенно ничего не добавляют к нашим познаниям относительно трех других факторов или различных возможных взаимодействий. Аналогичные рассуждения применимы и к комбинациям $P_1\ Q_2\ R_1\ S_1$, $P_1\ Q_1\ R_2\ S_1$ и $P_1\ Q_1\ R_1\ S_2$. Лишь одна комбинация $P_1\ Q_1\ R_1\ S_1$ была использована несколько раз. В случае же факторного испытания результат каждого опыта используется в нескольких случаях, притом каждый раз различным путем.

Распространение факторного метода испытаний на более чем 4 фактора или на факторы с более чем двумя уровнями очевидно.

Наконец, заметим, что большие по объему факторные испытания требуют значительного числа отдельных опытов; например, испытание 2^4 или $2\times2\times2\times2$ требует $2^4=16$ опытов. Это число может превышать объем блока. Поэтому разработаны специальные методы обхода таких затруднений посредством комбинирования 16 опытов в два блока по 8 опытов каждый, или в 4 блока по 4 опыта каждый. Эти методы рассмотрены в главе XIV.

ГЛАВА II

ОСНОВНЫЕ ПОНЯТИЯ СТАТИСТИКИ

(A) ТЕРМИНОЛОГИЯ

Подобно другим наукам, статистика имеет свою терминологию и свои обозначения. Они могут смутить новичка, но владение ими важно для правильного применения статистических методов.

Мы прилагаем статистические методы к группам чисел, полученных в результате измерений, относящихся к какому-либо свойству продукта или процесса производства — количества, чистоты, или таких качеств, как срок службы (в днях или тоннах пропускаемого материала) какой-либо части оборудования без повреждений. В другом случае, на каком-либо станке изготовляются детали, и мы измеряем их размеры, вес или твердость. Эту переменную мы будем называть „зависимой“; нас часто интересует влияние на нее условий, при которых она получается. Эти условия обычно могут быть представлены в виде ряда „независимых переменных“, которые мы назовем „факторами“. Так, в химических процессах условия могут определяться значениями температуры, давления, продолжительности реакции, пропорционального состава реагентов. Заданное значение независимой переменной называется „уровнем“. Данное наблюдение над зависимой переменной называется „индивидуальным испытанием“ или „индивидуумом“.

Важным является понятие „совокупности“, или, как ее часто называют, „генеральной совокупности“, представляющей собой, вообще говоря, большое количество индивидуумов, взятых из данного источника. Часто желают определить среднее значение какого-либо свойства всей партии „индивидуумов“ на основании наблюдений над небольшим количеством их; эта часть, выделенная из всей партии, называется „выборкой“.

Черточка над буквенным обозначением указывает на то, что подразумевается среднее значение: $\bar{x}$ является средним значением рассматриваемой совокупности значений x. Знак $\sum$ обозначает суммирование ряда членов.

Несколько трудным является понятие „числа степеней свободы“. Оно тесно связано с количеством наблюдений; вообще говоря, это — разность между количеством наблюдений и количеством наложенных на систему ограничений. Точное определение, однако, дать затруднительно; поэтому подробное пояснение этого понятия будет даваться для каждого частного применения. Заметим, что сходные понятия знакомы тем, кто имеет дело с „правилом фаз“ в области физической химии и занимается квантовой механикой молекул.

(B) ВЕРОЯТНОСТЬ

Вероятность выражается числами от 0 до 1; значение, равное нулю, указывает, что событие наверняка не произойдет (или что гипотеза неверна); значение, равное единице, указывает, что событие наверное произойдет (или что гипотеза верна). Значение 0,5 указывает, что наступление или ненаступление события одинаково вероятны (или что гипотеза в одинаковой мере может считаться верной или неверной).

В практике величина вероятности часто выражается в процентах; например, вероятность 0,05 считается 5%-ной вероятностью.

(C) БЕСКОНЕЧНЫЕ СОВОКУПНОСТИ. КРИТЕРИИ ЗНАЧИМОСТИ *

Согласно Фишеру („Статистические методы для исследовательских работников“), „основной для всей статистической работы является идея бесконечной совокупности с заданными в ней распределениями частот одной или более характеристик“. Ограниченные сведения, полученные из наблюдений над какой-нибудь функцией или индивидуумами данного типа, „позволяют нам составить некоторое представление о бесконечной гипотетической совокупности, из которой взята наша выборка, и, таким образом, о вероятной природе будущих выборок, к которым должно будет применяться наше заключение. Если вторая выборка опровергает это ожидание, мы заключаем, говоря на языке статистики, что она взята из иной совокупности и что обработка, которой подверглась вторая выборка, обнаружила существенное отличие этой второй совокупности от первой. Испытания этого рода могут быть названы „критериями значимости“; с помощью таких критериев можно установить, является ли вторая выборка значимо отличающейся от первой или нет“.

Применяя самые обычные критерии такого рода, мы допускаем, что средние („критерий t“ Стьюдента) или дисперсия („критерий z“

* См. комментарии редактора в конце книги. (*Прим. ред.*)

Фишера) — одни и те же для обеих выборок; затем, приняв эту гипотезу, мы вычисляем вероятность появления результатов, которыми мы фактически располагаем.

Если эта вероятность превышает 0,05 (5%), нет серьезных оснований сомневаться в гипотезе. Если же эта вероятность равна 0,05, то данный результат получался бы в среднем только в одном случае из двадцати, если верна наша гипотеза, т. е. если действительно обе выборки извлечены из совокупностей с одинаковыми средними, дисперсиями или распределениями. Поэтому обычно принимают, что в таком случае наша гипотеза неверна.

Меньшие, чем 0,05, значения вероятности служат, конечно, еще более определенным указанием того, что гипотеза неправильна. Таким образом, малое значение вероятности соответствует высокой степени значимости различия двух выборок.

(D) УРОВНИ ЗНАЧИМОСТИ

Какой уровень вероятности следует считать значимым — зависит от характера испытаний. Для многих целей принят 5%-ный уровень; однако следует отчетливо представлять себе, что при этом в одном из каждых 20 случаев мы будем утверждать наличие эффекта, не существующего в действительности. Если такой процент ошибки считается слишком большим, следует принять более высокий, например, 1%-ный, уровень значимости. В этом случае, однако, возникает опасение, что мы будем игнорировать существующие в действительности эффекты. Вероятно, наилучшим практическим компромиссом будет признание пригодными результатов с уровнем значимости в 5% или немного ниже. В окончательных выводах и заключениях следовало бы все же предпочесть более высокий уровень значимости. Его можно достичь путем увеличения числа наблюдений, если эффект существует в действительности.

Во всем последующем изложении мы будем употреблять слово „зна́чимый" в смысле „статистически-значимый". Слово „значимый" будет означать, что рассмотренные данные достаточны для утверждения, что соответствующий эффект существует с таким-то уровнем значимости; это не означает, что эффект будет непременно иметь большую важность с практической точки зрения, хотя, разумеется, это может быть и так. Верно и обратное: какой-либо из эффектов может оказаться при наличных данных незначимым, но причиной этого может быть просто недостаточность данных для обнаружения значимости; сам же по себе эффект может иметь величайшую практическую важность. В общем, наша цель заключается в оценке каждого эффекта с установлением так называемых „доверительных пределов", в которых

находится величина эффекта (с заданным уровнем вероятности); в некоторых случаях решение такого рода задачи отнюдь не просто.

Например, если величина эффекта должна достигать 15% для того, чтобы иметь практическую важность, а наши данные дают оценку эффекта в 5% ± 2%, то эффект является статистически-значимым, но не имеющим практической важности. С другой стороны, если наши данные привели к оценке эффекта в 5% ± 10%, то эффект является статистически-незначимым по нашим данным, но может все же быть важным практически.

(E) ЧИСЛОВЫЕ РАСЧЕТЫ

Применение статистических методов требует, вообще говоря определенного количества числовых расчетов. Значительная часть последних сводится к суммированию чисел и их квадратов; это легко выполнимо с помощью таблиц квадратов (таблиц Барлоу, например) и простого арифмометра. Для выполнения большого количества расчетов по корреляции необходима счетная машина, на которой можно производить умножение и деление. Существует ряд типов счетных машин; можно сказать с уверенностью, что чем совершеннее машина, тем точнее и быстрее можно выполнить работу. Электрическая машина с автоматическим делением и умножением очень желательна.

(F) МЕРЫ РАССЕИВАНИЯ

В статистической работе приходится часто иметь дело с колебаниями какого-либо свойства системы, например, в выпуске продукции или в качестве какого-либо химического продукта.

Мерой рассеивания или разброса в группе наблюдений могут служить различные числа:

(a) „широта", т. е. разность между наибольшим и наименьшим значениями; обозначается символом w;

(b) „среднее отклонение", т. е. среднее отклонений значений от их среднего;

(c) „дисперсия", т. е. частное от деления суммы квадратов отклонений значений от их средней величины на общее количество наблюдений без единицы; дисперсия обозначается символом σ^2.

Так, для нижеследующих результатов наблюдения

$$9,\ 11,\ 14,\ 11,\ 13,\ 12,\ 10,\ 13,\ 12,\ 15:$$

(a) Широта равна $15-9=6$.

(b) Среднее значение равно $(9+11+14+11+13+12+10+13+ +12+15):10=12$; отклонения равны: $12-9=3$; $12-11=1$; $14-12=2$; $12-11=1$; $13-12=1$; $12-12=0$; $12-10=2$; $13-12=1$; $12-12=0$; $15-12=3$.

Среднее отклонение равно $(3+1+2+1+1+0+2+1+ +0+3):10=1{,}4$.

(c) Дисперсия σ^2 определена так:

$$\sigma^2=\frac{\Sigma(x-\bar{x})^2}{n-1},$$

где $\bar{x}$ — среднее значение x-ов, n — число наблюдений, $(x-\bar{x})^2$ — квадраты отклонений; значит,

$$\sigma^2=\frac{3^2+1^2+2^2+1^2+1^2+0+2^2+1^2+0+3^2}{10-1}=\frac{30}{9}=3{,}33.$$

Из этих мер рассеивания широта в силу ее арифметической простоты часто применяется в работе по контролированию процессов (диаграмма контроля качества, см. главу VIII) при выборках небольшого объема ($n\leqslant 10$). Очевидно, что понятие широты не использует всей информации, даваемой выборкой, поскольку промежуточные значения не входят в ее определение.

Среднее отклонение имеет также достоинство арифметической простоты, но его свойства таковы, что математическое использование становится затруднительным; поэтому оценка по среднему отклонению должна считаться устаревшей.

Дисперсия является, вообще говоря, наиболее предпочтительной мерой рассеивания, так как ее свойства хорошо известны и на ней основаны распространенные критерии значимости. Дисперсия обладает также ценным свойством аддитивности, т. е. если процесс подвержен дисперсиям $\sigma_1^2, \sigma_2^2, \ldots$, вызываемым несколькими независимыми причинами, то полная дисперсия σ_T^2 равна сумме частичных дисперсий: $\sigma_T^2=\sigma_1^2+\sigma_2^2+\ldots$. Последнее соотношение особенно важно, когда полную дисперсию системы разлагают на компоненты, определяющиеся различными частями процесса. Это рассматривается подробно в главе VII. Следующим преимуществом дисперсии как меры рассеивания является ее максимальная „эффективность“; это означает, что с помощью ее из данных извлекается больше информации о рассеивании, чем с помощью других мер.

Квадратный корень из величины дисперсии называется „стандартным отклонением“ * и обозначается символом σ.

* Величину σ называют также „средним квадратическим отклонением“ или „стандартом“. (*Прим. перев.*)

Соотношение между широтой и стандартным отклонением рассмотрено в главе VIII, (E).

(G) ВЫЧИСЛЕНИЕ ДИСПЕРСИИ

При вычислении дисперсии часто бывает утомительно подсчитывать среднее $\bar{x}$, находить отклонения, возводить их в квадраты и суммировать. Мы используем поэтому алгебраически равноценную формулу:

$$\sigma^2=\frac{\Sigma x^2-\frac{(\Sigma x)^2}{n}}{n-1}.$$

Для приведенного выше примера имеем:

$$\sigma^2=\left[9^2+11^2+14^2+\ldots+15^2-\frac{(9+11+\ldots+15)^2}{10}\right]:(10-1)= =\left(1470-\frac{120^2}{10}\right):9=3{,}33.$$

Для облегчения арифметических вычислений нередко бывает целесообразно принять какое-либо произвольно выбранное из данных наблюдений значение за нуль и соответственно пересчитать все значения относительно нового начала отсчета. Так, для того же примера примем за начало отсчета число 9; данные получат вид: 0, 2, 5, 2, 4, 3, 1, 4, 3, 6.

Дисперсия этих значений будет равна:

$$\sigma^2=\left[(0^2+2^2+5^2+\ldots+6^2)-\frac{(0+2+5+\ldots+6)^2}{10}\right]:(10-1)= =\left(120-\frac{30^2}{10}\right):9=3{,}33.$$

Как видим, перенос начала отсчета (нуля) не отражается на величине дисперсии. Так это и должно быть, поскольку рассеивание измеряется относительно *среднего значения*.

Если бы в нашем примере все числа были соответственно больше на 100 единиц, то сумма квадратов была бы порядка сотен тысяч и вычисления были бы довольно утомительны; их можно было бы намного облегчить, либо сместив нуль на 100 единиц и перейдя таким образом к первому из изложенных выше расчетов, в котором сумма квадратов равнялась 1470, либо сместив нуль на 109 единиц, как это было сделано в последнем расчете, в котором сумма квадратов оказалась равной 120.

Смещение начала отсчета, т. е. нуля, является в значительной мере делом вкуса. Оно требует внимания, но зато, будучи проделано, намного облегчает последующие вычисления.

Получаются два преимущества, которые мы отметим в применении к тому же разобранному примеру: во-первых, мы приходим к цифрам, меньшим 100, которые при известном навыке нетрудно возводить в квадрат в уме; во-вторых, квадраты двузначных чисел имеют не более четырех знаков, что облегчает перенос их на вычислительные приборы.

(H) ОПРЕДЕЛЕНИЕ ДИСПЕРСИИ

Формально дисперсия определяется не как

$$\frac{1}{n-1}\sum(x-\bar{x})^2,$$

как было дано выше, а

$$\frac{1}{n}\sum(x-\bar{x})^2;$$

именно в последнем виде дисперсия рассматривается в ряде статистических работ. В практических применениях, однако, мы будем пользоваться выражением со знаменателем $n-1$.

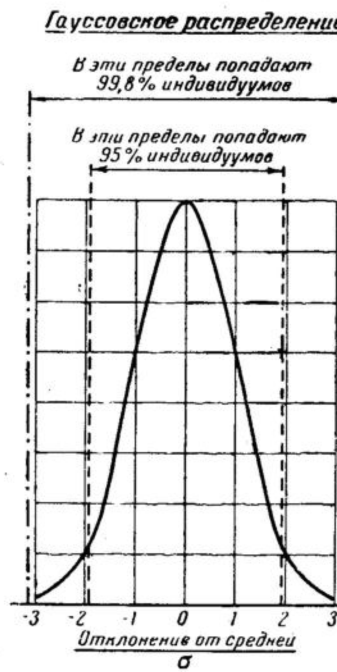
Причина этого расхождения определений следующая. При вычислении $\sum(\bar{x}-x)^2$ мы должны были бы подсчитать сумму квадратов отклонений от истинного среднего совокупности, из которой взята выборка. Мы же имеем только среднее значение выборки, которое, вообще говоря, не совпадает с истинным средним всей совокупности. Сумма квадратов отклонений от среднего выборки, как правило, будет меньше, чем сумма квадратов отклонений от истинного среднего; убедимся в этом на примере, предположив, что имеются два наблюдения 1 и 3 и что истинное среднее совокупности равно 1. Сумма квадратов отклонений от истинного среднего равна $(1-1)^2+(3-1)^2=4$; исходя из истинного среднего, при определении дисперсии мы взяли бы в качестве знаменателя 2 и получили бы дисперсию $4:2=2$. Сумма же квадратов отклонений от среднего значения выборки равна $(2-1)^2+(3-2)^2=2$. Если взять в качестве знаменателя 2, значение дисперсии получится 1; если же поставить в знаменатель $n-1=1$, получим значение дисперсии 2. Мы видим, что внесение в знаменатель выражения дисперсии $n-1$ вместо n способствует компенсации неточности, внесенной применением среднего значения выборки вместо истинного среднего.

Величина $n-1$ носит название числа степеней свободы дисперсии; она используется при сравнении одной дисперсии с другой с помощью критериев значимости (глава IV).

(I) РАСПРЕДЕЛЕНИЯ

Общеизвестно, что нельзя изготовить вполне тождественные изделия. Последовательно производимые индивидуальные предметы слегка отличаются друг от друга размерами, качеством и т. д. Аналогично, при выполнении повторных определений какой-либо физической величины, например выпуска продукции, получаемые значения не точно одинаковы; расхождения будут не только вследствие возможных изменений в процессе, но также и из-за ошибок в измерениях.

Можно доказать теоретически, что при очень большом числе очень малых ошибок, оказывающих независимое друг от друга влияние на результат, распределение получит форму, изображенную на фиг. 2. Это — так называемое гауссовское распределение.



Фиг. 2

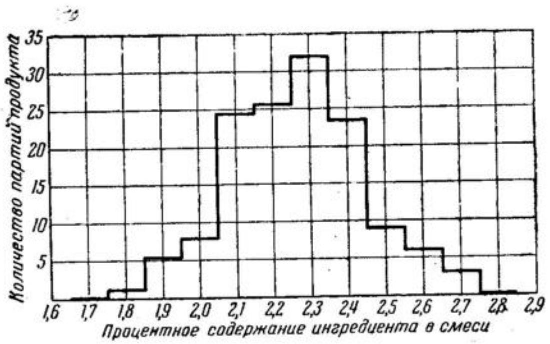
Гауссовское распределение имеет следующие ценные свойства:

(a) Оно полностью определяется двумя статистическими параметрами: 1) средним значением, фиксирующим положение распределения вдоль оси x, и 2) стандартным отклонением, фиксирующим его ширину или разброс.

(b) Выведено соотношение с расчетными таблицами между отклонениями от среднего, выраженными как кратное от σ, и частотой, с которой эти отклонения происходят. Так, в частности, 95% индивидуумов должны находиться между $-1{,}96\sigma$ и $+1{,}96\sigma$ по обе стороны от среднего, а 99,8% индивидуумов — между $-3{,}09\sigma$ и $+3{,}09\sigma$. Подобным же образом, если мы получили отклонение от среднего, равное d, то, вычислив величину $\frac{d}{\sigma}$ — так называемое „стьюдентово t", — с помощью специальных таблиц сможем определить, насколько часто должно встречаться такое отклонение.

На фиг. 3 показано типичное распределение из практики химической промышленности. Оно относится к количеству одного из ингредиентов в смеси из четырех составляющих. Величина ординаты в каком-либо интервале абсцисс указывает количество партий продукта, имевших содержание ингредиента в заданном интервале.

Видно, что, хотя опытное распределение и не соответствует точно теоретической кривой Гаусса, все же в общих чертах они схожи. Следует иметь в виду, что приведен результат всего лишь 138 наблюдений: ожидаемое число партий в любом заданном интервале довольно мало, поэтому должны обнаружиться случайные отклонения от ожидаемого значения. Если бы общее количество наблюдений было 1000, мы могли бы рассчитывать на более точное подобие опытного распределения гауссовскому; при 10 000 наблюдений сходство было бы еще более близким.



Фиг. 3

Многие из проверок с помощью критериев значимости основаны на допущении, что переменная распределена по гауссовскому закону.

В действительности опытные распределения не имеют точно гауссовской формы, однако почти неизменно приближаются к ней достаточно точно. К счастью, отступление от гауссовского закона должно быть очень значительным, для того чтобы критерии значимости оказались заметно неточными. Соответствующие расхождения для любого распределения, имеющего максимум около середины и спадающего по обе стороны от него до нуля, совершенно незначительны.

(J) РАСПРЕДЕЛЕНИЯ ГРУППИРОВАННЫХ ЧАСТОТ

Если количество наблюдений, стандартное отклонение которых подлежит определению, велико, например 40 или более, то необходимые выкладки могут быть сокращены путем применения так называемого „группирования частот".

В качестве иллюстрации используем данные из предыдущего параграфа (см. табл. 2.1).

ТАБЛИЦА 2.1

X	f	x	fx	fx^2
1,75 — 1,84	1	−4	−4	16
1,85 — 1,94	6	−3	−18	54
1,95 — 2,04	8	−2	−16	32
2,05 — 2,14	24	−1	−24	24
			−62	
2,15 — 2,24	26	0	0	0
2,25 — 2,34	32	1	32	32
2,35 — 2,44	23	2	46	92
2,45 — 2,54	9	3	27	81
2,55 — 2,64	6	4	24	96
2,65 — 2,74	3	5	15	75
	138		+144 −62	502
			+82	

Подразделяем значения переменной (в данном случае процентный состав ингредиента) примерно на 10—20 одинаковых групп (столбец X). Подсчитываем количество наблюдений по каждой из этих групп (столбец f). Этот столбец соответствует вертикальной оси на фиг. 2. Выбираем новое начало отсчета где-либо около середины интервала изменения X и наносим шкалу в новом масштабе (столбец x), единица которого составляет $^1/_{10}$ старой единицы. Дальнейшие арифметические операции ясны. В единицах x получаем $\sigma_x = 1{,}82$. В старых единицах X $\sigma_X = 0{,}182$. Поскольку $\bar{x} = 0{,}59$, среднее значение находится на 0,59 единиц x или на 0,06 единиц X выше нового начала отсчета. Последнее было взято при $X = 2{,}20$, следовательно, среднее равно $2{,}20 + 0{,}06 = 2{,}26$ единиц X.

Вычисления:

$$\sigma_x^2 = \frac{502 - \frac{82^2}{138}}{138 - 1} = \frac{502 - 48{,}72}{137} = 3{,}31;$$

$$\sigma_x = 1{,}82;$$

$$\sigma_X = 0{,}182;$$

$$\bar{x} = \frac{82}{138} = 0{,}59;$$

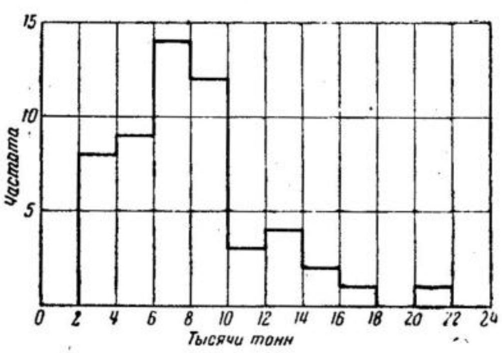
$$\bar{X} = 2{,}20 + \frac{0{,}59}{10} = 2{,}26.$$

При рассмотрении свойств гауссовского распределения было указано, что 95% общего числа наблюдений будут находиться в пределах $\pm 1{,}96\sigma$ от среднего. Здесь эти пределы равны $2{,}26 \pm \pm 1{,}96 \cdot 0{,}182$, т. е. соответственно 1,91 и 2,61. Рассматривая таблицу, видим, что два крайних интервала содержат числа наблюдений 1 и 3, т. е. в сумме 4 из 138, или $2{,}9\%$. Однако интервалы 1,85—1,94 и 2,55—2,64, вероятно, содержат наблюдения, лежащие вне пределов 1,91 и 2,61, так что действительное совпадение с гауссовским распределением будет достаточным.

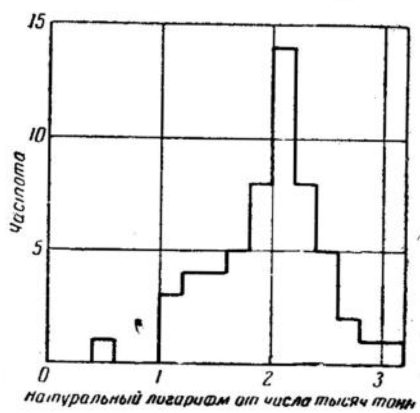
Аналогично, следует ожидать очень небольшого (порядка 2 из 1000) количества наблюдений, лежащих вне пределов $2{,}26 \pm \pm 3{,}09 \cdot 0{,}182$, т. е. соответственно 1,70 и 2,82. В таблице мы не находим ни одного наблюдения, выходящего за эти пределы, что хорошо согласуется с прогнозом, основанным на гауссовском распределении.

(К) ЛОГАРИФМИЧЕСКИ-НОРМАЛЬНЫЕ РАСПРЕДЕЛЕНИЯ

Для некоторых типов данных полезно рассматривать вместо переменной ее логарифм. Иногда наблюдаются переменные, которые подвержены тем большему рассеиванию, чем больше их значения. Например, в опытах по коррозии наблюдается, что при сильной коррозии цифры-дубликаты начинают более резко



Фиг. 4



Фиг. 5

расходиться друг с другом. В подобных случаях следует ожидать, что кривая распределения будет асимметричной, с длинным „хвостом" в сторону увеличения абсциссы. Однако отклонение, пропорциональное среднему нелогарифмированных данных, становится постоянным в логарифмическом масштабе. На фиг. 4 показано распределение количества обработанного материала в тысячах

тонн на производственную единицу кислотного завода до выхода единицы из строя.

На фиг. 5 то же распределение приведено в логарифмическом масштабе. Один индивидуум слева позволяет заподозрить наличие „хвоста" на низких абсциссах, но рассмотрение графика в целом показывает, что логарифмическое распределение едва ли значительно отличается от нормального.

В случае, когда мы имеем переменную с явно логарифмически-нормальным распределением и желаем применить какой-нибудь критерий значимости, например сравнивая два средних, целесообразно, пожалуй, работать с логарифмами переменной. Пример этого приведен в главе XII, (К).

ГЛАВА III

ЗНАЧИМОСТЬ СРЕДНИХ ЗНАЧЕНИЙ

(А) ЗНАЧИМОСТЬ ОДИНОЧНОГО СРЕДНЕГО

Пусть имеется совокупность, распределенная со стандартным отклонением σ от ее среднего $\overline{X}$. Пусть делаются выборки объемом в n индивидуумов. Первая выборка будет иметь среднее $\overline{x}_1$, вторая — среднее $\overline{x}_2$, не обязательно одинаковое с $\overline{x}_1$, и т. д. Ценным свойством этих средних является то, что они имеют распределение относительно среднего всей совокупности $\overline{X}$ со стандартным отклонением $\frac{\sigma}{\sqrt{n}}$.

Выбираем наудачу какой-нибудь индивидуум; если его величина отличается от среднего совокупности $\overline{X}$ более, чем на $1{,}96\sigma$, заключаем, что данный индивидуум вряд ли принадлежит к рассматриваемой совокупности. Аналогично, если в случае выборки в n индивидуумов обнаружено, что ее среднее отличается более чем на $1{,}96 \frac{\sigma}{\sqrt{n}}$ от средней совокупности $\overline{X}$, то следует заключить, что вряд ли эта выборка взята из рассматриваемой совокупности.

Отношение отклонения среднего $\overline{x}$ данной выборки, состоящей из n индивидуумов, от истинного значения среднего $\overline{X}$ всей совокупности к стандартному отклонению $\frac{\sigma}{\sqrt{n}}$ известно под названием „стьюдентова t":

$$t = \frac{\overline{x} - \overline{X}}{\sigma / \sqrt{n}}.$$

Рассмотрим характерный пример. Два исследователя A и B анализировали серию химических составов в отношении процентного содержания одного из ингредиентов и получили результаты, приведенные в табл. 3.1.

ТАБЛИЦА 3.1

Исследователи	1	2	3	4	5	6	7	8	9	10
A	7	9	8	10	8	11	9	8	9	8
B	11	7	10	10	9	10	10	9	11	11
$A-B=x$	-4	2	-2	0	-1	1	-1	-1	-2	-3

Если в приемах исследования не было постоянных различий, то следует ожидать, что разность между значениями, полученными обоими исследователями по каждому составу, должна в среднем равняться нулю.

В последней строке таблицы выписаны значения этой разности; требуется установить, значимо ли отличается их среднее от нуля.

Среднее значение $\overline{x}$ для величины $A-B$ есть $-1{,}10$, а предполагаемое значение $\overline{X}$ равно нулю.

Вычислим стандартное отклонение разностей $A-B$.

$$\sigma=\sqrt{\frac{(-4)^2+2^2+\ldots+(-3)^2-\frac{(-4+2+\ldots-3)^2}{10}}{10-1}}=1{,}79.$$

Стандартное отклонение среднего $\overline{x}$ равно $\frac{\sigma}{\sqrt{n}}=\frac{1{,}79}{\sqrt{10}}=0{,}565.$

Величина t равна $\frac{1{,}10}{0{,}565}=1{,}95.$

Значение, принимавшееся нами для σ, является, конечно, только оценкой истинной величины σ генеральной совокупности. Будучи основано только на десяти индивидуумах, оно не так точно, как было бы в случае сотни или тысячи индивидуумов. Табл. I в приложении, дающая величины значимых t, вычислена с учетом этого соображения. Из таблицы видно, что при бесконечном числе степеней свободы (здесь число степеней свободы на единицу меньше количества наблюдений) величина t, равная 1,96, должна встретиться всего один раз из двадцати ($P=0{,}05$), если испытываемая гипотеза правильна. В данном примере количество наблюдений 10, так что число степеней свободы равно 9. Для получения 5%-ного уровня значимости требуется иметь величину $t=2{,}26$. Полученное же нами значение равно 1,95; оно меньше значения, соответствующего 5%-ному уровню значимости и больше значения 1,83 для 10%-ного уровня. Значимость полученного результата находится поэтому где-то между 5% и 10%, т. е. наша гипотеза об отсутствии различия в работе обоих исследователей достаточно хорошо согласуется с данными. Однако, если значение t было хотя бы немного больше, следовало бы уже заключить, что гипотеза неверна и что почти наверняка имеется различие между исследователями.

(В) ДОВЕРИТЕЛЬНЫЕ ПРЕДЕЛЫ ДЛЯ ОДИНОЧНОГО СРЕДНЕГО

Пусть мы показали, что отличие среднего величины $A-B$ от нуля незначимо. Это могло произойти либо потому, что исследуемая величина не отличается от нуля, либо из-за недостаточности данных для правильной оценки. Выясним теперь, которая из этих альтернатив имеет место.

Мы хотим определить пределы, в которых мы можем с достаточной уверенностью полагать находящейся истинную величину среднего, основываясь на данных наших наблюдений.

Прежде всего решим, что считать „достаточной уверенностью"? Можно полагать результат удовлетворительным, если мы окажемся правы 19 раз из 20 (что эквивалентно „95%-ной уверенности").

Находим по табл. I в приложении значение t для 9 степеней свободы и уровня значимости 5% (95% шансов оказаться правым эквивалентно 5% шансов сделать ошибку). Это значение t равно 2,26. Тогда пределы $\pm L$ по обе стороны от выборочного среднего равны:

$$\pm L=\pm t\frac{\sigma}{\sqrt{n}}=\pm 2{,}26\cdot 0{,}565=\pm 1{,}28.$$

Здесь σ — стандартное отклонение совокупности, а n — число индивидуумов (объем) выборки. Таким образом, можно сказать с достаточной уверенностью (в пределах избранного уровня вероятности), что истинное значение среднего находится между $-1{,}10+1{,}28$ и $-1{,}10-1{,}28$, т. е. между 0,18 и $-2{,}38$. Эти пределы называются 95%-ными доверительными пределами.

Если число случаев 19 из 20 для нас недостаточно, мы можем потребовать, например, 999 из 1000 ($99{,}9\%$); по таблице I значение t для 9 степеней свободы и уровня 0,001 равно 4,78. Пределы тогда равны $-1{,}10\pm 4{,}78\cdot 0{,}565$, т. е. $-3{,}80$ и 1,60.

Рассмотрим действительную широту диапазонов между пределами (при уровне 95%, $99{,}9\%$ или любом другом). Эти пределы фиксируют степень точности, с которой определено среднее. Если диапазон узок, среднее определено с большой точностью, если широк — точность ниже. На этом основании решают, получился ли результат незначимым в силу действительного равенства среднего с ожидаемым значением или из-за недостаточности данных. В данном случае 95% шансов за то, что среднее лежит где-то между 0,18 и —2,38; такое определение представляется довольно неточным *.

Бо́льшая точность значения среднего может быть получена путем увеличения количества наблюдений n, поскольку стандартное отклонение среднего (или стандартная ошибка) равно $\frac{\sigma}{\sqrt{n}}$. Удваивая количество наблюдений, мы суживаем диапазон между пределами приблизительно в $\sqrt{2}$ раз, т. е. до 0,707 прежней величины. Взяв учетверенное количество наблюдений, получим диапазон, суженный вдвое.

(C) СРАВНЕНИЕ ДВУХ СРЕДНИХ

Пусть имеются две выборки — две группы чисел и требуется установить, является ли отличие их средних значимым, или же можно считать, что обе выборки извлечены из одной и той же совокупности.

Пусть:

$$x_1,\ x_2, \ldots,\ x_{n_1} \text{ и } x'_1,\ x'_2, \ldots,\ x'_{n_2}$$

эти выборки. Метод обработки зависит от их объема. Выборки можно считать большими, если общее количество индивидуумов $n_1 + n_2$ больше 30.

(I) *Малые выборки.* Вычисляем следующие величины:

$$\bar{x} = \frac{\Sigma x}{n_1}, \quad \bar{x}' = \frac{\Sigma x'}{n_2};$$

$$\sigma^2 = \frac{\Sigma x'^2 - \frac{(\Sigma x')^2}{n_2} + \Sigma x^2 - \frac{(\Sigma x)^2}{n_1}}{n_1 + n_2 - 2};$$

$$t = \frac{\bar{x} - \bar{x}'}{\sigma} \sqrt{\frac{n_1 \cdot n_2}{n_1 + n_2}}.$$

Обращаемся к таблице значений t, считая число степеней свободы равным $n_1 + n_2 - 2$. Если найденное значение t превосходит табличное для 5%-ного уровня значимости, с этой степенью уверенности можно считать, что наши выборки извлечены из различных совокупностей.

* Более точно следовало бы сказать: если мы утверждаем, что среднее лежит между 0,18 и —2,38, то вероятность правильности нашего утверждения равна 0,95.

Рассмотрим пример. Приведенные ниже данные представляют собой количество кислоты в сотнях тонн, получавшееся от двух установок A и B до выхода их из строя. Есть ли основания полагать, что одна установка лучше другой?

A	71, 67, 33, 79, 42
B	73, 80

Вычисляем
$$\bar{x}_A = \frac{71 + 67 + 33 + 79 + 42}{5} = 58{,}4;$$

$$\bar{x}_B = \frac{73 + 80}{2} = 76{,}5;$$

$$\sigma^2 = \frac{(71^2 + 67^2 + 33^2 + 79^2 + 42^2) - \frac{292^2}{5} + (73^2 + 80^2) - \frac{153^2}{2}}{5 + 2 - 2} = 319{,}2;$$

$$\sigma = 17{,}9;$$

$$t = \frac{76{,}5 - 58{,}4}{17{,}9} \cdot \sqrt{\frac{5 \cdot 2}{5 + 2}} = 1{,}21 \text{ при } 5 + 2 - 2 = 5 \text{ степенях}$$

свободы.

По таблице значений t находим, что при 5 степенях свободы t должно быть не ниже 2,57, чтобы считаться значимым при уровне в 5%. Таким образом, нет оснований считать совокупности различными.

Если обратиться к более подробным таблицам t, чем та, которая приведена в приложении, то мы увидим, что получившееся значение соответствует уровню значимости примерно в 25%. Иначе говоря, в условиях отсутствия разницы между двумя установками при повторных выборках мы получали бы столь же большую разность между ними, как и теперь, один раз в каждых четырех случаях. Получившийся же результат является лишь очень слабым указанием на возможное превосходство установки B.

Мы показали, что разность наших средних не отличается значимо от нуля. Далее можно установить доверительные границы, в пределах которых разность будет находиться с избранной вероятностью. Поступаем как и прежде.

Стандартная ошибка разности средних двух выборок, обозначаемая через σ_m, равна:

$$\sigma_m = \sigma \sqrt{\frac{n_1 + n_2}{n_1 n_2}} = 17{,}9 \sqrt{\frac{5 + 2}{5 \cdot 2}} = 14{,}9.$$

Кроме того, мы имеем из таблицы величину t для $n_1 + n_2 - 2 = 5$ степеней свободы и при данном уровне значимости. Выбирая величину последнего в 5% (95% шансов правильности предсказания, что истинная разница средних двух выборок находится между положенными границами), получим для 5 степеней свободы значение $t=2{,}57$.

Тогда вероятная величина максимального отклонения (при данном уровне значимости) составляет:

$$L=t\sigma_m=2{,}57\cdot 14{,}9=38{,}3.$$

Полученное же среднее значение разности было 18,1. Истинное значение в 95% случаев будет лежать в пределах, отличающихся от 18,1 на $\pm 38{,}3$, т. е. между $-20{,}2$ и 56,4. Отсюда видно, насколько неудовлетворительно сравнение, которое можно было сделать по приведенным данным. Может существовать достаточно заметная разница между двумя установками, но наличные данные недостаточны для получения ответа.

(II) *Большие выборки.* Для больших выборок применим тот же прием, что и для малых, с тем, однако, упрощением, что в этом случае нет необходимости непременно брать число $n-1$ вместо n. Вычисляем средние двух групп:

$$\bar{x}=\frac{\sum x}{n_1}, \quad \bar{x}'=\frac{\sum x'}{n_2}.$$

Далее вычисляем стандартные отклонения σ_1 и σ_2 обеих выборок. Стандартная ошибка разности средних обеих групп $\bar{x}-\bar{x}'$ выражается равенством:

$$\sigma_D^2=\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}.$$

Как и прежде, находим величину t,

$$t=\frac{\bar{x}-\bar{x}'}{\sigma_D},$$

и обращаемся к таблице величин t, беря нижнюю строку, отвечающую случаю бесконечного числа степеней свободы.

Так, если

$$\bar{x}=100, \quad n_1=40, \quad \sigma_1^2=120,$$

$$\bar{x}'=110, \quad n_2=30, \quad \sigma_2^2=60,$$

то

$$\sigma_D=\sqrt{\frac{120}{40}+\frac{60}{30}}=\sqrt{5}=2{,}236,$$

$$t=\frac{110-100}{2{,}236}=4{,}47.$$

Это — намного выше значения, соответствующего уровню значимости в $0{,}1\%$; таким образом, можно с уверенностью считать, что оба средних действительно отличаются друг от друга.

Для установления доверительных пределов $\pm L$, в которых истинное значение разности двух средних находится, например, в 95% случаев, берем значение t для бесконечного числа степеней свободы при заданном уровне значимости. Пределы эти равны $(\bar{x}-\bar{x}')\pm t\sigma_D$, в данном случае: $10\pm 1{,}96\cdot 2{,}24$, т. е. 5,6 и 14,4.

(D) ВЫВОДЫ

Вообще, в любом отчете об экспериментальной работе, когда в дополнение к полученным значениям требуется определить среднее, должны быть приведены его стандартное отклонение и количество наблюдений. Это позволит оценить уровень значимости результата.

Важным предварительным условием при расчете стандартного отклонения является то, что индивидуумы должны быть независимыми. Так, если сделали 10 партий какого-либо материала и взяли из каждой партии две выборки, нельзя рассчитывать стандартное отклонение при допущении, что сделано 20 наблюдений. Ясно, что наблюдения-дубликаты по каждой партии не являются независимыми друг от друга. Методы статистической обработки в таких условиях рассмотрены ниже, в главе XII, (J).

ГЛАВА IV

СРАВНЕНИЕ ДИСПЕРСИЙ

(A) СРАВНЕНИЕ ДВУХ ДИСПЕРСИЙ

Для сравнения средних двух групп чисел применяется „критерий t" Стьюдента. В некоторых случаях, однако, требуется сравнить рассеивание двух групп чисел; это является особой проблемой.

Критерий значимости для проблем подобного типа разработан Фишером. (Собственно говоря, Фишер брал натуральный логарифм отношения квадратных корней из дисперсий, обозначаемый им через z; мы же здесь будем рассматривать просто отношение дисперсий, обозначаемое через F.)

Предположим для примера, что имеются два альтернативных способа проведения некоторого технико-химического анализа. Один или оба способа могут иметь систематическую „пристрастность"; но это несущественно, так как нам требуется лишь метод сравнения.

В табл. 4.1 даны результаты шести анализов одной и той же выборки по способу A и семи анализов по способу B.

ТАБЛИЦА 4.1

Способ A . . .		95,6	94,9	96,2	95,1	95,8	96,3
Способ B . . .	93,3	92,1	94,7	90,1	95,6	90,0	94,7

Средние значения по способам A и B равны соответственно 95,65 и 92,93. Если бы желали проверить значимость этого различия, следовало бы применить „критерий t" Стьюдента. Здесь же мы сравним дисперсии

$$\sigma_A^2 = \frac{(5{,}6^2+4{,}9^2+\ldots+6{,}3^2)-\frac{(5{,}6+4{,}9+\ldots+6{,}3)^2}{6}}{6-1} = 0{,}324,$$

$$\sigma_B^2 = \frac{(3{,}3^2+2{,}1^2+\ldots+4{,}7^2)-\frac{(3{,}3+2{,}1+\ldots+4{,}7)^2}{7}}{7-1} = 5{,}14.$$

Затем вычисляем отношение большей дисперсии к меньшей, обозначаемое через F. С этим отношением связаны два значения числа степеней свободы — n_1 для большей дисперсии и n_2 — для меньшей.

В данном примере отношение дисперсий равно $\frac{5{,}14}{0{,}324} = 15{,}9$ при $n_1 = 6$ и $n_2 = 5$. В приложении помещены таблицы отношений дисперсий для ряда значений n_1 и n_2 и для уровней значимости в 20, 5,1 и 0,1%. Если полученное значение отношения дисперсий больше приведенного в таблице для соответствующих степеней свободы, это означает, что полученный результат более значим, чем уровень значимости в таблице.

Так, для чисел степеней свободы $n_1 = 6$, $n_2 = 5$ и для уровней значимости в 20, 5, 1 и 0,1% значения отношений дисперсий соответственно равны 2,2, 5, 10,7 и 28,8. Полученное нами отношение дисперсий равно 15,9; его значимость находится поэтому между уровнями 1 и 0,1% (см. текст ниже).

Следовательно, допуская, что наша гипотеза верна, мы вряд ли получили бы такой результат. В соответствии с этим весьма вероятно, что гипотеза неверна. Наша гипотеза состояла в том, что обе группы цифр были получены из одной и той же совокупности, т. е. что изменчивости в обоих способах были одними и теми же. Поскольку гипотеза оказалась неверной, можно

утверждать, что изменчивости обоих способов значительно отличаются друг от друга.

Здесь возникает несколько деликатный вопрос. Таблицы дисперсий составлены для проверки того, является ли дисперсия A большей, чем B, причем заранее известно, что значение A выведено определенным образом, который отличает его от B. В данном же случае мы испытываем, является ли дисперсия A большей, чем дисперсия A'; до окончания подсчетов мы не знаем, которой дисперсии приписать значение A и которой A'. Уже потом мы обозначаем через A бóльшую по величине дисперсию. Это обстоятельство вызывает изменение уровней значимости с 20, 5, 1 и 0,1% соответственно до 40, 10, 2 и 0,2%.

(B) ОСРЕДНЕНИЕ НЕСКОЛЬКИХ ДИСПЕРСИЙ

В развитие предыдущего примера предположим, что имелся ряд определений изменчивости каждого способа по ряду различных выборок. Предположим, что вторая выборка, подвергнутая анализу по методу B, дала результаты 93,1, 91,2, 92,6. Ее дисперсия тогда равна

$$\frac{3{,}1^2+1{,}2^2+2{,}6^2-\frac{(3{,}1+1{,}2+2{,}6)^2}{3}}{3-1} = 0{,}97$$

(взяв новое начало отсчета, мы вычли из всех данных 90).

Среднее взвешенное значение двух дисперсий, взятое с учетом числа степеней свободы, равно

$$\frac{2\cdot 0{,}97+6\cdot 5{,}14}{2+6}.$$

Это среднее значение совпадает с тем, которое получается путем суммирования квадратов и чисел степеней свободы для обеих выборок и равно

$$\frac{30{,}82+1{,}94}{6+2}.$$

Это — правильный метод вычисления средней дисперсии. Неверно: (a) брать среднее значение дисперсий без учета числа степеней свободы, (b) брать среднее значение стандартных отклонений.

Если вместо дисперсий имеются стандартные отклонения, следует их возвести в квадрат и взять среднее, как указано.

Число степеней свободы средней дисперсии, правильно рассчитанной, принимается равным сумме чисел степеней свободы дисперсий, из которых она вычислена. В рассматриваемом примере число степеней свободы средней дисперсии составляет $6+2=8$.

Можно также указать, что для подсчета средней дисперсии двух выборок обе дисперсии должны быть рассчитаны отдельно и затем объединены. Неверно было бы объединить все $7+3=10$ наблюдений и вычислить их дисперсию, потому что полученное таким путем значение дисперсии включало бы, помимо требуемой дисперсии повторных измерений по одной и той же выборке, также дисперсию между выборками, что не относится к делу.

(С) СРАВНЕНИЕ НЕСКОЛЬКИХ ДИСПЕРСИЙ

В Фишеровском критерии для отношения дисперсий сравниваются только две величины дисперсии и не более. Предположим, что имеется группа в десять станков, на которых обрабатываются партии какого-то изделия и определяется некоторое качество x в каждой партии; пусть имеются основания полагать, что на некоторых из станков обработка производится более правильно, чем на других, т. е., другими словами, величина рассеивания в их продукции меньше.

Критерий значимости в проблемах такого типа известен под названием критерия Бартлета.

Во втором столбце нижеприведенной таблицы представлены числовые данные, характеризующие свойство x по шести взятым подряд партиям изделий, обработанных на станках за номерами от 1 до 10. В целях упрощения вычислений данные отсчитаны от искусственной нулевой точки; в действительности же каждое число в этом столбце на 2200 единиц больше.

Станок	Данные	Σx	Σx^2	$\frac{(\Sigma x)^2}{6}$	$\Sigma (x-\overline{x})^2$	σ_i^2	$\ln (\sigma_i^2)$
1	28, 14, 15, 2, 9, 2	70	1294	816,7	477,3	95,46	4,5587
2	11, 3, 0, 9, 2, 12	37	359	228,2	130,8	26,16	3,2642
3	7, 0, —1, —1, 0, 1	6	52	6,0	46,0	9,20	2,2192
4	8, 0, 3, —1, 6, —1	15	111	37,5	73,5	14,70	2,6878
5	10, 5, 3, 0, 0, 1	19	135	60,2	74,8	14,96	2,7054
6	4, 8, 7, 0, —1, 1	19	131	60,2	70,8	14,16	2,6502
7	11, 2, 3, 11, 10, 0	37	355	228,2	126,8	25,36	3,2331
8	15, 10, 0, 8, 10, 10	53	589	468,2	120,8	24,16	3,1847
9	7, 7, 4, 9, 5, 6	38	256	240,7	15,3	3,06	1,1184
10	10, 2, —1, 10, —1, 7	27	255	121,5	133,5	26,70	3,2847
						253,92	28,8906

Столбцы, озаглавленные $\sum x$, $\sum x^2$, $(\sum x)^2/6$ и $\sum (x-\overline{x})^2$, содержат последовательные этапы вычисления дисперсии для

десяти станков. Величины дисперсий даны в предпоследнем столбце.

Мы хотим проверить гипотезу, состоящую в том, что эти десять значений дисперсии могли быть получены из одной и той же совокупности, т. е. что партии изделий, полученных от десяти станков, имеют одну и ту же изменчивость.

Находим натуральный логарифм каждой из дисперсий (последний столбец) и складываем их. Обозначим эту сумму через $\sum \ln \sigma_i^2$. Определяем среднее значение дисперсии: $\frac{253{,}92}{10}=25{,}392$; обозначаем это среднее через S^2. Находим натуральный логарифм его, равный 3,2343. Пусть k — число сравниваемых дисперсий (в данном случае 10), а n — число степеней свободы этих индивидуальных дисперсий (в данном случае 5).

Далее, рассчитываем две величины B и C:

$$B=kn\ln S^2-n\sum \ln \sigma_i^2=10\cdot 5\cdot 3{,}2343-$$
$$-5\cdot 28{,}8906=17{,}18,$$

$$C=1+\frac{k+1}{3nk}=1+\frac{10+1}{3\cdot 10\cdot 5}=1{,}073.$$

Если $\frac{B}{C}$ превосходит значение χ^2 при уровне значимости в 5% для $k-1$ степеней свободы, имеются основания считать с данным уровнем значимости, что дисперсии не взяты из однородной совокупности.

В данном случае $\frac{B}{C}=\frac{17{,}18}{1{,}073}=16$ для $10-1=9$ степеней свободы. Для 9 степеней свободы при уровнях значимости в 10, 5 и 2% величина χ^2, или здесь $\frac{B}{C}$, должна достигать соответственно 14,68, 16,92 и 19,68. Полученное нами значение не достигает 5%-ного уровня, но находится вблизи от него; таким образом, мы получаем указание на то, что станки значимо отличаются по изменчивости изделий.

Ясно, что если наибольшая и наименьшая дисперсии не отличаются значимо при проверке по методу отношения дисперсий, то дисперсии, имеющие промежуточные значения, также не могут отличаться друг от друга значимо; всю группу можно считать принадлежащей к единой совокупности. В таких случаях нет надобности применять критерий Бартлета.

Более общая форма критерия Бартлета, когда сравниваемые дисперсии имеют различные числа степеней свободы, здесь не будет рассматриваться.

(D) ДОВЕРИТЕЛЬНЫЕ ПРЕДЕЛЫ ДЛЯ ВЕЛИЧИНЫ ДИСПЕРСИИ

Доверительные пределы для величины дисперсии определяются так же, как и доверительные пределы для значений средних.

(I) *Малые выборки.* Желая вычислить вероятный верхний предел истинного значения дисперсии (обозначим его через φ) при уровне значимости, допустим, в 5%, пользуются формулой:

$$\varphi = \frac{n}{\chi^2}\sigma^2,$$

где σ^2 — значение дисперсии, подсчитанное по данной выборке с n степенями свободы; χ^2 берем соответствующим n степеням свободы и уровню значимости в 5%.

Итак, допустим, что имеется дисперсия σ^2, равная 10, вычисленная по выборке объемом в 10 индивидуумов (9 степеней свободы); тогда для уровня значимости в 5% имеем:

$$\chi^2 = 16{,}92,$$

$$\varphi = \frac{9}{16{,}92}\cdot 10 = 5{,}31.$$

Это — нижний предел. Соответствующий верхний предел получается из 95% уровня χ^2; величина χ^2 в этом случае при 9 степенях свободы равна 3,32.
Отсюда:

$$\varphi = \frac{9}{3{,}32}\cdot 10 = 27{,}1.$$

Рассматривая полученные пределы совместно, имеем вероятность в 5% того, что истинное значение дисперсии меньше, чем 5,31; с другой стороны, имеется вероятность в 5% того, что оно превосходит число 27,1. Общая вероятность того, что истинное значение дисперсии находится вне этих пределов, равна $5\% + 5\% = 10\%$. Поэтому, наши доверительные пределы должны оцениваться, вообще говоря, в 10%, а не в 5%.

Если желают иметь доверительные пределы ближе к 5%, можем взять величины χ^2 для уровней в 98% и 2%, что даст доверительные пределы 4%.

(II) *Большие выборки.* Для больших групп (свыше 30 индивидуумов) можно использовать тот факт, что стандартное отклонение стандартного отклонения σ составляет $\frac{\sigma}{\sqrt{2n}}$, где n — число наблюдений.

Так, при выборке из 50 индивидуумов со стандартным отклонением в 10,0 стандартное отклонение стандартного отклонения выборки σ_σ равно $\frac{10}{\sqrt{2\cdot 50}} = 1$. Доверительные пределы в 95% поэтому равны $10 \pm 2{,}01\cdot 1$ (2,01 — значение t при уровне значимости в 5% и при 50 степенях свободы).

ГЛАВА V

КРИТЕРИЙ χ^2

(А) ВВЕДЕНИЕ

Переходим к другому типу проблем, требующему приемов, отличных от рассмотренных ранее. Критерий χ^2 („хи-квадрат") применяется вообще в таких проблемах, где требуется определить, значимо ли отличается частота появления некоторого события от ожидаемого значения.

Поэтому этот критерий применим к явлениям такого рода, как появление герба или решетки при бросании монеты, частота случаев, когда железные бочки лопаются при перевозке, количество несчастных случаев в различных сменах, количество дефектных отливок, полученных различными методами.

Применяя критерий, рассчитывают ожидаемые частоты событий на основании теории, подвергаемой проверке; затем проверяют, значимо ли отличается наблюденная частота от ожидаемой. Если отличие значимо, то наша теория едва ли верна; если отличие незначимо, то наша теория находится в достаточном соответствии с данными, а отклонения могли произойти из-за случайных ошибок.

Величина χ^2 определяется так:

$$\chi^2 = \sum\left[\frac{(O-E)^2}{E}\right],$$

где O — фактически полученное (наблюденное) количество определенных результатов наблюдений (или исходов опыта), E — ожидаемое число этих результатов. Суммирование производится по всем результатам наблюдений (исходам опыта).

Обычно не рекомендуется применять критерий χ^2 в случаях, когда какое-либо E меньше 5. Этот случай будет рассмотрен особо в главе V, (H).

Теперь же рассмотрим ряд примеров.

(B) ТАБЛИЦА 1×2 („ОДИН НА ДВА")

Пусть мы бросили монету 100 раз, при этом 60 раз выпала решетка и 40 раз — герб. Можно ли сказать, что монета наверняка несимметрична?

Здесь имеются два возможных результата, или два „класса", — появление решетки или герба; по гипотезе, которую мы испытываем, математическое ожидание E для каждого класса равно 50.

Следует сделать поправку на то, что действительное распределение не является непрерывным, в то время как χ^2 является

4 Зак. 1691

непрерывной переменной. Так называемая „поправка на непрерывность" состоит в уменьшении на 0,5 значений, бо́льших ожидаемого, и в увеличении на 0,5 значений, меньших ожидаемого. Таким образом имеем

$$\chi^2 = \frac{(59{,}5-50)^2}{50} + \frac{(40{,}5-50)^2}{50} = 3{,}61.$$

Соответствующее число степеней свободы находится как число классов, значения которых могут быть заданы произвольно. В данном случае, имея сумму, мы можем заполнить произвольно только один из классов, поскольку, по выполнении этого, второй класс определяется автоматически; если монета бросалась 100 раз и 60 раз выпала решетка, то количество выпадений герба должно быть 40. Поэтому, обращаясь к таблице χ^2 для одной степени свободы, находим, что для достижения уровня значимости в 5% χ^2 должно быть равно или больше 3,84.

Следовательно, наш результат не достигает значимости с уровнем в 5%, и нельзя утверждать, что монета наверняка несимметрична.

Можно, пожалуй, подчеркнуть, что выведенное нами значение вероятности относится не к выпадению решетки точно 60 раз, а к выпадению решетки 60 или более раз, или к выпадению герба 60 или более раз. Вероятность выпадения решетки точно 60 раз при симметричной форме монеты равна:

$$\frac{100!}{60!\,40!} \cdot \frac{1}{2^{100}} = 0{,}010843867$$

(подобное вычисление может быть облегчено, если пользоваться таблицами степеней чисел и таблицами факториалов, содержащимися, например, в таблицах Барлоу).

Вероятность выпадения решетки точно 61 раз равна:

$$\frac{100!}{61!\,39!} \cdot \frac{1}{2^{100}} = 0{,}007110733.$$

Продолжая вычисления и просуммировав значения вероятностей выпадения решетки 60, 61, 62 и т. д. раз, получим сумму вероятностей 0,028443968. Это будет величина вероятности выпадения решетки 60 или более раз. Аналогичным путем вычисляем вероятность выпадения герба 60 или более раз и складываем ее с полученной ранее вероятностью выпадения решетки 60 или более раз; получим общую величину вероятности 0,056887436, или $5{,}69\%$.

Обращаясь к более подробной таблице χ^2 для одной степени свободы (например, в книге G. U. Yule and M. G. Kendall, An Introduction to the Theory of Statistics, таблица 4А в приложении), получаем величину вероятности $5{,}8\%$. Оба метода расчета, таким образом, дают хорошо совпадающие результаты.

(С) ТАБЛИЦА ЗНАЧЕНИЙ χ^2

В табл. II, в приложении, приведены значения χ^2 для величин P, бо́льших 0,10, а именно, для 0,50, 0,90, 0,95, 0,98, 0,99.

Последние значения относятся к случаям, когда совпадение получается лучшим, чем оно должно быть. Может показаться удивительным, что так случается. Однако рассмотрим случай, когда монета подбрасывалась 1000 раз в целях выяснения, являлась ли она несимметричной. Математическое ожидание для каждого из двух классов равно 500. Если мы получим в результате испытания 500 выпадений каждого класса, совпадение было бы замечательным. Затем предположим, что эксперимент был бы повторен много раз; тогда мы должны ожидать получения экспериментальных цифр с частотами, соответствующими значениям χ^2, отвечающим в свою очередь величинам P, разбросанным в диапазоне 0,95—0,05, причем в одном случае из 40 выходящим за верхний предел, и в одном случае из 40 — за нижний предел. Наиболее часто появляющимся будет значение χ^2, соответствующее $P=0{,}50$. Близкие совпадения должны случаться только с частотой, определяемой проверкой по методу χ^2; если они оказываются более частыми, можно с достаточным основанием подозревать, что данные „подтасованы".

(D) ТАБЛИЦА $1 \times n$ („ОДИН НА n")

Приведем пример такой таблицы для $n=3$.

Смена	Количество несчастных случаев в заданный период времени
A	1
B	7
C	7

Мы высказываем гипотезу, что нет связи между количеством несчастных случаев и сменой; проверим это помощью критерия χ^2. По гипотезе, математическое ожидание частоты для каждого класса равно 5,

$$\chi^2 = \frac{(5-1)^2}{5} + \frac{(7-5)^2}{5} + \frac{(7-5)^2}{5} = 4{,}8.$$

Число степеней свободы — 2 (при заданном количестве классов 3 мы можем заполнить два класса произвольно, но третий класс в таком случае определится однозначно). Обращаясь к таблице значений χ^2, находим, что при двух степенях свободы и для получения уровня значимости в 5% величина χ^2 должна быть равна или выше 5,99. Поскольку полученное значение χ^2

ниже этого значения, наша гипотеза имеет основание; другими словами, нет достаточных причин утверждать, что три смены различны в отношении подверженности несчастным случаям.

Такой результат, возможно, удивит не-статистика, который, наверно, пришел бы к заключению, что смены определенно отличаются друг от друга. Важно заметить, что статистик не утверждает, что три смены не различаются: он просто устанавливает, что свидетельство их различия недостаточно. Конечно, мы могли бы сформулировать гипотезу, что количества несчастных случаев находятся в отношении 1 : 7 : 7. Тогда χ^2 равнялось бы нулю, т. е. не было бы доказательства, что количества несчастных случаев отличаются от этого отношения. Поэтому очевидно, что данные просто недостаточны для вывода какого-либо заключения.

Предположим, что мы подождали еще некоторое время и число несчастных случаев за весь период стало соответственно 2, 14 и 14. Тогда наша таблица примет вид:

Смена	Количество несчастных случаев за весь период
A	2
B	14
C	14

Проверяем снова нашу гипотезу, что нет связи между количеством несчастных случаев и сменой.

Общее количество несчастных случаев равно 30; согласно нашей гипотезе, математическое ожидание E равно 10 для каждой категории.

Соответственно этому

$$\chi^2=\frac{(2-10)^2}{10}+\frac{(14-10)^2}{10}+\frac{(14-10)^2}{10}=9{,}6.$$

Такое значение χ^2 при двух степенях свободы больше значения, соответствующего уровню значимости в 1%; иначе говоря, по нашей гипотезе, такое значение могло бы быть получено реже, чем 1 раз из 100.

Таким образом, дополнительные данные теперь сделали очень вероятным предположение, что смены различаются в отношении подверженности несчастным случаям.

В табл. 5.1 приведены некоторые данные о процессе размалывания в ролле*. Испытывались четыре разных метода загрузки ролла; количество проведенных циклов показано в первой строке.

* Машина для размалывания бумажной массы. (*Прим. перев.*)

Во второй горизонтальной графе показано количество происшедших засорений между валом и станиной.

ТАБЛИЦА 5.1

Метод загрузки	A	B	C	D	Итого
Количество циклов . . .	8	10	9	13	40
Количество происшедших засорений	5	8	9	10	32
Ожидавшееся количество засорений	6,4	8	7,2	10,4	32

Мы желаем проверить гипотезу, что нет связи между частотой случаев засорений и методом загрузки, т. е., что частота случаев засорений независима от метода загрузки. Согласно этой гипотезе, частота засорения должна быть одинакова при любом методе, т. е. должна равняться $\frac{32}{40}=0{,}8$ на цикл. Третья строка в таблице содержит значения математического ожидания числа засорений для каждого из четырех методов загрузки. Вычисляем χ^2:

$$\chi^2=\frac{1{,}4^2}{6{,}4}+\frac{0^2}{8}+\frac{1{,}8^2}{7{,}2}+\frac{0{,}4^2}{10{,}4}=0{,}77.$$

Мы использовали одну степень свободы, принимая ограничение, что ожидаемый итог равен действительному (если бы мы имели какую-либо величину ожидания *a priori*, например, признали бы гипотезу, что случалось одно засорение за цикл, тогда мы не потеряли бы этой степени свободы). Обращаясь к таблице χ^2 для трех степеней свободы, видим, что для достижения уровня значимости в 5% мы должны были бы получить $\chi^2=7{,}82$. Ясно, что приведенные данные прекрасно согласуются с гипотезой, что частота засорений не зависит от метода загрузки.

Заметим, что табл. 5.1 кажется с виду таблицей $2\times n$ [см. далее раздел (F)]. Но в таблице $2\times n$ сравниваемые частоты относятся к одному и тому же классу событий. В нашем же случае факт осуществления цикла не относится к тому же классу, что и факт засорения. Поэтому имеем дело с таблицей $1\times n$, содержащей сведения о происшедших засорениях; другая горизонтальная графа выписана только с целью вывода ожидаемых частот.

(E) ТАБЛИЦА 2×2 („ДВА НА ДВА“)

Одна из часто встречающихся проблем может быть выражена в такой форме:

В производственном процессе получается a случаев брака в выборке из $a+b$ индивидуумов; в процессе другого типа получается c случаев брака в выборке из $c+d$ индивидуумов. Являются ли обе совокупности различными или одинаковыми?

Такого рода проблема может встретиться, например, в производстве кордита*, где могут считаться браком патроны, имеющие, допустим, наружные пороки; мы заинтересованы выяснить, дает ли один из методов набивки прессового цилиндра меньше брака, чем другой.

Точно такая же по форме проблема возникает при проверке изделий с помощью набора постоянных калибров вместо применения непрерывно-регулируемых калибров; если изделие не проходит испытания по тому или иному калибру, оно является браком, если проходит — считается удовлетворительным.

Допустим, что в первой выборке в 1000 изделий было 100 случаев, т. е. 10% брака; в другой выборке в 500 изделий было 60 случаев, т. е. 12% брака.

Оправдывается ли с достаточной уверенностью предположение, что совокупности, представляемые этими двумя партиями, различны?

Составим таблицу наблюдений в форме:

Процесс	Брак	Удовлетворит. издел.	Итого
A	$a=100$	$b=900$	$a+b=1000$
B	$c=60$	$d=440$	$c+d=500$
Итого	$a+c=160$	$b+d=1340$	1500

Подсчитываем величину математического ожидания для каждой ячейки: она прямо пропорциональна итогам соответствующих строк и столбцов и обратно пропорциональна общему итогу; таким образом, математические ожидания для a, b, c и d составляют:

$$\frac{160\cdot 1000}{1500}=106{,}7;\quad \frac{1340\cdot 1000}{1500}=893{,}3;\quad \frac{160\cdot 500}{1500}=53{,}3;$$
$$\frac{1340\cdot 500}{1500}=446{,}7.$$

Соответственно этому,

$$\chi^2=\frac{6{,}7^2}{106{,}7}+\frac{6{,}7^2}{893{,}3}+\frac{6{,}7^2}{53{,}3}+\frac{6{,}7^2}{446{,}7}=1{,}4.$$

* Бездымный нитроглицериновый порох. (*Прим. перев.*)

Число степеней свободы — единица (при наличии итогов по строкам и столбцам мы можем заполнить произвольно только одну ячейку, так как данные в других этим однозначно определяются). Полученное значение χ^2 соответствует примерно уровню значимости в 20%; таким образом, кажущееся превосходство процесса A этой проверкой не подтверждается.

В вышеуказанном примере мы опустили поправку на непрерывность: при таких больших числах это допустимо, поскольку получающаяся разница очень невелика.

В примере мы вычислили математические ожидания для каждой ячейки отдельно, затем подсчитали χ^2 как сумму отдельных членов. Однако можно было бы воспользоваться удобной формулой:

$$\chi^2=\frac{(ad-bc)^2\cdot N}{(a+b)(b+d)(c+d)(a+c)},$$

где

$$N=a+b+c+d.$$

Подставляем цифры вышеприведенного примера:

$$\chi^2=\frac{(100\cdot 440-900\cdot 60)^2\cdot 1500}{160\cdot 1340\cdot 500\cdot 1000}=1{,}4.$$

Результат тот же, что и ранее.

Если объемы выборок малы, важно вводить поправку на непрерывность, как показано в следующем примере.

В результате прессования одного замеса пороховой массы получено 9 бракованных прессовок из 45 (20% брака); при прессовании другого замеса получено 29 бракованных прессовок из 87 (33% брака).

	Замес A	Замес B	Итого
Удовлетворительные результаты	$58=a$	$36=b$	94
Брак	$29=c$	$9=d$	38
Итого	87	45	132

Как и ранее, метод внесения поправки на непрерывность заключается в прибавке $\frac{1}{2}$ к числам тех ячеек, которые ниже ожидаемых, и в вычитании $\frac{1}{2}$ из чисел тех ячеек, которые выше ожидаемых. Итоги от этого, конечно, не изменятся:

$$\chi^2=\frac{(58{,}5\cdot 9{,}5-28{,}5\cdot 35{,}5)^2\cdot 132}{94\cdot 45\cdot 38\cdot 87}=1{,}96,$$

Для одной степени свободы этот результат совершенно незначим. Если бы мы пренебрегли поправкой на непрерывность, величина χ^2 получилась бы 2,57 — в данном случае тоже незначимая. Однако в пограничных случаях разница может оказаться существенной.

(F) ТАБЛИЦА $2 \times n$ („ДВА НА n")

Данные табл. 5.2 относятся к двум типам происшествий, двум определенным типам аварий A и B, которые могут происходить в некотором производственном процессе.

Процесс может вестись на трех сортах сырья — L, M и N.

ТАБЛИЦА 5.2

	L	M	N	Итого
A	42	13	33	88
B	20	8	25	53
Итого . .	62	21	58	141

Выдвигаем гипотезу, что нет связи между родом аварии и сортом сырья. Другими словами, каждая из аварий происходит одинаково часто при всех трех сортах сырья. Проверим гипотезу с помощью критерия χ^2.

Математическое ожидание для ячейки LA составляет $\frac{62 \cdot 88}{141} = 38{,}7$; соответствующий член в выражении χ^2 равен $\frac{(42-38{,}7)^2}{38{,}7} = 0{,}28$.

Аналогичным путем находим, что математическое ожидание для ячейки MA составляет $\frac{21 \cdot 88}{141}$ и т. д.

Число степеней свободы для таблицы $n_1 \times n_2$ равно

$$(n_1 - 1)(n_2 - 1);$$

в данном случае оно равно

$$(2-1)(3-1) = 2.$$

Величина χ^2 после проведения подсчета по всем шести ячейкам будет равна 1,5, т. е. намного ниже той, которая соответствует уровню значимости в 5%. Поэтому можно заключить, что данные прекрасно согласуются с нашей гипотезой.

Подходя к задаче не статистически, мы подсчитали бы отношения $\frac{A}{B}$ для сортов сырья L, M и N: $\frac{42}{20} = 2{,}1$; $\frac{13}{8} = 1{,}63$; $\frac{33}{25} = 1{,}32$. Вероятно, мы решили бы, что эти отношения значительно отличаются друг от друга. Применение проверки с помощью критерия χ^2 показывает, однако, что подобное заключение было бы неоправданным.

Как для таблицы 2×2, так и для таблицы $2 \times n$ существует удобный механический способ подсчета χ^2. Пусть величины частот в парах ячеек равны a и A, b и B и т. д.; соответствующие итоги равны n и N. Подсчитываем для каждой пары:

$$p = \frac{a}{a+A} \quad \text{и} \quad ap = \frac{a^2}{a+A}$$

и для итога

$$P = \frac{n}{n+N}.$$

Тогда

$$\chi^2 = \frac{1}{P(P-1)} [\Sigma(ap) - nP].$$

В нашем случае три пары ячеек дают такие значения ap:

$$\frac{42^2}{62} = 28{,}452, \quad \frac{13^2}{21} = 8{,}048, \quad \frac{33^2}{58} = 18{,}776.$$

Сумма $\Sigma(ap)$ равна 55,276, $P = \frac{88}{141} = 0{,}624$, $nP = \frac{88^2}{141} = 54{,}972$.

Отсюда $\chi^2 = (55{,}276 - 54{,}972) \cdot \frac{1}{0{,}624 \cdot 0{,}376} = 1{,}51$, как и прежде

(G) ТАБЛИЦА $n_1 \times n_2$ („n_1 НА n_2")

Приведем пример такой таблицы для $n_1 = 3$, $n_2 = 6$:

	a	b	c	d	e	f	Итого
A	11	23	8	5	18	18	83
B	17	29	10	17	7	15	95
C	6	21	8	24	15	9	83
Итого . . .	34	73	26	46	40	42	261

Таблица показывает результаты производства некоторого продукта в трех сменах — A, B и C; продукт классифицирован по качеству на 6 сортов от a до f.

Каждая из трех смен в среднем снабжалась сырьем одного и того же качества.

Возник вопрос, не производит ли одна смена продукт лучшего качества, чем другие, т. е. не имеется ли сравнительного избытка некоторых качеств у продукции какой-либо смены.

Для проверки гипотезы об отсутствии связи между качеством продукции и сменой подсчитываем величину χ^2 по данным таблицы. Математическое ожидание по ячейке Aa равно $\frac{83\cdot 34}{261}=$ $=10{,}8$; соответствующий член, входящий в выражение χ^2, равен $\frac{(11-10{,}8)^2}{10{,}8}=0{,}004$. Те же величины для ячейки Ba равны $\frac{95\cdot 34}{261}=$ $=12{,}38$ и $\frac{(17-12{,}38)^2}{12{,}38}=1{,}72$.

Проведя подсчеты для всех ячеек и просуммировав члены, входящие в выражение для χ^2, получаем $\chi^2=26{,}3$. Число степеней свободы задано числом независимых путей, которыми таблица данных может быть заполнена при заданных итогах. Для таблицы $n_1\times n_2$ это число равно $(n_1-1)(n_2-1)$ — в данном случае $5\cdot 2=10$. Обращаясь к таблице значений χ^2 в строке для 10 степеней свободы, видим, что вероятность получения столь большого значения χ^2 менее $0{,}01=1\%$.

Поэтому, гипотеза об отсутствии связи между качеством продукции и сменой оказывается неправильной, т. е. такая связь существует.

Проверка помощью критерия χ^2, определяющая значимость связи между двумя переменными, не дает никакой информации о типе связи. Последнее можно найти, только обращаясь к исходной таблице данных. Для этого удобно внести в каждую ячейку значение ожидаемой частоты наряду со стоящим там значением наблюденной частоты. Сделав это для настоящего примера, найдем, что смена A выработала некоторый относительный избыток качеств „e" и „f", смена B — качества „a" и смена C — качества „d".

(H) УСЛОВИЕ, ЧТО МАТЕМАТИЧЕСКОЕ ОЖИДАНИЕ ЧАСТОТЫ НЕ ДОЛЖНО БЫТЬ МЕНЕЕ 5

Как упоминалось в начале этой главы, не следует вести проверку с помощью критерия χ^2 в случаях, когда математическое ожидание частоты для какого-либо класса менее 5.

Иногда это означает, что проверку нельзя сделать. В иных случаях это затруднение можно обойти.

Рассмотрим следующую таблицу:

	L_1	L_2	M	N_1	N_2	Итого
A	5	37	13	28	5	88
B	3	17	8	20	5	53
Итого ..	8	54	21	48	10	141

Эта таблица содержит данные, приведенные в табл. 5.2 раздела (F). Математическое ожидание для ячеек L_1B и N_2B менее 5. Применить проверку к данным этой таблицы в том виде, как она представлена, нельзя. Однако, если мы объединим данные столбцов L_1 и L_2, а также N_1 и N_2, получим совершенно удовлетворительную таблицу 2×3, рассмотренную в разделе (F).

ГЛАВА VI

РАСПРЕДЕЛЕНИЕ ПУАССОНА

(A) ВВЕДЕНИЕ

Рассмотрим некоторый производственный процесс, в котором одна и та же операция повторяется большое число раз каждый день, причем иногда, случайно, происходит воспламенение.

Если количество воспламенений за отдельные дни выписать в календарном порядке, получится такой ряд чисел:

0 0 1 0 0 2 0 3 0 0 0 1 3 0 1 4 0 0 0 2...

В некоторые дни воспламенений не было, в другие дни их было сравнительно много. Рассмотрим вопрос, была ли вероятность воспламенения постоянной в течение всего периода времени. В первом приближении мы могли бы принять, что дни, когда происходило 0 или 1 воспламенение, характеризовались малой вероятностью возникновения воспламенения, а дни с 3, 4 и более воспламенениями — более высокой вероятностью. Затем мы стали бы искать причины изменения этой вероятности, т. е. исследовать, чем дни с высоким количеством воспламенений отличались от других дней, имея конечной целью внести должные поправки в производственный процесс. С другой стороны, если вероятность возникновения воспламенения в действительности постоянна, мы напрасно потратили бы время на выяснение, чем одни дни отличались от других.

Если принять, что вероятность воспламенения в любой момент мала, но конечна по величине и постоянна, то в течение конечного периода времени шансы возникновения происшествия могут быть довольно значительными. Происшествие может быть любого типа: здесь мы рассматриваем самовоспламенение в производственном процессе; в другом случае может рассматриваться появление бракованных изделий в массовом производстве.

В этом последнем случае следует обращать должное внимание на предварительное условие, что вероятность события (например, вероятность получения бракованного экземпляра при случайном выборе) в каждый момент должна быть мала. В практике приближение считается достаточно хорошим, если процент брака менее 10.

(В) ЧИСЛО СОБЫТИЙ ЗА ПРОМЕЖУТОК ВРЕМЕНИ

Обращаясь к нашему примеру самопроизвольного воспламенения, подсчитаем количество дней, в течение которых не было ни одного происшествия, затем количество дней, в течение которых было по 1, по 2 и т. д. происшествий. Если наша гипотеза правильна, то относительные частоты должны быть такими, как показано в табл. 6.1, где e — основание натуральных логарифмов, m — средняя частота, а, например, 4! обозначает произведение $4 \cdot 3 \cdot 2 \cdot 1$.

Ряд чисел в правом столбце известен под названием ряда Пуассона. Сумма относительных частот равна 1; действительно:

$$e^{-m} + \frac{m}{1!} e^{-m} + \frac{m^2}{2!} e^{-m} + \frac{m^3}{3!} e^{-m} + \ldots =$$

$$= e^{-m} \left[1 + \frac{m}{1!} + \frac{m^2}{2!} + \frac{m^3}{3!} + \ldots\right] = e^{-m} \cdot e^{m} = 1.$$

ТАБЛИЦА 6.1

Количество происшествий за один день	Относительная частота
0	e^{-m}
1	$\frac{m}{1!} e^{-m}$
2	$\frac{m^2}{2!} e^{-m}$
3	$\frac{m^3}{3!} e^{-m}$
4	$\frac{m^4}{4!} e^{-m}$
и т. д.	и т. д.

В первых двух столбцах табл. 6.2 сведены результаты наблюдений над производственным процессом за 201 день. Событием являлось воспламенение, возникавшее в процессе, который повторялся большое число раз каждый день. В течение периода в 201 день было 150 таких событий, распределение которых и приведено в табл. 6.2.

102 дня прошло без воспламенений, в 59 было по 1 воспламенению, в 31 — по 2 воспламенения и т. д. Общее число дней: $102 + 59 + 31 + 8 + 0 + 1 = 201$. Общее число событий (воспламенений): $102 \cdot 0 + 59 \cdot 1 + 31 \cdot 2 + 8 \cdot 3 + 0 \cdot 4 + 1 \cdot 5 = 150$.

ТАБЛИЦА 6.2

Количество происшествий за один день	Количество событий такого типа	Математическое ожидание числа событий
0	102	95,3
1	59	71,1
2	31	26,5
3	8 } 9	6,6 } 8,1
4	0	1,3
5	1	0,2
6	0	0
↓		
Итого	201	201,0

Мы имеем 150 происшествий за 201 день, так что средняя дневная частота $m = 0,7463$. По таблице натуральных логарифмов $e^m = 2,109$; тогда $e^{-m} = 0,4742$. Начальный член ряда, соответствующий 0 происшествий за день, равен $Ne^{-m} = 201 \cdot 0,474 = 95,3$; следующий член равен $\frac{m}{1!} Ne^{-m} = 0,7463 \cdot 95,3 = 71,1$; третий член равен $\frac{m^2}{2!} Ne^{-m} = 0,7463 \cdot \frac{71,1}{2} = 26,5$ и т. д.

Эти значения математических ожиданий по Пуассону приведены в последнем столбце табл. 6.2.

Для оценки значимости отклонения наблюденных частот от ожидаемых сделаем проверку с помощью критерия χ^2 обычным путем, как для таблицы $1 \times n$.

Слагаемое, входящее в χ^2 от первой ячейки, равно

$$\frac{(102 - 95,3)^2}{95,3} = 0,47.$$

Аналогичные вычисления делаются и для других ячеек. При проведении проверки по методу χ^2 важно объединить достаточное количество ячеек в конце таблицы, чтобы удовлетворить условию, что математическое ожидание частоты в каждой ячейке должно быть не меньше 5.

Следующим важным условием при проведении проверки с помощью критерия χ^2 является то, что мы должны принять число степеней свободы равным $n - 2$. Причина потери двух степеней свободы, а не одной, как в обычной таблице $1 \times n$, в том, что мы должны не только согласовать итоги (теряя при этом одну степень свободы, как и в таблице $1 \times n$), но и использовать одну постоянную для получения соответственной формы распределения; последнее обстоятельство отнимает вторую степень свободы.

В вышеуказанном примере величина χ^2 получается равной 3,4 при 2-х степенях свободы. Это соответствует вероятности в 15%, так что можно считать, что гипотеза удовлетворительно согласуется с данными. (Величина P должна быть 5% или меньше, для того чтобы опровергнуть гипотезу о согласованности между наблюденным и ожидаемым распределением.)

Заметим, что форма распределения по Пуассону зависит от величины средней частоты m. Если m меньше единицы, как было в примере, то каждый последующий член в распределении будет меньше предыдущего; первый член будет при этом наибольшим.

Если m больше 1, но меньше 2, то второй член будет больше первого, поскольку отношение второго члена к первому равно $\frac{m}{1}$; но второй член будет больше третьего, поскольку отношение третьего члена ко второму равно $\frac{m}{2}$; следовательно, второй член является в данном случае максимальным. Если m более 2 и менее 3, то третий член будет больше второго и четвертого.

Табл. 6.3 составлена по данным предыдущего примера. Промежуток взят трехдневный, так что средняя частота составляет теперь 2,2537.

ТАБЛИЦА 6.3

Количество происшествий за трехдневные промежутки	Количество событий такого типа		Математическое ожидание события	
	8		7,04	
1	20		15,86	
2	13		17,87	
3	12		13,42	
4	6		7,56	
5	5	8	3,41	5,24
6	2		1,28	
7	0		0,41	
8	1		0,11	
9	0		0,03	

Видно, что ожидаемые частоты близко соответствуют наблюденным. Величина χ^2 равна 4,49; при 4-х степенях свободы это соответствует вероятности в 40%.

1345 к

(С) РАСПРЕДЕЛЕНИЕ ПРОМЕЖУТКОВ ВРЕМЕНИ

Выше рассматривалось количество происшествий за данный промежуток времени. Несколько иным подходом к проблеме служит рассмотрение длительности промежутка между двумя событиями. Промежуток может меняться от нуля до десяти и более дней. Если принять, что вероятность происшествия постоянна в течение всего периода, можно предсказать относительные частоты различных длительностей интервала.

Пусть m является обратной величиной средней длительности такого промежутка, а N — общее количество наблюденных промежутков. Тогда число промежутков в пределах длительности от t_1 до t_2 определяется формулой:

$$N(e^{-mt_1} - e^{-mt_2}),$$

где e — основание натуральных логарифмов.

В табл. 6.4 приведено распределение промежутков времени по данным, послужившим для составления предыдущей таблицы.

ТАБЛИЦА 6.4

Промежутки времени (в часах)	Наблюденное количество промежутков	Ожидаемое количество промежутков
0 — 7,9	48	44,46
8 — 15,9	36	34,63
16 — 23,9	27	26,95
24 — 31,9	24	21,00
32 — 39,9	10	16,44
40 — 47,9	13	12,64
48 — 55,9	10	10,01
56 — 63,9	8	7,64
64 — 71,9	9	6,03
72 — 79,9	3	4,76
80 и более	13	16,45
Итого	201	201,01

Расчеты облегчаются, если принять 8 часов за единицу времени. Используя более мелкую разбивку по времени, чем указано в табл. 6.4, мы получили бы, что средняя длительность промежутка составляет 3,9950 единиц времени, а m равно 0,2503.

Математическое ожидание для первого промежутка равно:

$$N(e^{-mt_1} - e^{-mt_2}) = 201 \cdot (e^{-0{,}2503 \cdot 0} - e^{-0{,}2503 \cdot 1}) = 44{,}46.$$

Математическое ожидание для второго промежутка

$$N(e^{-mt_2} - e^{-mt_3}) = 201 \cdot (e^{-0,2503 \cdot 1} - e^{-0,2503 \cdot 2}) = 34,63.$$

Эти математические ожидания даны в последнем столбце табл. 6.4.

Согласие с гипотетическим распределением можно проверить с помощью критерия χ^2. Величина χ^2 оказалась равной 6,13, что при $11-2=9$ степенях свободы соответствует вероятности в 75%. Совпадение, таким образом, является превосходным.

Распределение событий по времени, выраженное в такой форме, любопытно в том смысле, что чаще всего встречается кратчайший промежуток.

Интересно отметить, что с помощью этих двух методов проверяются весьма различные аспекты распределения. Большое количество длительных промежутков (например, в 4 дня и более) проявилось бы в настоящем методе в виде большого числа в последней строке; поскольку же ожидаемое число в этой строке мало, то соответствующее слагаемое в выражении для χ^2 окажется довольно значительным. В первом же из описанных методов, однако, подобные данные попали бы в рубрику дней с нулевым количеством событий; поскольку эта рубрика сама обширна, ее увеличение должно быть очень велико, для того чтобы привносимое ею слагаемое в χ^2 увеличилось заметно. Ясно, что ряд дней с нулевым количеством событий более заметен, когда такие дни следуют друг за другом подряд, нежели в случаях, когда они перемежаются днями с событиями. Последняя форма обработки данных более приспособлена для обнаружения этого вида отклонения от распределения по Пуассону. С другой стороны, первая форма обработки более приспособлена к обнаружению скоплений большого количества событий в течение короткого промежутка времени.

ГЛАВА VII

ДИСПЕРСИОННЫЙ АНАЛИЗ

(А) ВВЕДЕНИЕ

Ценным свойством дисперсии является то, что если в процессе имеется несколько факторов, каждый из которых вносит какую-то долю в величину дисперсии конечного продукта, то общая дисперсия равна сумме составляющих дисперсий. При этом предполагается, что факторы действуют независимо.

Это положение не столь очевидно, как может показаться. Так, если мы используем величину стандартного отклонения, как меру изменчивости, было бы неверно сказать, что стандартное отклонение конечного продукта равно сумме стандартных отклонений, вносимых отдельными факторами.

Указанное свойство аддитивности дисперсии позволяет применять технический прием, известный под названием „дисперсионного анализа“. Этот прием состоит в разложении дисперсии процесса по составляющим факторам, относительная важность которых может тогда быть оценена.

Дисперсионный анализ может применяться в различных формах в зависимости от структуры исследуемого процесса; выбор надлежащей формы является обычно одной из главных трудностей в практическом применении анализа. В настоящей главе будут рассмотрены только две простейших формы; в одной из последующих глав мы попытаемся рассмотреть общий прием обработки более сложных случаев, часто встречающихся в практике.

(В) ДИСПЕРСИОННЫЙ АНАЛИЗ ПО ПАРТИЯМ ПРОДУКТА И МЕЖДУ НИМИ

Допустим, имеется некоторый процесс, состоящий из двух стадий. В стадии A производится операция A над сырым материалом, поступающим партиями по 300 фунтов (136 *кг*) каждый раз. Эта операция не является совершенно точно воспроизводимой и вызывает дисперсию σ_A^2 в конечном продукте. В стадии B процесса каждая партия в 300 фунтов делится на три партии по 100 фунтов каждая, которые подвергаются затем операции B.

Операция B также не является совершенно точно воспроизводимой и вызывает дисперсию σ_B^2 в конечном продукте. Проблема состоит в рассмотрении общей величины дисперсии конечного продукта и разложении ее на компоненты σ_A^2 и σ_B^2. Выполнив это, мы будем в состоянии решить, какая часть процесса потребует более тщательного контроля для того, чтобы уменьшить величину общей дисперсии.

Для определения σ_A^2 и σ_B^2 мы приступаем прежде всего к расчету величины $n\sigma_A^2+\sigma_B^2$ и σ_B^2 (в общем случае, когда каждая основная партия подразделяется на n более мелких). Таким путем становится осуществимой оценка значимости σ_B^2.

Функция $n\sigma_A^2+\sigma_B^2$ имеет $m-1$ степеней свободы, где m — число основных партий; σ_B^2 имеет $m(n-1)$ степеней свободы. Поскольку σ_A^2 существует, величина $n\sigma_A^2+\sigma_B^2$ должна быть больше σ_B^2; мы можем легко проверить это по методу отношения дисперсий.

В табл. 7.1 приведены числовые значения рассматриваемого свойства продукта для каждой партии (всего их mn, в нашем случае $3 \cdot 10 = 30$); значения даны столбцами, по 3 в каждом; каждый столбец содержит данные по трем партиям, полученным из одной основной.

Для упрощения арифметических вычислений из всех данных была вычтена одна и та же постоянная; очевидно, что это смещение нулевой точки никоим образом не повлияет на относительную изменчивость.

ТАБЛИЦА 7.1

	−3 −2 −3	−4 −3 −5	−3 −1 −4	−1 3 2	4 3 6	−2 3 1	1 0 1	2 1 1	−1 −1 2	−1 1 1	Общий итог
Итого	−8	−12	−8	4	13	2	2	4	0	1	−2

Вычисления ведутся так:

(1) Складываются квадраты индивидуумов:

$$(-3)^2+(-4)^2+(-3)^2+\ldots+1^2+2^2+1^2=204.$$

(2) Складываются квадраты итогов по столбцам; полученный общий итог делится на число индивидуумов в каждом столбце:

$$\frac{(-8)^2+(-12)^2+(-8)^2+\ldots+0^2+1^2}{3}=160{,}67.$$

(3) Квадрат общего итога по всем индивидуумам делится на общее число индивидуумов:

$$\frac{(-3-4-3-\ldots+1+2+1)^2}{30}=0{,}13.$$

Затем выписывается таблица этого дисперсионного анализа:

ТАБЛИЦА 7.2

Источники дисперсии	Сумма квадратов	Число степеней свободы	Средние квадратов	Компоненты дисперсии
Между столбцами	(2) — (3) = 160,54	$m-1=9$	17,84	$n\sigma_A^2+\sigma_B^2$
По столбцам . . .	(1) — (2) = 43,33	$mn-m=20$	2,17	σ_B^2
Итого	(1) — (3) = 203,87	$mn-1=29$		

m — число столбцов,

n — число индивидуумов в столбце,

σ_A^2 — дисперсия вследствие различий между столбцами,

σ_B^2 — дисперсия по столбцам.

Три суммы квадратов являются разностями между соответствующими величинами (1), (2) и (3), рассчитанными ранее.

Числа степеней свободы выводятся следующим образом:

(a) *Между столбцами.* Число столбцов $=m$, поэтому число степеней свободы $=m-1$. (Число степеней свободы обычно равно числу путей, какими рассматриваемая группа может быть произвольно заполнена; если итог задан, то, по заполнении $m-1$ классов, последний класс однозначно определяется; таким образом, можно заполнить произвольно только $m-1$ классов.)

(b) *По столбцам.* В каждом столбце n индивидуумов; каждый столбец поэтому дает $n-1$ степеней свободы; всего имеется m столбцов, поэтому общее число степеней свободы $=m(n-1)=$ $=mn-m$.

(c) *Общее.* Общее число степеней свободы равно $mn-1$, на единицу меньше общего числа наблюдений.

Среднее квадратов равно соответствующей сумме квадратов, деленной на число степеней свободы.

Слово „квадраты" в заголовках столбцов табл. 7.2 относится к квадратам отклонений от средних величин; таким образом, значения средних квадратов имеют некоторые из характеристик дисперсий.

В последней вертикальной графе таблицы показано, что среднее квадратов между столбцами оценивает величину $n\sigma_A^2+\sigma_B^2$, а среднее квадратов по столбцам оценивает σ_B^2. Таким образом, если σ_A^2 действительно существует, то первое значение средних квадратов должно быть заметно больше второго.

Мы желаем проверить, значимо ли превышает среднее квадратов между столбцами среднее квадратов по столбцам. Это можно выполнить, испытывая отношение дисперсий по Фишеру (см. выше, гл. IV, (A)).

Подсчитываем отношение большего среднего квадратов к меньшему, $\frac{17{,}84}{2{,}17}=8{,}22$, и обращаемся к таблице значений F для n_1 степеней свободы большей дисперсии (в данном случае $n_1=9$) и n_2 степеней свободы меньшей дисперсии (в данном случае $n_2=20$). Если наше значение F больше указанного в таблице, то результат более значим, чем уровень, приведенный в таблице. В данном случае, например при $n_1=8$ (в таблице нет значений для $n_1=9$) и $n_2=20$, величина F имеет значения 2,45, 3,56 и 5,44 для уровней значимости соответственно в 5, 1 и 0,1%. Наш результат (=8,22) свидетельствует о значимости выше 0,1%.

Установив таким образом, что среднее квадратов между столбцами значительно больше среднего квадратов по столбцам, т. е. что дисперсия между столбцами существует, приступаем к ее вычислению.

Имеем

$$n\sigma_A^2+\sigma_B^2=17{,}84,$$
$$\sigma_B^2=\ 2{,}17,$$

где n — число индивидуумов в столбце, в данном случае 3.

Поэтому

$$\sigma_A^2=\frac{17{,}84-2{,}17}{3}=5{,}22, \qquad \sigma_B^2=2{,}17.$$

Полная величина дисперсии в наших исходных данных, таким образом, составилась из двух компонент: $\sigma_A^2=5{,}22$ — вследствие различий между столбцами (в данном случае столбцы представляют основные партии сырья) и $\sigma_B^2=2{,}17$ — вследствие изменчивости по столбцам (в данном случае в пределах основных партий, т. е. между партиями продукта, составляющими основную).

(C) ИССЛЕДОВАНИЕ МНОГОСТУПЕННЫХ ПРОЦЕССОВ

Бóльшая часть общей величины изменчивости в предыдущем процессе возникает в стадии A и только меньшая часть — в стадии B. Таким образом, при попытке стандартизации процесса в целях уменьшения общей величины изменчивости главное внимание должно быть уделено стадии A.

Общая дисперсия процесса (в момент исследования) равна $\sigma_T^2=\sigma_A^2+\sigma_B^2=5{,}22+2{,}17=7{,}39$; таким образом, его стандартное отклонение равно $\sqrt{7{,}39}$. Наша оценка разброса, в пределах которого должны находиться 95% всех индивидуумов, выражается цифрой $\pm 1{,}96\cdot\sqrt{7{,}39}=\pm 5{,}3$, всего (в обе стороны) $2\cdot 5{,}3=10{,}6$. Это подтверждается и в результате рассмотрения исходных данных: из тридцати данных наблюдения наименьшее составляет -5, а наибольшее $+6$; общий разброс равен 11,0.

Допустим, что все усилия были направлены на уточнение процесса B; работа прошла вполне успешно, так что σ_B^2 уменьшилась до нуля. Полная величина дисперсии теперь стала бы $\sigma_A^2=5{,}22$, стандартное отклонение 2,29 и разброс 95% индивидуумов $\pm 1{,}96\cdot 2{,}29=\pm 4{,}5$; полный разброс стал бы равен 9. Поскольку полный разброс был вначале 10,6, улучшение оказалось незначительным.

Если, с другой стороны, усилия были бы направлены на уточнение стадии процесса A и нам удалось бы свести σ_A^2 к нулю, общая дисперсия была бы $\sigma_B^2=2{,}17$, а стандартное отклонение 1,47; разброс 95% индивидуумов стал бы $\pm 1{,}96\cdot 1{,}47=\pm 2{,}9$; полный разброс 5,8.

Улучшение, достигнутое уточнением стадии A, таким образом, намного более существенно по сравнению с тем, которое достигается улучшением стадии B. Ясно, что усилия, затраченные на исследование стадии B, были бы в значительной мере бесплодны. До проведения детального физико-химического исследования многоступенного процесса, очевидно, желательно определить путем показанного дисперсионного анализа, какие ступени больше всего влияют на изменчивость, и сосредоточить внимание на них. Выгода такого метода состоит в том, что стадии процесса, рассмотрение которых признано нецелесообразным, могут содержать большое число переменных, каждая из которых при ином подходе потребовала бы исследования.

Дисперсионный анализ заставляет вначале сосредоточиться на структуре процесса, прежде чем переходить к исследованию отдельных переменных.

(D) ДИСПЕРСИОННЫЙ АНАЛИЗ ПРИ НЕРАВНЫХ СТОЛБЦАХ

Следует заметить, что дисперсионный анализ между столбцами и по столбцам является, в сущности, развитием „критерия t“ Стьюдента. Вместо сравнения двух средних сравниваются m средних, где m — число столбцов. Можно было бы, казалось, сделать серию проверок с помощью „критерия t“ для средних значений столбцов, беря их попарно. Это потребовало бы большой затраты труда и, кроме того, вызвало бы ряд усложнений. Так, при наличии 20 сравнений и работая до получения уровня значимости в 5%, даже в случае, когда столбцы составлены из одной и той же совокупности, мы ожидали бы, что в среднем один из них дает значение t, достигающее 5%. При применении дисперсионного анализа эти трудности обходятся.

Число индивидуумов в каждом столбце иногда бывает неодинаково; тогда прием, показанный в разделе (B), непосредственно неприменим.

В нижеследующей таблице приведены данные о продукции, полученной от отдельных производственных единиц завода до выхода их из строя из-за коррозии. Данные распределены по литейным цехам.

Литейная	Полученная продукция	Итог	Среднее
A	84, 60, 40, 47, 34	265	53,0
B	67, 92, 95, 40, 98, 60, 59, 108, 86	705	78,3
C	46, 93, 100	239	79,7

С первого взгляда можно было бы притти к выводу, что тигли литейной A дают в среднем более низкую производительность, чем тигли литейных B и C.

Анализ начинают с расчета следующих величин:

(1) Сумма квадратов индивидуумов:

$$84^2+60^2+40^2+\ldots+93^2+100^2=95\,709.$$

(2) Сумма квадратов итогов каждой строки, деленных на числа индивидуумов в строках:

$$\frac{265^2}{5}+\frac{705^2}{9}+\frac{239^2}{3}=88\,310.$$

(3) Квадрат общего итога, деленный на общее число индивидуумов:

$$\frac{(265+705+239)^2}{17}=85\,981.$$

Дисперсионный анализ проводится так:

Источники дисперсии	Суммы квадратов	Числа степеней свободы	Средние квадратов
Между литейными . .	(2) — (3) = 2329	2	1164
По литейным	(1) — (2) = 7399	14	528,5
Итого	(1) — (3) = 9728	16	

Число степеней свободы определяется так:

(a) Между литейными — на единицу меньше числа литейных.

(b) По литейным — по каждой литейной на единицу меньше числа наблюдений, относящегося к ней, т. е. $4+8+2=14$.

(c) Итоговое — на единицу меньше общего количества наблюдений.

Сопоставляя среднее квадратов между литейными и по литейным, получаем отношение дисперсий $\frac{1164}{528,5}=2,2$ при числах степеней свободы $n_1=2$, $n_2=14$. Эта цифра значительно ниже соответствующей уровню значимости в 5%; таким образом, нет достаточных оснований считать, что тигли литейной A имеют более низкую производительность, чем тигли других литейных.

В данном случае мы более всего заинтересованы получить действительные средние по трем литейным; мы применили дисперсионный анализ для проверки значимости кажущегося различия между средними. В более общем случае, при проведении анализа по партиям и между партиями, как это делалось в двух предыдущих разделах, может понадобиться вычислить обе компоненты дисперсии. Среднее квадратов по литейным, как и прежде, определяет дисперсию по литейным, т. е. $\sigma_A^2=528,5$. Расчет дисперсии между литейными осложняется, если размеры партий (столбцов) неодинаковы.

Если N — общее число наблюдений, n_i — число наблюдений по каждой партии для $i=1, 2, \ldots, k$, где k — число партий, то среднее квадратов между литейными (между партиями или столбцами) определяется, как

$$\frac{N^2-\sum n_i^2}{N(k-1)}\cdot\sigma_B^2+\sigma_A^2.$$

Заметим, что в случае, когда все партии (столбцы) одинакового размера, т. е. все $n_i=n$, уравнение упрощается [см. гл. VII, (B)]:

$$\frac{(nk)^2-kn^2}{nk(k-1)}\cdot\sigma_B^2+\sigma_A^2=n\cdot\sigma_B^2+\sigma_A^2.$$

В нашем примере величины n_i составляют 5, 9, 3; $N=5+9+3=17$; $k=3$. Поэтому:

$$1164,5=\frac{17^2-(5^2+9^2+3^2)}{17\cdot(3-1)}\sigma_B^2+\sigma_A^2=5,118\sigma_B^2+\sigma_A^2.$$

Вычитая $\sigma_A^2=528,5$, получим $\sigma_B^2=124,3$.

Следует подчеркнуть еще раз, что подсчет компонент дисперсии здесь сделан только ради иллюстрации метода расчета. Практически для данной задачи этого не стали бы делать, потому что, во-первых, мы более заинтересованы в величинах средних, и, во-вторых, эффект настолько незначим, что его расчет не имел бы смысла, так как компоненты по настоящим данным неотличимы от нуля. Кроме того, нежелательно применять оценку дисперсии между партиями (столбцами) при наличии только трех партий; вычисленная при этом дисперсия имела бы слишком большую ошибку.

Наконец, здесь мы имеем такой пример, когда неразумно было бы предполагать существование бесконечной совокупности литейных, из которых данные три являются выборкой. Этот вопрос рассмотрен в дальнейшем, в главе XII, (B).

(E) РАЗЛОЖЕНИЕ ДИСПЕРСИИ НА СОСТАВЛЯЮЩИЕ ПО СТРОКАМ И СТОЛБЦАМ И ОСТАТОЧНУЮ ДИСПЕРСИЮ

Часто индивидуумы могут быть распределены по категориям не только по партиям (столбцам), но одновременно и по строкам. Так, в предыдущем примере первый индивидуум в каждом столбце (каждой партии) прошел один из станков, все вторые индивидуумы — другой из станков, все третьи индивидуумы — третий станок. Таким образом, каждая строка соответствует определенному станку. Между этими станками может существовать систематическое различие, и мы желаем проверить, существует ли это различие в действительности, и, в случае положительного ответа, оценить его влияние на общую величину дисперсии конечного продукта.

Формально будем иметь дело с двойной классификацией; каждый индивидуум принадлежит к заданной строке и заданному столбцу.

Другим примером двойной классификации является выпуск продукции с различных станков различными сменами рабочих. Третий пример — в экспериментах по нитрации целлюлозы могли применяться n_1 различных концентраций кислоты (соответствуют строкам) и n_2 различных значений температуры кислоты (соответствуют столбцам).

Возвращаясь к примеру раздела (В) с партиями продукта, разбиваемыми на три более мелкие партии, подсчитываем следующие величины:

(1) Сумма квадратов всех индивидуумов (это делалось и прежде — см. величину (1) в предыдущем анализе).

(2) Сумма квадратов итогов по строкам, деленная на число индивидуумов в каждой строке (это — новая величина, в данном примере равная 8,4).

(3) Сумма квадратов итогов по столбцам, деленная на число индивидуумов в каждом столбце (то же, что и величина (2) в предыдущем анализе).

(4) Квадрат общего итога по всем индивидуумам, деленный на их общее число (величина (3) в предыдущем анализе).

Затем составляем таблицу дисперсионного анализа:

Источники дисперсии	Суммы квадратов	Числа степеней свободы	Средние квадратов	Компоненты дисперсии
Между строками	$(2)-(4)=8,40-0,13$	$n_1-1=2$	4,13	$n_2\sigma_1^2+\sigma_0^2$
Между столбцами	$(3)-(4)=160,67-0,13$	$n_2-1=9$	17,84	$n_1\sigma_2^2+\sigma_0^2$
Остаточная	$(1)+(4)-(2)-(3)$	$(n_1-1)(n_2-1)=18$	1,94	σ_0^2
Итого	$(1)-(4)=203,87$	$n_1n_2-1=29$		

n_1 — число строк $=3$,

n_2 — число столбцов $=10$,

σ_1^2 — дисперсия из-за различий между строками,

σ_2^2 — дисперсия из-за различий между столбцами,

σ_0^2 — остаточная дисперсия.

Остаточная сумма квадратов может быть получена так, как указано в таблице; на практике легче подсчитать суммы квадратов по строкам и столбцам и найти остаток как разность между их суммой и итоговой суммой квадратов.

Для проверки значимости значения среднего квадратов между строками сравниваем эту величину с остатком с помощью отношения дисперсий. В данном случае отношение дисперсий составляет $\frac{4,13}{1,94}=2,12$ при числах степеней свободы $n_1=2$; $n_2=18$. Обращаясь к табл. III в приложении, находим, что это отношение не является значимым. Поэтому наша гипотеза об отсутствии систематического различия между строками (станками) достаточно согласуется с данными. Если величина среднего квадратов между строками оказалась бы значимой, можно было бы приступить к определению величины σ_1^2, т. е. дисперсии из-за различий между строками (станками), из системы уравнений: $n_2\sigma_1^2+\sigma_0^2=4,13$, $\sigma_0^2=1,94$; откуда $\sigma_1^2=\frac{4,13-1,94}{n_2}$.

ГЛАВА VIII

ДИАГРАММА КОНТРОЛЯ КАЧЕСТВА

(А) ВВЕДЕНИЕ

Диаграмма контроля качества получила в последние годы широкое применение как средство контроля качества продукции, особенно в производственных процессах легкой промышленности, как-то: в работе автоматических станков, прессов и т. п. До сих пор она не получила широкого применения в химической промышленности, хотя в некоторых производственных процессах применение ее, повидимому, целесообразно. Здесь диаграмма контроля качества будет рассмотрена только применительно к исследовательской работе*.

Диаграмма контроля качества предназначается для обнаружения присутствия „определимых причин дисперсии" в ряде индивидуумов. Для этого выбираются „рациональные подгруппы" (т. е. группы индивидуумов, которые должны быть более сходны друг с другом, чем генеральная совокупность, как-то: индивидуумы, полученные от данного станка или данным рабочим, полученные за данный день или от данной партии сырья, и т. п.); из рациональных подгрупп получают меру дисперсии в совокупности; затем исследуют, отличаются ли средние подгрупп больше,

* Литература о промышленном контроле качества сейчас весьма обширна: см. руководство W. A. Shewhart, Economic Control of Quality of Manufactured Product, 1931; краткие отчеты British Standard 600R „Quality Control Charts", B. P. Duddling & W. J. Jennet, 1942; British Standart 1008 „Quality Control": A First Guide to Quality Control for Engineers, E. H. Sealy, H. M. S. O., 1946. Более полное изложение: L. E. Simon, An Engineer's Manual of Statistical Methods, 1941. Изложение методов контроля качества в применении к стандартным химическим анализам дает H. E. MacColl, *Chemistry and Industry*, Dec. 9th 1944, p. 418.

чем следует на основании предположения, что все индивидуумы взяты из одной и той же совокупности (иначе говоря, проверяют, что нет значимой разницы между средними подгрупп). Очевидно, что результат этой проверки будет тем же, что и от дисперсионного анализа между столбцами и по столбцам.

В применяемой на практике диаграмме контроля качества оценка дисперсии в пределах подгрупп обычно выводится не из дисперсии, а из средней широты в пределах подгрупп; при этом имеются коэффициенты для пересчета средней широты в величину стандартного отклонения (см. раздел (E)). По получении величины σ (стандартного отклонения в пределах подгрупп), можно утверждать, что средние выборок объема n будут иметь распределение со стандартным отклонением $\frac{\sigma}{\sqrt{n}}$. Таким образом, 95% будут находиться в пределах $1{,}96\frac{\sigma}{\sqrt{n}}$ по обе стороны от среднего $\overline{X}$, принимаемого равным среднему совокупности. Это означает, что 1 из 40 находится ниже $\overline{X}-1{,}96\frac{\sigma}{\sqrt{n}}$ и 1 из 40 — выше $\overline{X}+1{,}96\frac{\sigma}{\sqrt{n}}$. Эти числа служат в диаграммах контроля качества „внутренними“ пределами; значения $3{,}09\frac{\sigma}{\sqrt{n}}$ служат „внешними“ пределами; последние указывают, что 99,8% не выходят за эти пределы, т. е. что 1 из 1000 лежит за нижним пределом и 1 из 1000 — за верхним пределом.

Если среднее какой-либо выборки лежит вне этих пределов, вероятно, что исходные допущения не оправдываются. Этими исходными допущениями является, конечно, то, что выборка взята из совокупности со стандартным отклонением σ и генеральным средним $\overline{X}$. Первое допущение проверяется в контрольной диаграмме на широтах подгрупп (см. следующий раздел); поэтому вероятно, что среднее выборки не отличается значимо от величины $\overline{X}$.

(B) ИЗМЕНЧИВОСТЬ В ПРЕДЕЛАХ ПАРТИИ. КОНТРОЛЬНАЯ ДИАГРАММА ДЛЯ ШИРОТЫ

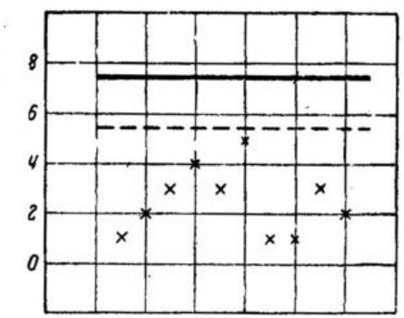
Обратимся к данным табл. 7.1, глава VII, (B). Проверим гипотезу, что изменчивости в пределах каждого столбца (партии) принадлежат к однородной совокупности; другими словами — что партии не отличаются по их внутренней изменчивости (см. табл. 8.1).

Подсчитываем широту (диапазон) w для каждой партии и наносим результаты на диаграмму (см. фиг. 6). Затем подсчитываем величину средней широты $\overline{w}$, в данном случае равную 2,50, и наносим на диаграмму внутреннюю контрольную линию $D'_{0,975}\overline{w}=2{,}17\cdot 2{,}50=5{,}4$ и внешнюю контрольную линию $D'_{0,999}\overline{w}=2{,}98\cdot 2{,}50=7{,}5$. Коэффициенты D берутся из таблицы V (в приложении) для данного объема выборки — 3. Если партии сходны по внутренней изменчивости, то в среднем не более 1 из 40 точек будут лежать вне внутренней контрольной линии и только 1 из 1000 точек — за внешней контрольной линией. Выпадение двух точек за внутреннюю контрольную линию (в пределах двух или трех точек) или одной точки за внешнюю контрольную линию поэтому явится веским свидетельством того, что партии нельзя считать имеющими одну и ту же внутреннюю изменчивость.

ТАБЛИЦА 8.1

Данные	−3 −2 −3	−4 −3 −5	−3 −1 −4	−1 3 2	4 3 6	−2 3 1	1 0 1	2 1 1	−1 −1 2	−1 1 1
$\overline{x}$	−2,67	−4,00	−2,67	1,33	4,33	0,67	0,67	1,33	0	0,33
w	1	2	3	4	3	5	1	1	3	2
$\sum x$	−8	−12	−8	4	13	2	2	4	0	1
$\sum x^2$	22	50	26	14	61	14	2	6	6	3
$\sum(x-\overline{x})^2$	0,67	2,00	4,67	8,67	4,67	12,67	0,67	0,67	6,00	2,67
σ_i^2	0,33	1,00	2,33	4,33	2,33	6,33	0,33	0,33	3,00	1,33

Контрольная диаграмма для диапазона



Фиг. 6

Из фиг. 6 видно, что в рассматриваемом случае нет вообще точек, лежащих за контрольными линиями, так что нет причины предполагать, что партии отличаются друг от друга по внутренней изменчивости.

(C) СРАВНЕНИЕ КОНТРОЛЬНОЙ ДИАГРАММЫ ДЛЯ ШИРОТЫ С ПРОВЕРКОЙ С ПОМОЩЬЮ КРИТЕРИЯ БАРТЛЕТА

Отвлекаясь от темы, заметим, что с помощью контрольной диаграммы для широты проверяется та же гипотеза, что и с помощью критерия Бартлета (см. гл. IV, (C)).

Применяя последний, прежде всего подсчитывают дисперсию каждой партии σ_i^2 вышеуказанным путем, т. е. определяя $\sum x$, $\sum x^2$ и $\sum(x-\overline{x})^2$. Сумма натуральных логарифмов дисперсий σ_i^2 в пределах партий равна: $\sum \ln \sigma_i^2 = 3{,}06039$. Число степеней свободы для индивидуальных дисперсий $n=2$. Среднее дисперсий по партиям $S^2=2{,}164$; число сравниваемых дисперсий $k=10$. Отсюда получается:

$$B=kn\ln S^2-n\sum\ln\sigma_i^2=10\cdot 2\cdot\ln 2{,}164-2\cdot 3{,}06039=9{,}31862;$$
$$C=1+\frac{k+1}{3nk}=1+\frac{10+1}{3\cdot 2\cdot 10}=1{,}183.$$

Отношение $\frac{B}{C}=\frac{9{,}319}{1{,}183}=7{,}87$ имеет распределение, как величина χ^2 при $k-1=9$ степенях свободы. Оно намного меньше значения, соответствующего уровню значимости в 10%, так что нет причин предполагать, что партии отличаются одна от другой по внутренней изменчивости.

Проверка с помощью критерия Бартлета приводит к тому же результату, что и контрольная диаграмма для широты, но требует большего объема расчетов; с другой стороны, ее преимуществами являются большая чувствительность и точность.

(D) ИЗМЕНЧИВОСТЬ МЕЖДУ ПАРТИЯМИ. КОНТРОЛЬНАЯ ДИАГРАММА ДЛЯ СРЕДНИХ

Мы хотим установить, отличаются ли партии значимо по их средним, учитывая наблюденную величину изменчивости индивидуумов.

Подсчитываем среднее по каждой партии (первая строка в нижней части табл. 8.1) и наносим их на диаграмму (фиг. 7).

Фиг. 7

Подсчитываем среднее $\bar{X}=-0{,}67$ и проводим внутренние контрольные линии $\bar{X}\pm A'_{0,025}\bar{w}=-0{,}67\pm 0{,}668\cdot 2{,}5=1{,}60$ и $-1{,}74$ и внешние контрольные линии $\bar{X}\pm A'_{0,001}\bar{w}=-0{,}67\pm 1{,}054\cdot 2{,}5=2{,}57$ и $-2{,}70$. Если изменчивость между партиями не больше той, которая вызывается изменчивостью в пределах партий, одна точка из 20 будет лежать вне внутренних контрольных линий и две из 1000 — за внешними контрольными линиями. В данном случае четыре из 10 точек находятся вне внутренних контрольных линий и две из 10 — за внешними контрольными линиями. Поэтому очевидно, что гипотеза об отсутствии изменчивости между партиями несостоятельна.

Этот вывод вполне согласуется с достигнутым с помощью дисперсионного анализа в главе VII, (В). При применении контрольной диаграммы требуется намного меньше арифметической

работы, но дисперсионный анализ имеет преимущество в том, что он дает числовую оценку относительной важности изменчивости между партиями и внутри партий, а также значимости утверждения о существовании изменчивости между партиями. Поэтому для детальной и точной работы предпочтительнее дисперсионный анализ; контрольная же диаграмма весьма полезна для предварительных исследований.

Важным предварительным условием при применении контрольных диаграмм должно быть наличие не менее 10, а лучше не менее 20 партий; в противном случае оценка средней широты $\bar{w}$ будет не очень точной.

(Е) ПЕРЕХОД ОТ ШИРОТЫ К СТАНДАРТНОМУ ОТКЛОНЕНИЮ

Часто желают узнать стандартное отклонение ряда выборок из n индивидуумов, взятых из некоторой совокупности. Можно, конечно, подсчитать дисперсию для каждой выборки и взять среднее значение (поскольку все выборки состоят из n индивидуумов, простое среднее будет тем же, что и подсчитанное из сумм квадратов с учетом степеней свободы).

Однако потребуется гораздо меньше вычислений, если подсчитать величину широты для каждой выборки, вычислить ее среднюю величину и разделить ее на коэффициент d_n, значения которого приведены в последнем столбце табл. V в приложении; значение d_n выбирают в соответствии с объемом наличной выборки. Полученный результат будет стандартным отклонением. Таким образом, стандартное отклонение получается лишь приближенно; им можно пользоваться, когда число выборок достаточно велико, например не меньше 20. При этом считается, что эта оценка соответствует бесконечному числу степеней свободы.

За исключением случая, когда объем выборки равен 2, мы несколько жертвуем точностью в оценке σ; компенсацией служит облегчение расчета. Однако потеря информации не очень серьезна — менее 20% для выборок размером 10 и еще меньше для меньших выборок. При условиях, в которых применяется данный метод, это не имеет особенного значения, поскольку для больших выборок наша оценка будет достаточно точной почти для всех целей.

Читатель заметит, что именно на d_n основан выбор коэффициентов $A'_{0,025}$ и $A'_{0,001}$, используемых в построении контрольных линий для средних. Поэтому ясно, что для выборки объема n контрольные линии для пределов 1 из 40 (1 величина из 40 за верхним пределом и 1 из 40 за нижним пределом; всего 1 из 20 вне пределов) должны соответствовать значению $1{,}96\frac{\sigma}{\sqrt{n}}$, где σ — стандартное отклонение, 1,96 — величина t для бесконечного числа степеней свободы (считается, что величина σ точно известна, так как мы берем большое число выборок, предпочтительно не менее 20) и для уровня вероятности в 5%, n — объем выборки. Если, например, $n=4$, а $\overline{w}$ — величина средней широты, то пределы должны быть при $1{,}96\,\frac{\overline{w}}{2{,}059}\cdot\frac{1}{\sqrt{4}}=0{,}476\,\overline{w}$, где 2,059 — значение d_n при $n=4$. Можно видеть, что 0,476 и является значением $A'_{0{,}025}$, приведенным в табл. V в приложении.

ГЛАВА IX

СВЯЗИ МЕЖДУ ДВУМЯ ПЕРЕМЕННЫМИ

(A) ВВЕДЕНИЕ *

Часто возникает вопрос, является ли видимая связь между двумя переменными значимой, и если эта связь оказывается значимой, то какова наилучшая форма ее представления.

Имеющиеся для решения этого вопроса статистические методы не вполне удовлетворительны. Общий прием состоит в проверке того, могут ли данные быть представлены уравнением прямой. $y=a+bx$, значимы ли отклонения от прямолинейного закона, и, при утвердительном ответе на последний вопрос, могут ли эти отклонения быть выражены уравнением типа $y=a+bx+cx^2$ или $y=a+bx+cx^2+dx^3$, и т. д. Практическая работа по подсчетам чрезвычайно трудоемка на всех этапах, кроме первого, т. е. определения, могут ли данные быть приближенно связаны линейной зависимостью $y=a+bx$**.

Неудовлетворительность такого приема объясняется тем, что не существует причины *a priori*, почему сложная система должна определяться линейным соотношением между двумя из ее переменных, или даже более сложным законом типа $y=a+bx+cx^2$ и т. д. Прием более всего подходит к такому случаю, когда данные, нанесенные на график, представляют „эллиптическое облако“, и возникает сомнение, отражает ли эта форма действительную зависимость или случайности выборки. В другом случае может не возникать сомнений в существовании зависимости и желают лишь получить наилучшую оценку последней.

* Исчерпывающее изложение этих вопросов см. у W. E. Deming, Statistical Adjustment of Data, 1943.

** Исключением является случай, когда значения независимого переменного — равноотстоящие; см. R. A. Fisher, Statistical Methods for Research Workers, 27, 28, а также G. Egloff and R. C. Kuder, *Journ. of Phys. Chemistry*, **46** (1942), p. 926, L. H. C. Tippet, Methods of Statistics, 9, 3.

Если данные, нанесенные на график, лежат на плавной кривой сложной формы, то рассматриваемый прием неприменим. Поэтому в практике следует применять настоящий статистический прием только в случае, когда пары наблюдений над двумя переменными по нанесении на график располагаются приблизительно прямолинейно.

(B) ПРЕОБРАЗОВАНИЕ ПЕРЕМЕННЫХ

Теория рассматриваемого процесса может заставить предполагать *a priori*, что соотношение между переменными должно иметь вид

$$Y=A\cdot B^x.$$

Путем логарифмирования эта зависимость преобразовывается в линейную

$$\log Y=\log A+x\log B.$$

Заменяя $\log Y=y$, $\log A=a$, $\log B=b$, получаем

$$y=a+bx.$$

В другом случае, когда подозревают, что соотношение имеет вид

$$Y=AX^B,$$

также прибегают к логарифмированию:

$$\log Y=\log A+B\log X.$$

Заменяя $\log Y=y$, $\log A=a$, $\log X=x$, получим

$$y=a+Bx.$$

(C) КОЭФФИЦИЕНТ КОРРЕЛЯЦИИ

Для проверки значимости предполагаемого линейного соотношения подсчитывается коэффициент корреляции r:

$$r=\frac{\Sigma(x-\bar{x})(y-\bar{y})}{\sqrt{\Sigma(x-\bar{x})^2\,\Sigma(y-\bar{y})^2}},$$

где знак Σ обозначает суммирование по всем парам наблюдений. В случае, когда соотношение между данными может быть точно представлено прямой линией, $r=\pm 1$; знак плюс соответствует положительному наклону прямой, знак минус — отрицательному. Если же связи между переменными вообще не существует, то $r=0$. Однако даже при отсутствии связи между переменными обычно получается значение r, отличное от нуля, поскольку случайность выборки может вызвать те или иные отклонения.

Тем не менее, можно проверить, является ли наблюденное значение r бóльшим, чем то, которое получилось бы случайно, при отсутствии корреляции. Обращаемся к табл. IV значений r (в приложении), беря число степеней свободы на два меньше, чем число пар наблюдений (причина, что оно на *два* меньше числа пар наблюдений, лежит в том, что учитываемые данные связаны соотношением с двумя постоянными a и b). Если наше значение r превосходит данное в таблице, например для уровня значимости в 5%, то мы имеем 19 шансов против 1, что корреляция действительно имеет место.

В качестве иллюстрации приведем статистическую обработку данных, которые предположительно связаны зависимостью вида $y=a+bx$. Было сделано 30 наблюдений; часть данных представлена следующей таблицей:

y	1,66	2,0	1,66	1,78	1,89	1,82	...	2,07	2,29
x	6	8	17	16	19	16	...	13	12

Величина y представляет собой вес патрона кордита, сделанного из нитроцеллюлозы с ацетоном вязкости x.

Для облегчения арифметических вычислений переменные были преобразованы; величина y в действительности является весом патрона в унциях (1 унция $=28{,}3$ г) минус 13 фунтов 1 унция (5,917 кг), а величина x — вязкость в сантипуазах минус 50 $\left(1 \text{ пуаз} = \frac{\text{дина} \cdot \text{сек.}}{\text{см}^2}\right)$.

Отметив данные x и y на чертеже, получим график с большим рассеиванием, но с определенным указанием на то, что более высокие значения вязкости вызывают снижение веса патронов.

Подсчитываем следующие выражения:

$$\sum(x-\bar{x})^2=\sum x^2-\frac{(\sum x)^2}{n}=(6^2+8^2+\ldots+12^2)-$$
$$-\frac{(6+8+\ldots+12)^2}{30}=3213,$$

$$\sum(y-\bar{y})^2=\sum y^2-\frac{(\sum y)^2}{n}=(1{,}66^2+2^2+\ldots+2{,}29^2)-$$
$$-\frac{(1{,}66+2+\ldots+2{,}29)^2}{30}=9{,}060,$$

$$\sum(x-\bar{x})(y-\bar{y})=\sum xy-\frac{(\sum x)(\sum y)}{n}=(1{,}66\cdot 6+2\cdot 8+\ldots)-$$
$$-\frac{(1{,}66+2+\ldots)(6+8+\ldots)}{30}=-47{,}84.$$

Определяем коэффициент корреляции:

$$r=\frac{\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\sum(x-\bar{x})^2\sum(y-\bar{y})^2}}=\frac{-47{,}84}{\sqrt{9{,}060\cdot 3213}}=-0{,}28.$$

Из таблицы r (табл. IV в приложении) видим, что вероятность получения такого значения r в отсутствии какой-либо корреляции составляет около 0,10, т. е. 1 раз из 10 мы могли бы получить столь же большое значение r даже в отсутствии корреляции. Доказательство наличия корреляции, таким образом, недостаточно.

(D) УРАВНЕНИЕ ЛИНИИ РЕГРЕССИИ

Если значение r окажется значимо бо́льшим нуля, можно рассмотреть вопрос несколько глубже; хотя в рассмотренном примере r и не является значимо бо́льшим нуля, мы используем его данные для иллюстрации вопроса.

Желая использовать данные для предсказания значений y, соответствующих заданным значениям x, подсчитываем:

$$a=\frac{\sum y}{n}=\bar{y}=\frac{49{,}58}{30}=1{,}65,$$

$$b=\frac{\sum(y-\bar{y})(x-\bar{x})}{\sum(x-\bar{x})^2}=\frac{-47{,}84}{3213}=-0{,}0149,$$

$$\bar{x}=\frac{\sum x}{n}=\frac{670}{30}=22{,}3.$$

Формула для вывода наиболее вероятного значения y, соответствующего заданному значению x (так называемая регрессия y на x), тогда имеет вид:

$$y=a+b(x-\bar{x})=1{,}65-0{,}0149(x-22{,}3)=1{,}98-0{,}0149\ x.$$

Эта линия регрессии y на x является линией, дающей минимум (для всех точек) суммы квадратов отклонений, измеренных параллельно оси y в масштабе y. Это дает лучшую оценку для y по заданному значению x.

Если желают предсказать значение x по заданному значению y, используют подобное же уравнение, вычислив

$$a'=\frac{\sum x}{n},\quad b'=\frac{\sum(y-\bar{y})(x-\bar{x})}{\sum(y-\bar{y})^2},\quad \bar{y}=\frac{\sum y}{n},$$

$$x=a'+b'(y-\bar{y}).$$

Эта линия дает минимум суммы квадратов отклонений в каждой точке, измеренных параллельно оси x в масштабе x. Она дает лучшую оценку для x по заданному значению y.

Не-статистики часто возражают против двух „лучших" линий, т. е. двух линий регрессии. Они считают, что должна существовать одна линия, представляющая „истинное" соотношение. Этот взгляд может быть правильным, но мы не касаемся „истинных" соотношений; мы стремимся сделать наилучшую возможную оценку величины y по заданному значению x, используя определенное количество экспериментальных данных для выяснения характера соотношения. В одном случае мы считаем известным x и желаем оценить y, в другом случае полагаем известным y и желаем оценить x. Это — две различные операции, поэтому неудивительно, что в них применяются две различные функции.

(E) ОСТАТОЧНАЯ ДИСПЕРСИЯ ОТНОСИТЕЛЬНО ЛИНИИ РЕГРЕССИИ

Допустим, что имеется линия регрессии y на x и точки, разбросанные по обе стороны от нее. Величина стандартного отклонения σ_r этого разброса, измеренная в масштабе y параллельно оси y, дается уравнением (числовые данные взяты из предыдущего примера):

$$\sigma_r = \sqrt{1-r^2}\sqrt{\frac{\sum(y-\bar{y})^2}{N-2}} = \sqrt{1-0{,}280^2}\cdot\sqrt{\frac{9{,}06}{28}} = 0{,}546.$$

Это число представляет собой меру полезности уравнения линии регрессии при предсказании значений y по заданным значениям x.

Предположим, что желательно установить доверительные пределы в 95% по обе стороны линии регрессии, в которых должны лежать 95% всех точек; берем значение t для $N-2=28$ степеней свободы и уровня значимости в 5% и подсчитываем величину $t\cdot\sigma_r = 2{,}05\cdot 0{,}546 = 1{,}12$. Затем проводим две линии, параллельных линии регрессии; одна из них смещена на 1,12 единиц y вниз, а другая — на 1,12 единиц y вверх. Теперь, если использовать линию регрессии для предсказания значения y по данному значению x, это будут доверительные пределы, между которыми с вероятностью 95% правильности нашего предсказания, должно находиться значение y *.

(F) ПРИМЕНЕНИЕ ДИСПЕРСИОННОГО АНАЛИЗА К ИССЛЕДОВАНИЮ РЕГРЕССИИ

Другим возможным методом проверки значимости регрессии является разложение дисперсии зависимого переменного y на составляющие факторы.

Общая сумма квадратов для величины y равна $\sum(y-\bar{y})^2 = 9{,}060$. Сумма квадратов, учитываемых линией регрессии, составляет

$$\frac{[\sum(y-\bar{y})(x-\bar{x})]^2}{\sum(x-\bar{x})^2} = \frac{47{,}84^2}{3213} = 0{,}7123.$$

* Это верно лишь приблизительно; уточнение см. в разделе (H).

Вносим эти цифры в таблицу дисперсионного анализа:

Источники дисперсии	Сумма квадратов	Числа степеней свободы	Средние квадратов
Регрессия	0,7123	1	0,7123
Остаток	8,3477	28	0,2981
Итого	9,06	29	—

Число степеней свободы равно 1 для линии регрессии, общему числу пар наблюдений минус 1 — для итога; разность их будет числом степеней свободы для остатка. Аналогично, сумма квадратов для остатка вычисляется как разность между итоговым значением и значением для линии регрессии.

Значимость линии регрессии оценивается путем обычного сравнения ее среднего квадратов с остаточным значением. В данном случае она лежит между 20 и 5%-ными уровнями значимости.

Остаток — таков же, как и остаточная дисперсия относительно линии регрессии, подсчитанная в предыдущем разделе. Поэтому квадратный корень из остаточной дисперсии составляет $\sqrt{0{,}2981} = 0{,}546$, как подсчитано и ранее.

Заметим, что результаты, полученные рассматриваемым путем, тождественны с найденным путем подсчета коэффициента корреляции.

(G) СРАВНЕНИЕ КОЭФФИЦИЕНТОВ РЕГРЕССИИ

Может понадобиться установить, значимо ли отличаются друг от друга два коэффициента регрессии. Допустим, что проведены эксперименты над некоторой системой: (a) при обходе по часовой стрелке, (b) при обходе против часовой стрелки. Для обеих серий экспериментов вычерчены графики, и возникло подозрение, что результаты различаются.

Можно сделать проверку, пользуясь тем, что дисперсия коэффициента регрессии равна остаточной дисперсии величины y, деленной на $\sum(x-\bar{x})^2$, т. е. в данном случае $\frac{0{,}2981}{3213}$. Квадратный корень из этого числа, равный 0,0096, дает стандартное отклонение коэффициента регрессии; число степеней свободы на 2 меньше общего числа наблюдений, т. е. равно 28.

Если два коэффициента регрессии b_1 и b_2 имеют дисперсии σ_1^2 и σ_2^2, подсчитываем „Стьюдентово t" в виде

$$t = \frac{b_1 - b_2}{\sqrt{\sigma_1^2+\sigma_2^2}}.$$

Число степеней свободы для этой величины t равно сумме чисел степеней свободы обеих индивидуальных дисперсий.

(H) ТОЧНАЯ ФОРМУЛА ДЛЯ ОСТАТОЧНОЙ ДИСПЕРСИИ ОТНОСИТЕЛЬНО ЛИНИИ РЕГРЕССИИ

Как указано ранее, формула для остаточной дисперсии относительно линии регрессии, выведенная в разделе (E), не вполне точна. Если σ_r^2 — остаточная дисперсия, определенная выше, и σ_b^2 — дисперсия самого коэффициента регрессии, вычисленная, как указано в разделе (G), то точная формула для σ_{r1}^2 будет следующей:

$$\sigma_{r1}^2 = \sigma_r^2 + X^2\sigma_b^2,$$

где X — расстояние в единицах x от $\bar{x}$, средней точки x.

Например, в рассмотренном случае (раздел (D)), величина $\bar{x}$ равна 22,3. В разделе (C) было подсчитано: $\sum(x-\bar{x})^2 = 3213$; откуда $\sigma_x^2 = 110{,}79$. Величина σ_x получилась порядка 10; поэтому, для выбора точки в сторону одного из концов шкалы x, можно взять $\bar{x} + 2\sigma_x = 22{,}3 + 2 \times 10 = 42{,}3$; это соответствует $X = 20$. Подставляя значения σ_r и σ_b, получаем:

$$\sigma_{r1}^2 = 0{,}2981 + 20^2 \times 0{,}00009278 = 0{,}3352,$$

откуда

$$\sigma_{r1} = 0{,}579.$$

В данном случае применение приближенной формулы дало результат, близкий к точному; разумеется, чем ближе мы подходим к средней точке x, тем меньше значение X и тем менее существенно внесение поправки.

(I) ПРИМЕНЕНИЕ ДИСПЕРСИОННОГО АНАЛИЗА К ПРОВЕРКЕ ЛИНЕЙНОСТИ

Нередко возникает вопрос: можно ли сказать, что группа наблюдений, которая в первом приближении могла быть представлена прямолинейной зависимостью, значимо отличается от последней. Этот вопрос легко разрешается, если данные имеют симметричную форму, удобную для приложения дисперсионного анализа.

Рассмотрим следующие данные:

x	10	20	30	40
y	92,8 93,0	94,0 94,3	95,1 94,8	94,9 94,8
$\bar{y}$	92,9	94,15	94,95	94,85

Величина x может рассматриваться как значение температуры в градусах Цельсия в определенной стадии производственного процесса; два значения y указывают количество продукции, получившейся при данной температуре в двух повторных экспериментах. Интересно выяснить влияние температуры на выход продукции. В первом приближении кажется, что y растет линейно по мере повышения температуры x; но значение y при $x = 40$ меньше, чем при $x = 30$. Проверим, значимо ли это видимое отклонение от линейности.

Для упрощения работы сделаем замену переменных: $X = \frac{x}{10}$; $Y = (y - 92) \cdot 10$. Таблица примет вид:

X	1	2	3	4
Y	8 10	20 23	31 28	29 28

В первую очередь разложим дисперсию данных на составляющие между столбцами (заданные значения X) и по столбцам.

Подсчитываем расчетные величины:

$$(1) = 8^2 + 10^2 + \ldots + 28^2 = 4463,$$

$$(2) = \frac{1}{2}[(8+10)^2 + (20+23)^2 + (31+28)^2 + (29+28)^2] = 4451{,}5,$$

$$(3) = \frac{1}{8}(8 + 10 + 20 + \ldots + 28)^2 = 3916{,}1.$$

Составляем таблицу:

Источники дисперсии	Сумма квадратов	Числа степеней свободы	Средние квадратов
Между значениями температуры	(2)—(3) = 535,4	3	178,5
По значениям температуры	(1) — (2) = 11,5	4	2,87
Итого	(1)—(3) = 546,9	7	

Сопоставляя среднее квадратов между температурами со средним квадратов по температурам, видим, что различие между ними значимо.

Подразделим теперь суммы квадратов между значениями температуры на составляющие, определяемые линией регрессии и отклонениями от нее. Подсчитываем:

$$\sum X = 2(1+2+3+4) = 20,$$

$$\sum X^2 = 2(1^2+2^2+3^2+4^2) = 60,$$

$$\sum (X-\overline{X})^2 = 60 - \frac{20^2}{4\cdot 2} = 10,$$

$$\sum XY = 1(8+10) + 2(20+23) + 3(3+28) + 4(29+28) = 509,$$

$$\sum (Y-\overline{Y})(X-\overline{X}) = \sum XY - \sum X \frac{\Sigma Y}{N} =$$

$$= 509 - 20 \cdot \frac{177}{8} = 66{,}5.$$

Коэффициент 2 в выражениях для $\sum X$ и $\sum X^2$ учитывает наличие двух наблюдений величины Y для каждого значения X. Сумма квадратов, относящаяся к линии регрессии:

$$\frac{[\Sigma (Y-\overline{Y})(X-\overline{X})]^2}{\Sigma (X-\overline{X})^2} = \frac{66{,}5^2}{10} = 442{,}22.$$

Так как общее значение суммы квадратов, относящейся к различиям в значениях температуры, равно 535,4, то сумма квадратов, определяемая отклонениями от линии регрессии, составляет

$$535{,}4 - 442{,}2 = 93{,}2.$$

Число степеней свободы для линии регрессии равно 1, для отклонений от нее — числу, принадлежащему строке „между значениями температуры“, минус 1.

Вносим вычисленные результаты в таблицу:

Источник дисперсии	Суммы квадратов	Числа степеней свободы	Средние квадратов
Между значениями температуры:			
Регрессия	442,2	1	442,2
Отклонения от регрессии . . .	93,2	2	46,6
По значениям температур (остаток)	11,5	4	2,87
Итого	546,9	7	—

ТАБЛИЦА 9.1

																			Ay	By	Cy	Dy	Ey	F	G
				1															25	1	−9	−9	81	−3	+27
																			24	0	−8	0	0	0	
	1				1														23	2	−7	−14	98	−8	+56
																			22	0	−6	0	0	0	
			1		1		1												21	3	−5	−15	75	−6	+30
		1			1	1													20	3	−4	−12	48	−8	+32
				1	2			1											19	4	−3	−12	36	−6	+18
		1	1	2	2	2	1	1											18	10	−2	−20	40	−20	+40
	1		1	1	1	2	3	2		1									17	12	−1	−12	12	−12	+12
	1		2		3	3	2	1	2	1		1							16	16	0	−94	0	−10	
					3		2	1	3			1	1						15	11	1	11	11	+12	+12
		1	1		2	4	3	1	1	1							1		14	15	2	30	60	−1	−2
				1	1	2	2	1	2	1	1								13	11	3	33	99	+5	+15
			1					2		1		1	1						12	6	4	24	96	+12	+48
						1				1				1					11	3	5	15	75	+9	+45
						1	1					1		1					10	4	6	24	144	+11	+66
Ax	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26			$101 = N$		+137	$875 = \Sigma y^2$		+401
Bx	3	3	7	6	17	16	15	10	8	6	1	4	2	2	0	0	1	$101 = N$				−94		−74	−2
Cx	−6	−5	−4	−3	−2	−1	0	1	2	3	4	5	6	7	8	9	10							+49	
Dx	−18	−15	−28	−18	−34	−16	−129	10	16	18	4	20	12	14	0	0	10	104							
Ex	108	75	112	54	68	16	0	10	32	54	16	100	72	93	0	0	100	$915 = \Sigma x^2$				$\Sigma y = +43$		$\Sigma x = -25$	$+399 = \Sigma xy$

$$\Sigma x = +104 - 129 = -25.$$

Общая значимость линии регрессии может быть проверена путем сравнения значений средних квадратов регрессии и остатка. В данном случае она, очевидно, весьма значима.

Для проверки того, значимо ли отклонение от прямой линии, подсчитываем отношение дисперсий отклонения от регрессии и остатка: $\frac{46{,}6}{2{,}87}=16{,}2$; числа степеней свободы $n_1=2$, $n_2=4$. Это имеет уровень значимости почти в 1%; с этой степенью уверенности можно принять, что данные не могут быть полностью представлены прямой линией.

Наилучшая оценка коэффициента регрессии Y на X дается равенством:

$$b=\frac{\sum(Y-\overline{Y})(X-\overline{X})}{\sum(X-\overline{X})^2}=\frac{66{,}5}{10}=6{,}65.$$

(J) РАСЧЕТ КОЭФФИЦИЕНТА КОРРЕЛЯЦИИ, КОЭФФИЦИЕНТОВ РЕГРЕССИИ И Т. Д. ПО ГРУППИРОВАННЫМ ДАННЫМ

При большом числе пар наблюдений, например 100, расчет коэффициента корреляции по способу, описанному в разделе (С), потребует значительного труда; работа может быть облегчена путем группирования данных по принципу, разобранному в главе II, (J).

Подразделяем данные по каждой оси координат на 10 или 20 одинаковых групп; подсчитываем число наблюдений в каждой ячейке.

Данные в табл. 9.1 относятся к производственному процессу, включающему образование эмульсии в воде. Независимая переменная x является функцией качества воды (жесткости), а зависимой переменной y является качество конечного продукта.

Табл. 9.1 является в сущности графиком, в котором Ax представляет собой ось x, Ay — ось y. Число, соответствующее каждой комбинации значений Ax и Ay, показывает число наблюдений, при которых эта комбинация наблюдалась (незаполненное место указывает, что таких результатов при наблюдениях не получалось). В графе B дана частота, с которой наблюдались соответствующие количества. Так, для $Ax=10$ видно, что имеются точки с абсциссой $Ax=10$ и ординатами $Ay=23$, 17 и 16; соответствующая запись $Bx=3$.

В графах C приведен условный масштаб со смещенным началом координат, введенный для облегчения арифметических расчетов. В графах D и E приведены величины $B\times C$ и $B\times C^2$. Общее число наблюдений, равное сумме частот, записанных в графах B, составляет 101. Таким образом имеем:

$$\sum(x-\overline{x})^2=\sum x^2-\frac{(\sum x)^2}{N}=915-\frac{25^2}{101}=908{,}8.$$

$$\sum(y-\overline{y})^2=\sum y^2-\frac{(\sum y)^2}{N}=875-\frac{43^2}{101}=856{,}7.$$

Подсчет члена $\sum(x-\overline{x})(y-\overline{y})$ более затруднителен. Для каждого значения Cy умножаем частоту каждой ячейки на ее значение Cx и затем суммируем полученные числа для всех ячеек, соответствующих данному значению Cy. Результат записывается в графу F. Так, для $Cy=-3$ заполнены три ячейки, первая — с единичной частотой при $Cx=-3$, вторая — с частотой 2 при $Cx=-2$, третья — с единичной частотой при $Cx=1$. Таким образом, $F=1\cdot(-3)+2\cdot(-2)+1\cdot1=-6$.

Для проверки точности значений, записанных в графу F, следует иметь в виду, что сумма всех значений должна быть равна $\sum x$. Произведения Cy и F записаны в графу G; сумма ее значений равна $\sum xy$. Чтобы иметь отклонения от средних, а не от произвольных нулевых точек, вычитаем корректирующую величину:

$$\frac{\sum x\sum y}{N}=\frac{(-25)\cdot(+43)}{101}=-10{,}64.$$

Итак, получаем:

$$\sum(x-\overline{x})(y-\overline{y})=399{,}0-(-10{,}6)=409{,}6.$$

Коэффициент корреляции равен:

$$R=\frac{\sum(x-\overline{x})(y-\overline{y})}{\sqrt{\sum(x-\overline{x})^2\sum(y-\overline{y})^2}}=\frac{409{,}6}{\sqrt{908{,}8\cdot856{,}7}}=0{,}466.$$

Для 99 степеней свободы его значимость превосходит уровень в 0,1%. Коэффициент регрессии y на x можно рассчитать, как и прежде:

$$b=\frac{\sum(x-\overline{x})(y-\overline{y})}{\sum(x-\overline{x})^2}=\frac{409{,}6}{908{,}8}=0{,}451.$$

Уравнение регрессии:

$$y=\frac{\sum y}{N}+b\left(x-\frac{\sum x}{N}\right)=0{,}451\,x+0{,}537.$$

(K) КОРРЕЛЯЦИЯ И ПРИЧИННАЯ ЗАВИСИМОСТЬ

Значимость коэффициента корреляции указывает на связь между двумя переменными: важно, однако, иметь в виду, что отсюда не следует прямо наличие причинной зависимости.

Пусть, например, мы располагаем данными о заводе, охватывающими период в несколько месяцев; из них находим, что температура T_1 на определенной стадии производственного процесса изменяется довольно значительно. Мы можем нанести на график величину T_1 и зависимую переменную y, получить нечто сходное с линейным соотношением, рассчитать коэффициент корреляции и найти его значимым.

На этой стадии было бы опрометчиво допустить, что y непременно является функцией T_1, так как может случиться, что действительно воздействующей переменной была температура T_2 в какой-то иной стадии производственного процесса, и что T_1 была связана с T_2 некоторым общим фактором (например, погодой).

Такого типа ошибка может всегда произойти, когда данными являются показатели работы действующего предприятия. Ошибка менее вероятна, если данные получены из искусственно поставленного эксперимента. Так, если в эксперименте мы повышаем температуру T_1 до избранного значения и работа завода улучшилась, судя по значению зависимой переменной y (выпуск продукции, качество и т. п.), то для практических надобностей может быть достаточно сказать: „ведите работу при этом значении T_1“. Однако это еще не может служить доказательством причинного влияния T_1 на y, так как условия завода могут быть таковы, что повышение T_1 до нового значения одновременно отразилось на действительно воздействующую переменную T_2. Например, в процессе, действующем на основе принципа встречного движения, повышение температуры в одном месте вероятно отразится на температуре в других местах.

Соображения, приведенные в последних трех разделах, никоим образом не обесценивают желательность проверки значимости любого кажущегося соотношения. Они только указывают, что, если корреляция и значима, необходима осторожность в выводах о наличии причинного влияния одной переменной на другую.

(L) ЗАКЛЮЧЕНИЕ

Когда подозревают, что соотношение между двумя переменными приблизительно линейно, имеются в распоряжении три статистические величины, которые отражают наиболее важные свойства имеющихся данных:

(a) Коэффициент корреляции и соответствующее число степеней свободы являются удобной мерой степени прочности связи.

(b) Коэффициент регрессии b — мера наклона линии регрессии, т. е. среднего приращения величины y на единицу приращения величины x.

(c) Остаточная дисперсия σ_r — мера разброса значений y около линии регрессии, т. е. мера надежности оценки величины y для заданных значений x по линии регрессии.

(d) Если эксперимент был проведен с соответствующими повторениями, часто желательно применить дисперсионный анализ для оценки значимости отступления от линейного закона.

При применении уравнения регрессии следует всегда помнить, что оно действительно только в диапазоне независимой переменной, включающем данные, использованные в расчете. Крайне неблагоразумно было бы допускать экстраполяцию, кроме случаев, когда для такой экстраполяции имеется серьезное теоретическое обоснование. Для того чтобы предотвратить экстраполяцию, всегда рекомендуется указывать обследованный диапазон независимой переменной *.

ГЛАВА X

МНОЖЕСТВЕННАЯ КОРРЕЛЯЦИЯ

(A) ВВЕДЕНИЕ

В главе IX рассмотрено определение соотношения между зависимой и одной независимой переменными. Во многих случаях, однако, имеется более, чем одна, независимых переменных, представляющих практическую важность. В таких случаях необходимо изолировать эффект каждой независимой переменной от эффектов других.

В нижеследующем примере мы рассмотрим потери в работающей под давлением окислительной установке для производства азотной кислоты (около 55—60%) путем окисления аммиака воздухом. Азотные газы абсорбируются при пропускании вверх по поглотительной башне, азотная же кислота циркулирует вниз. Процесс абсорбции является экзотермическим, т. е. сопровождается выделением тепла, так что температура стремится возрасти. При высоких температурах абсорбция азотных газов менее эффективна, поэтому для снижения температуры в башне оборудованы охлаждающие змеевики с циркулирующей водой. Отдача тепла происходит также путем радиации в окружающую среду. Небольшое количество азотных газов не абсорбируется и уходит наружу через верхнюю часть башни. Эту потерю будем считать зависимой переменной. Уменьшение потери газа имеет первостепенное экономическое значение. Определим, прежде всего, ее зависимость от двух независимых переменных: температуры воздуха x_a и температуры охлаждающей воды x_w. В последующем разделе рассмотрим еще и третью независимую переменную — крепость азотной кислоты, циркулирующей в поглотительной башне.

* Всестороннее исследование вопросов, связанных с корреляцией, можно найти в книге Yule and Kendall, An Introduction to the Theory of Correlation.

(B) ДВЕ НЕЗАВИСИМЫЕ ПЕРЕМЕННЫЕ

Определим зависимость потерь в окислительной установке (обозначаем величину их через y) от температуры охлаждающей воды x_w и от температуры воздуха x_a.

Данные об этих трех переменных собраны за 139 дней.

Было принято, что величина потерь в первом приближении зависит линейно от двух независимых переменных; зависимость эта выражается уравнением вида

$$y = a_{y \cdot wa} + b_{yw \cdot a}(x_w - \bar{x}_w) + b_{ya \cdot w}(x_a - \bar{x}_a).$$

Смысл коэффициентов:

$a_{y \cdot wa}$ — значение, которое величина y имела бы, если бы x_a и x_w были постоянными и равнялись своим средним значениям;

$b_{yw \cdot a}$ — скорость возрастания y при увеличении x_w, если бы x_a была постоянна;

$b_{ya \cdot w}$ — скорость возрастания y при увеличении x_a, если бы x_w была постоянна.

Требуется найти значения этих трех коэффициентов, наиболее соответствующие имеющимся данным. Поступаем следующим образом.

Для облегчения подсчетов, особенно в случае, если не пользуются счетными приборами, применяются произвольные нулевые точки; в случае надобности избежать дробей, умножают соответствующие величины на 10 или на 100 и т. д.

Переменные, принятые в данном случае, таковы:

$y = 10 \times$ (потери в $\frac{г}{м^3}$ минус 3,0);

x_w — температура воды в градусах Цельсия минус 20° C;

x_a — температура воздуха в градусах Фаренгейта минус 50° F.

Для каждой системы значений y, x_w и x_a подсчитываем три произведения yx_w, x_wx_a и x_ay. Таблица наблюдений за 139 дней получилась в виде:

Дата	y	x_w	x_a	x_wx_a	x_ay	x_wy
30. XII. 1943	−20	−6	−11	66	220	120
29	−9	1	−3	−3	27	−9
28	−18	1	−2	−2	36	−18
27	−6	−1	−1	1	6	6
25	−3	2	−3	−6	9	−6
23	−5	−5	−10	50	50	25
22	−20	2	−20	−40	400	−40
21	+19	5	+14	70	266	+95
18	−4	−1	−14	14	56	4
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
	1479	406	1	3533	17 340	12 797

Внизу таблицы даны итоги по каждому столбцу. Кроме того, подсчитаны суммы квадратов результатов индивидуальных наблюдений по трем переменным:

$$\sum y^2 = 99\,077; \quad \sum x_w^2 = 3456; \quad \sum x_a^2 = 10\,525.$$

В таком виде все итоги относятся к отклонениям от произвольных нулевых точек. Их следует преобразовать в отклонения от средних с помощью корректирующих величин обычного типа. Так, например:

$$\sum (y - \bar{y})^2 = \sum (y^2) - \frac{(\sum y)^2}{N} = 99\,077 - \frac{1479^2}{139} = 83\,340;$$

$$\sum (y - \bar{y})(x_w - \bar{x}_w) = \sum (y \cdot x_w) - \frac{(\sum y)(\sum x_w)}{N} = $$

$$= 12\,797 - \frac{1479 \cdot 406}{139} = 8477.$$

Аналогичным путем подсчитываются и прочие величины как отклонения от средних. Будем обозначать суммирование отклонений от средних символом $\sum'$. Так,

$$\sum (y - \bar{y})(x_w - \bar{x}_w) = \sum' y\, x_w, \quad \sum (y - \bar{y})^2 = \sum' y^2 \text{ и т. д.}$$

В результате расчетов получено:

$$\sum' y^2 = 83\,340; \quad \sum' x_w^2 = 2270{,}1; \quad \sum' x_a^2 = 10\,525,$$
$$\sum' yx_w = 8477; \quad \sum' yx_a = 17\,329{,}4; \quad \sum' x_wx_a = 3530{,}1.$$

Коэффициенты регрессии $b_{yw \cdot a}$ и $b_{ya \cdot w}$ вычисляются из системы уравнений:

$$\sum' yx_w = b_{yw \cdot a} \sum' x_w^2 + b_{ya \cdot w} \sum' x_wx_a;$$
$$\sum' yx_a = b_{yw \cdot a} \sum' x_wx_a + b_{ya \cdot w} \sum' x_a^2.$$

Подставляя числовые данные, получим:

$$8477 = 2270\ b_{yw \cdot a} + 3530\ b_{ya \cdot w}. \qquad \text{(A)}$$
$$17\,329 = 3530\ b_{yw \cdot a} + 10\,525\ b_{ya \cdot w}. \qquad \text{(B)}$$

Решение этих уравнений:

$$b_{yw \cdot a} = 2{,}454, \quad b_{ya \cdot w} = 0{,}8234.$$

Уравнение для величины y получает вид:

$$(y - \bar{y}) = 2{,}454\ (x_w - \bar{x}_w) + 0{,}8234\ (x_a - \bar{x}_a)$$

или

$$y - \frac{1479}{139} = 2{,}454\left(x_w - \frac{406}{139}\right) + 0{,}8234\left(x_a - \frac{1}{139}\right),$$

или

$$y = 3{,}4665 + 2{,}454\ x_w + 0{,}8234\ x_a. \qquad \text{(C)}$$

Прежде чем итти дальше, следует проверить значимость двух коэффициентов регрессии $b_{yw \cdot a}$ и $b_{ya \cdot w}$. Для этого применяется дисперсионный анализ.

Сумма квадратов зависимой переменной y, очевидно, равна $\Sigma' y^2 = 83\,340$; затем подсчитываются суммы квадратов y, определяемые каждой из двух переменных независимо одна от другой. Они равны:

$$\frac{(\Sigma' y x_w)^2}{\Sigma' x_w^2} = \frac{8477^2}{2270} = 31\,656{,}2 \quad (\text{для } x_w)$$

и

$$\frac{(\Sigma' y x_a)^2}{\Sigma' x_a^2} = \frac{17329^2}{10\,525} = 28\,531{,}5 \quad (\text{для } x_a).$$

Бóльшая из этих двух величин записывается в первую строку ниже приведенной табл. 10.1.

Следующим шагом является подсчет суммы квадратов y, определяемой двумя независимыми переменными совместно. Эта сумма равна

$$b_{yw \cdot a}\, \Sigma' y x_w + b_{ya \cdot w}\, \Sigma' y x_a = 2{,}454 \cdot 8477 + 0{,}8234 \cdot 17\,329 = 35\,071.$$

Далее, сумма квадратов, определяемая добавлением второй независимой переменной x_a, равна разности между суммой квадратов, определяемой x_w и x_a совместно, и суммой квадратов, определяемой только x_w, т. е.

$$35\,071 - 31\,656{,}2 = 3414{,}8.$$

Записываем вычисленные величины в таблицу дисперсионного анализа.

ТАБЛИЦА 10.1

Источники дисперсии	Сумма квадратов	Числа степеней свободы	Средние квадратов
Вариация, определяемая величиной x_w	31 656	1	31 656
Увеличение, определяемое добавлением x_a	3415	1	3415
Итог, определяемый величинами x_w и x_a совместно	35 071		
Остаток	48 269	136	354,9
Итого	83 340	138	—

Числа степеней свободы для каждого из коэффициентов регрессии равны 1, оставшиеся степени относятся к остатку. Аналогично, остаточная сумма квадратов равна разности между общим итогом и итогом, определяемым двумя коэффициентами регрессии.

Для проверки на значимость второго (меньшего) члена регрессии сравнивается его среднее квадратов с остаточным средним. В данном случае это отношение дисперсий составляет $\frac{3415}{354{,}9} = 9{,}6$ при числах степеней свободы $n_1 = 1$, $n_2 = 136$. Это значительно выше 1%-го уровня значимости, так что едва ли можно усомниться в реальности эффекта величины x_a.

Для оценки величины y при заданных значениях x мы пользуемся уравнением

$$y = 3{,}466 + 2{,}454 x_w + 0{,}823 x_a.$$

Величина дисперсии оценки выражается остаточной дисперсией в вышеприведенной таблице, а именно числом 354,9. Стандартное отклонение выражается значением квадратного корня из этого числа, т. е. 18,8.

Часто полезно видоизменение расчета, введенное Фишером. В целях общности обозначим $x_1 = x_w$, $x_2 = x_a$.

Таким образом,

$$\Sigma' y^2 = 83340{,}0 \qquad \Sigma' y x_1 = 8477{,}0$$
$$\Sigma' x_1^2 = 2270{,}1 \qquad \Sigma' y x_2 = 17329{,}4$$
$$\Sigma' x_2^2 = 10525{,}0 \qquad \Sigma' x_1 x_2 = 3530{,}1.$$

Прежде всего решаем уравнения:

$$c_{a1}\, \Sigma' x_1^2 + c_{a2}\, \Sigma' x_1 x_2 = 1, \qquad \text{(D)}$$
$$c_{a1}\, \Sigma' x_1 x_2 + c_{a2}\, \Sigma' x_2^2 = 0. \qquad \text{(E)}$$

Величины c_{a1} и c_{a2} являются обычными частными коэффициентами регрессии; они могли бы быть полнее описаны символами $b_{y1 \cdot 2}$ и $b_{y2 \cdot 1}$, как делалось ранее. Подставляем данные нашего примера в (D) и (E):

$$2270{,}1\, c_{a1} + 3530{,}1\, c_{a2} = 1, \qquad \text{(F)}$$
$$3530{,}1\, c_{a1} + 10525\, c_{a2} = 0. \qquad \text{(G)}$$

Корни этой системы уравнений

$$c_{a1} = 0{,}9208 \cdot 10^{-3}, \qquad c_{a2} = -0{,}3088 \cdot 10^{-3}.$$

Далее решается система уравнений

$$c_{b1}\, \Sigma' x_1^2 + c_{b2}\, \Sigma' x_1 x_2 = 0, \qquad \text{(H)}$$
$$c_{b1}\, \Sigma' x_1 x_2 + c_{b2}\, \Sigma' x_2^2 = 1. \qquad \text{(I)}$$

Заметим, что левые части уравнений (H) и (I) тождественны с (D) и (E), правые же части поменялись местами.

После подстановки значений сумм получаем решение:

$$c_{b1} = -0{,}3088\cdot10^{-3},\quad c_{b2} = 0{,}1986\cdot10^{-3}.$$

Теперь подсчитываются полные значения b_1 и b_2:

$$b_1 = c_{a1}\Sigma' yx_1 + c_{a2}\Sigma' yx_2 =$$
$$= 0{,}9208\cdot10^{-3}\cdot8477 - 0{,}3088\cdot10^{-3}\cdot17\,329{,}4 = 2{,}454; \qquad (J)$$
$$b_2 = c_{b1}\Sigma' yx_1 + c_{b2}\Sigma' yx_2 =$$
$$= -0{,}3088\cdot10^{-3}\cdot8477 + 0{,}1986\cdot10^{-3}\cdot17329{,}4 = 0{,}824. \qquad (K)$$

Эти значения b_1 и b_2, разумеется, совпали с полученными ранее для $b_{yw\cdot a}$ и $b_{ya\cdot w}$. Отметим одно из принципиальных преимуществ последнего приема: здесь легко получить частные коэффициенты регрессии для второй зависимой переменной Y, не решая вновь систему уравнений. В настоящем примере мы, правда, занимаемся лишь одной зависимой переменной — комбинированными потерями газа. Однако можно заинтересоваться отношением количеств азотистых и азотных газов и колебаниями этого отношения в зависимости от двух независимых переменных. Для этого потребуется найти суммы произведений новой зависимой переменной Y с x_1 и x_2, $\Sigma' Yx_1$ и $\Sigma' Yx_2$ и подставить в уравнения (J) и (K). Дополнительная затрата труда, таким образом, очень невелика, особенно если сравнить с решением новых систем уравнений, как требовалось прежде.

Вторым преимуществом настоящего способа расчета является непосредственное получение стандартной ошибки каждого частного коэффициента регрессии. Для этого требуется знать величину остаточной дисперсии σ_r^2, которая уже имеется в табл. 10.1. В новых обозначениях получим:

$$\sigma_r^2 = \frac{1}{n-3}\left(\Sigma' y^2 - b_1\Sigma' yx_1 - b_2\Sigma' yx_2\right) = 354{,}9.$$

Тогда стандартная ошибка коэффициента b равна

$$\sigma_r\sqrt{c_{a1}} = \sqrt{354{,}9}\cdot\sqrt{0{,}9208\cdot10^{-3}} = 0{,}574.$$

То же для b_2:

$$\sigma_r\sqrt{c_{b2}} = \sqrt{354{,}9}\cdot\sqrt{0{,}1986\cdot10^{-3}} = 0{,}266.$$

Эти стандартные ошибки имеют числа степеней свободы равными $n-3$; поскольку n обычно велико, любой коэффициент регрессии, в пределах удвоенной величины своей стандартной ошибки, может считаться значимым при уровне 5%. Таким образом, эта процедура является методом проверки значимости, альтернативным к указанному ранее в табл. 10.1. Можно также установить вероятные пределы (95%) для b_1: $2{,}454 \pm 2\cdot0{,}574 =$ $= 1{,}306$ или $3{,}602$. Поступая таким образом, выясняют, сколько значащих цифр следует удержать при окончательном установлении уравнения регрессии.

(C) НЕОБХОДИМОСТЬ МНОЖЕСТВЕННОЙ РЕГРЕССИИ И ЧАСТНОЙ КОРРЕЛЯЦИИ

При выводе коэффициентов регрессии $b_{yw\cdot a}$ и $b_{ya\cdot w}$ мы фиксировали переменную в индексе, стоящем за точкой. По этой причине эти коэффициенты и называются „частными" коэффициентами регрессии.

Обычные коэффициенты регрессии определялись бы уравнениями:

$$b_{yw} = \frac{\Sigma' y\,x_w}{\Sigma' x_w^2} = \frac{8477}{2270} = 3{,}734,$$

$$b_{ya} = \frac{\Sigma' yx_a}{\Sigma' x_a^2} = \frac{17329}{10525} = 1{,}646.$$

Заметим, что эти обычные коэффициенты регрессии значительно больше соответствующих частных коэффициентов. Это легко объяснимо. Если подсчитать коэффициент корреляции x_w и x_a, получим:

$$r_{wa} = \frac{\Sigma' x_w x_a}{\sqrt{\Sigma' x_w^2\Sigma' x_a^2}} = \frac{3530}{\sqrt{2270\cdot10525}} = 0{,}7222.$$

Таким образом, налицо весьма значительная и очень важная корреляция между x_w и x_a. Конечно, это определяется природой процесса. Условия погоды, вызывающие понижение температуры воздуха x_a, повлекут и снижение температуры охлаждающей воды x_w.

Поэтому при расчете обычного коэффициента регрессии y на x, в силу того, что большие значения x_w будут стремиться ассоциироваться с большими значениями x_a, значения y для больших x будут стремиться стать бо̀льшими, чем следует, поскольку они также ассоциируются с большими значениями x_a. В частных же коэффициентах регрессии влияние другой переменной исключено, и они измеряют только эффект данной переменной.

В связи с этим иногда применяются частные коэффициенты корреляции. Если обычные коэффициенты корреляции равны:

$$r_{wa} = 0{,}7222,$$

$$r_{yw} = \frac{\Sigma' yx_w}{\sqrt{\Sigma' y^2\Sigma' x_w^2}} = \frac{8477}{\sqrt{83340\cdot2270}} = 0{,}6165,$$

$$r_{ya} = \frac{\Sigma' yx_a}{\sqrt{\Sigma' y^2\,\Sigma' x_a^2}} = \frac{17329}{\sqrt{83340\cdot10525}} = 0{,}5850,$$

то частные коэффициенты корреляции будут:

$$r_{yw\cdot a}=\frac{r_{yw}-r_{ya}r_{wa}}{\sqrt{(1-r_{ya}^2)(1-r_{wa}^2)}}=\frac{0{,}6165-0{,}5850\cdot0{,}7222}{\sqrt{(1-0{,}5850^2)\cdot(1-0{,}7222^2)}}=0{,}3458,$$

$$r_{ya\cdot w}=\frac{r_{ya}-r_{yw}r_{aw}}{\sqrt{(1-r_{yw}^2)(1-r_{aw}^2)}}=\frac{0{,}5850-0{,}6165\cdot0{,}7222}{\sqrt{(1-0{,}6165^2)\cdot(1-0{,}7222^2)}}=0{,}2567.$$

При проверке значимости обычного коэффициента корреляции число степеней свободы равно $n-2$, где n — число наблюдений. Для частного же коэффициента корреляции теряется еще одна степень свободы, так что число степеней свободы будет $n-3$.

В рассматриваемом случае, хотя коэффициенты и заметно уменьшаются при переходе от обычных к частным коэффициентам, их величины остаются еще значимыми при большом числе степеней свободы (139 — 3). В общем случае, однако, может оказаться, что положительный обычный коэффициент корреляции будет преобразован в отрицательный частный коэффициент корреляции; для этого (см. формулу для $r_{ya\cdot w}$) произведение $r_{yw}\cdot r_{aw}$ должно быть больше r_{ya}.

В рассматриваемом примере различие между обычными и частными коэффициентами корреляции хотя и было интересным для нас, но не имело непосредственного практического значения; в других случаях оно может быть важно. Предположим, например, что, основываясь на обычном коэффициенте регрессии 0,373, решили, что было бы экономично установить охлаждающее оборудование, с помощью которого средняя температура охлаждающей воды была бы снижена на 10° С. По обычной корреляции можно было бы ожидать уменьшения потерь на 3,73 единицы, что достаточно оправдывало бы установку. По частной же корреляции действительно достигнутое снижение потерь было бы 2,44 единицы; разность между 3,73 и 2,44 может оказаться достаточной, чтобы сделать установку неэкономичной.

(D) МНОЖЕСТВЕННАЯ КОРРЕЛЯЦИЯ ПРИ ТРЕХ НЕЗАВИСИМЫХ ПЕРЕМЕННЫХ

В рассмотренном примере существует еще третья переменная, влияние которой представляет значительный интерес. Это — крепость кислоты в поглотительной башне. Обозначим ее через x_s.

Определим коэффициенты a и b в уравнении

$$y=a_{y\cdot was}+b_{yw\cdot as}\cdot x_w+b_{ya\cdot ws}\cdot x_a+b_{ys\cdot aw}\cdot x_s,$$

в котором обозначения имеют тот же смысл, что и прежде: $a_{y\cdot was}$ — значение, которое имела бы величина y, если бы x_w, x_a и x_s поддерживались постоянными и равными их средним значениям, $b_{yw\cdot as}$ определяет скорость возрастания y при увеличении x_w, при x_a и x_s, постоянно равных их средним значениям.

Аналогично определяются $b_{ya\cdot ws}$ и $b_{ys\cdot aw}$.

Нам понадобятся следующие суммы, суммы квадратов и суммы произведений (полученные ранее мы вычисляем заново с бóльшим числом знаков):

$\sum y = 1479$	$\sum y^2 = 99077$	$\sum' y^2 = 83340{,}014$
$\sum x_w = 406$	$\sum x_w^2 = 3456$	$\sum' x_w^2 = 2270{,}1295$
$\sum x_a = 1$	$\sum x_a^2 = 10525$	$\sum' x_a^2 = 10524{,}9928$
$\sum x_s = 70$	$\sum x_s^2 = 36364$	$\sum' x_s^2 = 36328{,}7482$
$\sum yx_w = 12797$	$\sum' yx_w = 8477{,}0431$	
$\sum yx_a = 17340$	$\sum' yx_a = 17329{,}3597$	
$\sum yx_s = 12931$	$\sum' yx_s = 12186{,}1799$	
$\sum x_wx_a = 3533$	$\sum' x_wx_a = 3530{,}07914$	
$\sum x_ax_s = 5365$	$\sum' x_ax_s = 5364{,}4964$	
$\sum x_sx_w = 1943$	$\sum' x_sx_w = 1738{,}5396$	

Теперь требуется решить три системы трех уравнений (A), (B) и (C), приведенные ниже. Первая система уравнений имеет в правых частях (A), (B) и (C) соответственно 1, 0 и 0; решения этой системы будут c_{a1}, c_{b1} и c_{c1}. Вторая система имеет в правых частях 0, 1 и 0; ее решения будут c_{a2}, c_{b2} и c_{c2}. Третья система имеет в правых частях 0, 0 и 1; ее решения будут c_{a3}, c_{b3} и c_{c3}.

$$c_a\textstyle\sum' x_w^2+c_b\sum' x_wx_s+c_c\sum' x_wx_a=1,\ 0,\ 0 \qquad \text{(A)}$$
$$c_a\textstyle\sum' x_sx_w+c_b\sum' x_s^2+c_c\sum' x_sx_a=0,\ 1,\ 0 \qquad \text{(B)}$$
$$c_a\textstyle\sum' x_ax_w+c_b\sum' x_ax_s+c_c\sum' x_a^2=0,\ 0,\ 1 \qquad \text{(C)}$$

Решать эти уравнения можно как угодно, но систематический метод решения менее подвержен случайным ошибкам. Общие методы решения приведены в книге Уиттекера и Робинсона „Обработка наблюдений“ (E. T. Whittaker and G. Robinson: „The Calculus of Observations“ (Blackie), 3-е изд., гл. IX). Наиболее же удовлетворительный прием дан М. Дулитлом (M. H. Doolittle). Вписываем уравнения в графы (A′), (B′) и (C′) таблицы типа табл. 10.2.

Подставляя вышеуказанные числовые значения, получаем графы (A″), (B″), (C″). Подразумевается, что все цифры в столбце под заголовком c_a, являются значениями коэффициентов при c_a в соответствующих уравнениях; аналогично в столбцах c_b и c_c. Правые части уравнений временно увеличены в 10^4 раз, чтобы избежать в дробных уравнениях чрезмерно больших количеств нулей после запятой.

Затем записываем значения (A″) еще раз в графу (A‴); под ними записываем частные от деления их на коэффициент при c_a, взятый с обратным знаком (здесь это —2270,1295). Обозначаем эти величины через (D).

ТАБЛИЦА 10·2

	c_a	c_b	c_c	Первая система $\times 10^4$	Вторая система $\times 10^4$	Третья система $\times 10^4$
(A′)	$\sum' x_w^2$	$\sum' x_w x_s$	$\sum' x_w x_a$	10000	0	0
(B′)	$\sum' x_s x_w$	$\sum' x_s^2$	$\sum' x_s x_a$	0	10000	0
(C′)	$\sum' x_a x_w$	$\sum' x_a x_s$	$\sum' x_a^2$	0	0	10000
(A″)	2270,12950	1738,5396	3530,07914	10000	0	0
(B″)	1738,5396	36328,7482	5364,4960		10000	0
(C″)	3530,07914	5364,4960	10524,9928	0	0	10000
(A‴)	2270,12950	1738,5396	3530,07914	10000	0	0
(D)	—1,00000	— 0,76583279	— 1,55501223	— 4,4050351	0	0
(B‴	1738,5396	36328,7842	5364,4960	0	10000	0
(E)	— 1738,5396	— 1331,4306	— 2703,4503	— 7658,3279	0	0
(F)	0,0000	34997,3526	2661,0457	— 7658,3279	10000	0
(G)		— 1,0000	— 0,076035630	0,218825915	— 0,28573589	0
(C‴)	3530,07914	5364,4960	10524,9928	0	0	10000
(H)	— 3530,07914	— 2703,4502	— 5489,3161	— 15550,1223	0	0
(I)		— 2661,0458	— 202,3343	582,3058	— 760,35630	0
(J)	0,0000	0,0000	4833,3424	— 14967,8165	— 760,35630	10000
(K)			1,00000	— 3,0967838	— 0,1573148	2,0689617
(L_1)		— 1,0000	0,2354659	0,2188259		
(L_2)		— 1,0000	0,0119615		— 0,2857359	
(L_3)		— 1,0000	— 0,1573148			0,0
(L')		1,0000		0,0166400		
(L'_2)		1,0000			0,2976974	
(L'_3		1,0000				— 0,1573148
(M_1)	— 1,0000	— 0,0127435	4,8155366	— 4,4050351		
(M_2)	— 1,0000	— 0,2279864	0,2446251		0	
(M_3)	— 1,0000	0,1204768	— 3,2172607			0
(M'_1)	1,0000			9,2078282		
(M'_2)	1,0000				0,0166397	
(M'_3)	1,0000					— 3,0967839

Далее, записываем значения (B″) еще раз в графу (B‴). Умножаем (A‴) на коэффициент при c_b, стоящий в графе (D) (здесь — 0,76583279). Получаем значения (E). Складывая их с (B‴), получаем значения (F). Деля (F) на коэффициент при его первом члене, взятый с обратным знаком (т. е. на —34997,3536), получаем (G).

Затем записываем значения (C″) еще раз в графу (C‴). Умножаем (A‴) на коэффициент при c_c, стоящий в графе (D) (т. е. на — 1,55501223); получаем значения (H). Кроме того, умножаем (F) на коэффициент при c_c, стоящий в графе (G) и получаем значения (I). Складывая (C‴), (H) и (I), получаем значения (J). Деля последние на коэффициент при c_c (4833,3424), получим значения (K).

Три числовых значения K, стоящие в правой части таблицы, являются решениями для c_{c1}, c_{c2} и c_{c3}, увеличенными в 10^4 раз.

Беря теперь значения (G) и подставляя в столбце c_c решения для c_{c1}, c_{c2} и c_{c3}, получим значения (L_1), (L_2) и (L_3). Отсюда сразу получаем значения (L'_1), (L'_2) и (L'_3), являющиеся решениями для c_{b1}, c_{b2} и c_{b3}. [(L'_1) в правой части таблицы является разностью значения (L_1), стоящего в столбце c_c и значения (L_1), стоящего в столбце „первая система“; аналогично определены (L'_2) и (L'_3). — *Прим. перев.*]

Теперь подставляем решения для c_b и c_c в (D), получая величины (M_1), (M_2) и (M_3); от последних приходим к величинам (M'_1), (M'_2) и (M'_3), которые являются решениями для c_{a1}, c_{a2} и c_{a3}. [(M'_1) в правой части таблицы получено как разность между суммой значений (M_1), стоящих в столбцах c_b и c_c, и значением (M_1) в столбце „первая система“. Аналогично определены (M'_2) и (M'_3). — *Прим. перев.*]

Сводим полученные решения в матрицу, приведенную в таблице 10. 3.

ТАБЛИЦА 10.3

(Все значения умножены на 10^4.)

	a	*b*	*c*
1	9,2078282	0,0166400	—3,0967838
2	0,0166397	0,2976974	—0,1573148
3	—3,0967839	—0,1573148	2,0689617

Подобно исходным уравнениям, эта матрица симметрична: $c_{b1} = c_{a2}$; $c_{c1} = c_{a3}$; $c_{c2} = c_{b3}$.

Хорошей проверкой правильности решений явится подстановка в (C″) решений для c_a. В левой части уравнения мы получим:

$$(3530{,}07914 \cdot 9{,}207822 + 5364{,}4960 \cdot 0{,}0166397 - 10524{,}9928 \cdot 3{,}0967839) \cdot 10^{-4} = 0{,}00000025.$$

Эта величина достаточно близка к ожидавшемуся значению, равному нулю.

Теперь вычисляем коэффициенты регрессии:

$$b_{w\cdot as} = c_{a1} \Sigma' yx_w + c_{a2} \Sigma' yx_s + c_{a3} \Sigma' yx_a = (9{,}2078282 \cdot 8477{,}0431 + 0{,}0166397 \cdot 12186{,}1799 - 3{,}0967839 \cdot 17329{,}3597) \cdot 10^{-4} = 2{,}4592651.$$

$$b_{s\cdot aw} = c_{b1} \Sigma' yx_w + c_{b2} \Sigma' yx_s + c_{b3} \Sigma' yx_a = 0{,}1042687.$$

$$b_{a\cdot sw} = c_{c1} \Sigma' yx_w + c_{c2} \Sigma' yx_s + c_{c3} \Sigma' yx_a = 0{,}7685143.$$

Полезной проверкой последних операций явится подстановка значений в уравнение:

$$b_{w\cdot as} \Sigma' x_a x_w + b_{s\cdot aw} \Sigma' x_a x_s + b_{a\cdot ws} \Sigma' x_a^2 = \Sigma' yx_a.$$

После подстановки числовых значений получаем:

$$17329{,}3566 = 17329{,}3597.$$

Расхождение на 0,0031 незначительно и им можно пренебречь.

Величина остаточной дисперсии определяется из уравнения:

$$\sigma_r^2 = \frac{1}{n-4} \cdot (\Sigma' y^2 - b_{w\cdot as} \Sigma' yx_w - b_{s\cdot aw} \Sigma' yx_s - b_{a\cdot sw} \Sigma' yx_a) = \frac{1}{135}(83340{,}014 - 2{,}4592651 \cdot 8477{,}0431 - 0{,}1042687 \cdot 12186{,}1799 - 0{,}7685143 \cdot 17329{,}3597) = 354{,}846.$$

Стандартные ошибки трех коэффициентов регрессии определяются так:

$$\sigma_{w\cdot as} = \sigma_r \cdot \sqrt{c_{a1}} = \sqrt{354{,}847 \times 9{,}2078 \times 10^{-4}} = 0{,}5716,$$

$$\sigma_{s\cdot aw} = \sigma_r \cdot \sqrt{c_{b2}} = 0{,}1028,$$

$$\sigma_{a\cdot sw} = \sigma_r \cdot \sqrt{c_{b3}} = 0{,}2709.$$

Для проверки значимости каждого коэффициента вычисляют значение величины t, равное отношению коэффициента регрессии к величине его стандартной ошибки и разыскивают это значение в таблице t, при числе степеней свободы, равном числу степеней свободы остаточной дисперсии. Например, для коэффициента $b_{a\cdot sw}$ величина t равна $t = \frac{0{,}7685143}{0{,}2709} = 2{,}837$, при 135 степенях свободы. Обращаясь к таблице I в приложении, видим, что это значение t соответствует уровню значимости между 1% и $0{,}1\%$.

Доверительные пределы для каждой оценки можно вывести обычным путем. Для $b_{a\cdot sw}$ и доверительного уровня в 95% пределы эти таковы: $0{,}768514 \pm 1{,}98 \cdot 0{,}2709$, или между 0,232132 и 1,304896.

Отметим попутно, что мы сохранили чрезмерное количество значащих цифр; однако гораздо лучше опускать последние цифры в самом конце расчета, а не в начале его.

Окончательный вид уравнения регрессии:

$$y-\bar{y}=b_{w\,.\,as}(x_w-\bar{x}_w)+b_{s\,.\,aw}(x_s-\bar{x}_s)+b_{a\,.\,sw}(x_a-\bar{x}_a),$$

или:

$$y-\frac{1479}{139}=2{,}459265\left(x_w-\frac{406}{139}\right)+0{,}104269\left(x_s-\frac{70}{139}\right)+$$
$$+0{,}768514\left(x_a-\frac{1}{139}\right),$$

или, наконец:

$$y=2{,}459x_w+0{,}104x_s+0{,}768x_a+3{,}399.$$

(E) ЗАКЛЮЧЕНИЕ

Употребление множественной регрессии требует довольно громоздких арифметических вычислений; однако при наличии массы данных с несколькими независимыми переменными это единственный серьезный способ оценки влияния последних.

Так, например, в рассматривавшемся примере мы особенно интересовались ролью независимой переменной с наименьшим коэффициентом регрессии x_s.

Полностью исключив влияние двух других независимых переменных, мы получили оценку коэффициента регрессии, смогли проверить его значимость путем сравнения его дисперсии с остаточной дисперсией 354,8 и нашли его совершенно незначимым. Если бы мы пренебрегли двумя другими переменными, то сумма квадратов, относящаяся к x_s, получилась бы равной 4088 при величине остатка в $\frac{79252}{137}=578{,}5$. При таких условиях коэффициент регрессии получился бы значимым при уровне в 1%; это было бы совершенно ошибочным выводом.

Другая причина для использования множественной регрессии была подробно рассмотрена ранее; было указано на то, что можно получить вводящие в заблуждение результаты, если пользоваться простой регрессией, когда независимые переменные связаны друг с другом. В данном примере независимые переменные имели положительные коэффициенты регрессии и положительную корреляцию и тем самым привели бы к фиктивно-большим значениям обоих коэффициентов регрессии. Если же они имели бы отрицательную корреляцию, мы получили бы фиктивно-малые значения коэффициентов регрессии при пользовании обычной регрессией. Разумеется, такого явления не получится, если независимые переменные некоррелированы. Тогда единственным недостатком при применении обычной регрессии вместо частной (или множественной) регрессии было бы снижение точности оценки, поскольку остаток был бы намного выше.

Как и в случаях обычной регрессии, экстраполировать уравнения множественной регрессии рискованно, если нет на то очень серьезного теоретического обоснования. Для того чтобы предотвратить экстраполяцию, следует всегда указывать диапазон независимых переменных, встретившихся в данных, на котором выводились уравнения регрессии.

ГЛАВА XI

ОБЩИЙ ДИСПЕРСИОННЫЙ АНАЛИЗ

(А) ВВЕДЕНИЕ

Обстановка, в которой разумно применять дисперсионный анализ, может сложиться различными путями:

(1) Случайно: исследуемый процесс удобен по своей структуре. Такой пример рассмотрен в разделе (I) настоящей главы; там разбирается работа трех цехов, каждый из которых обрабатывает три сорта слабой кислоты. Факт наличия трех цехов не зависел от статистика; для него это было случайным стечением обстоятельств.

(2) Путем небольшого видоизменения существующего производственного процесса. Может быть, например, что в некотором процессе к каждой партии сырого материала добавляется переменное количество материала с неопределенной „историей“, поступающего во вторичную переработку; в целях исследования эти добавления могут быть временно прекращены.

(3) Путем специально проводимого эксперимента. Несколько факторов могут подаваться на различных уровнях, специально рассчитанных для проведения факторных испытаний.

(B) ВИДЫ АНАЛИЗОВ

В применениях дисперсионного анализа часто возникает затруднение в выборе надлежащей формы анализа. Следующее затруднение представляет методика применения критериев значимости. Мы попытаемся наметить здесь общие приемы исследования.

В последующих примерах будет рассматриваться гипотетический полимер, перерабатываемый в волокна: зависимой переменной будет прочность волокон на растяжение; будут рассматриваться различные модификации основной идеи с целью показать взаимные соотношения различных форм анализа. После каждой формы этого основного примера будет детально рассмотрен другой пример сходной структуры. Кроме того, ряд вопросов применения дисперсионного анализа рассматривается в главе XII.

Все виды анализа можно подразделить на два общих типа:

(I) Виды, в которых категории сохраняют свое взаимнооднозначное соответствие. Эти виды можно назвать чисто-факторными анализами.

(II) Виды, в которых некоторые категории не сохраняют своего взаимнооднозначного соответствия. Эти виды можно назвать неполными анализами.

Смысл этой классификации станет ясен ниже при детальном рассмотрении примеров.

(C) АНАЛИЗ С ДВУМЯ ФАКТОРАМИ

Часто бывает возможно классифицировать комплект индивидуумов двояким путем; так, можно иметь сырье, классифицированное по сортаменту g соответственно номинальной степени полимеризации. Пусть в нашем распоряжении имеется ряд партий материала каждого сорта; одна партия каждого сорта может обрабатываться при одной из t различных температур. Таким образом, имеются два фактора: сорт полимера — с g уровнями и температура обработки — с t уровнями.

Каждый индивидуум будет целиком определяться своими значениями g и t.

Такой пример был рассмотрен в главе VII, (E). В табл. 11.1 приведены числа степеней свободы и компоненты дисперсии.

ТАБЛИЦА 11.1

Источник дисперсии	Числа степеней свободы	Компоненты дисперсии
G	$g-1$	$t\sigma_g^2+\sigma_0^2$
T.	$t-1$	$g\sigma_t^2+\sigma_0^2$
Остаточная . .	$(g-1)(t-1)$	σ_0^2
Итого	$gt-1$	

Порядок проверки значимости очевиден. Если фактор G существует, т. е. σ_g^2 больше нуля, величина $t\sigma_g^2+\sigma_0^2$ должна быть больше σ_0^2. Значимость этого можно легко проверить по методу отношения дисперсий. Фактор T проверяется аналогичным путем.

(D) АНАЛИЗ С ТРЕМЯ ФАКТОРАМИ

В предыдущем разделе рассматривался случай g сортов полимера; по одной партии каждого сорта обрабатывалось при какой-либо из t различных температур. Предположим теперь, что каждый из s поставщиков снабжает нас полимером всех сортов и что t партий сырья каждого сорта от каждого поставщика обрабатываются на каждой из t температур.

Важно, чтобы сырье первого сорта от первого поставщика имело бы ту же номинальную степень полимеризации, что и сырье первого сорта от второго, третьего и т. д. поставщиков.

Аналогичное сходство номинальных свойств должно быть у всех сортов материала, получаемого от разных поставщиков. Таким образом, материал первого сорта от всех поставщиков имеет определенное общее свойство, отличающее его от всех вторых сортов, и т. д.

Всего теперь имеется gts индивидуумов; каждый из них однозначно определяется заданными значениями g, t и s. Таким образом, налицо три фактора: поставщики, сорта полимера и температуры обработки.

Техника вычислений при дисперсионном анализе при трех факторах приводится ниже. В табл. 11.2 приведены числа степеней свободы и компоненты дисперсии.

ТАБЛИЦА 11.2

Источники дисперсии	Числа степеней свободы	Компоненты дисперсии
G	$g-1$	$\sigma_0^2+s\sigma_{gt}^2+t\sigma_{gs}^2+ts\sigma_g^2$
T.	$t-1$	$\sigma_0^2+g\sigma_{ts}^2+s\sigma_{gt}^2+sg\sigma_t^2$
S.	$s-1$	$\sigma_0^2+t\sigma_{sg}^2+g\sigma_{st}^2+gt\sigma_s^2$
$G\times T$.	$(g-1)(t-1)$	$\sigma_0^2+s\sigma_{gt}^2$
$T\times S$.	$(t-1)(s-1)$	$\sigma_0^2+g\sigma_{st}^2$
$S\times G$.	$(s-1)(g-1)$	$\sigma_0^2+t\sigma_{sg}^2$
Остаточная	$(g-1)(t-1)(s-1)$	σ_0^2
Итого . . .	$gts-1$	

Проверка значимости должна быть начата с наименьшего взаимодействия путем сравнения его дисперсии с остаточной.

Так, если наименьшим взаимодействием является, например, $S\times G$, проверка делается путем определения отношения дисперсии $\sigma_0^2+t\sigma_{sg}^2$ к σ_0^2. Если это взаимодействие не является значимым, то присоединяем его к остатку и получаем новую оценку величины σ_0^2; аналогично оперируют и с другими взаимодействиями, беря их в порядке возрастания величин.

При проверке основных влияний может быть ряд случаев:

(I) Все три взаимодействия незначимы. Вычеркивая их компоненты дисперсии из табл. 11.2 и присоединяя их числа степеней свободы к остатку, получаем табл. 11.3.

ТАБЛИЦА 11.3

Источник дисперсии	Числа степеней свободы	Компоненты дисперсии
G	$g-1$	$\sigma_0^2+ts\sigma_g^2$
T	$t-1$	$\sigma_0^2+sg\sigma_t^2$
S	$s-1$	$\sigma_0^2+gt\sigma_s^2$
Остаточная	$gts-(g+t+s)+2$	σ_0^2
Итого ...	$gts-1$	

Проверка значимости трех основных эффектов ведется так, что для фактора G, очевидно, сравнивают $\sigma_0^2+ts\sigma_g^2$ с σ_0^2; аналогично поступают и для T и S.

(II) Одно из взаимодействий, например $G\times T$, значимо, другие же незначимы. Вычеркивая последние из табл. 11.2, получают табл. 11.4.

ТАБЛИЦА 11.4

Источники дисперсии	Числа степеней свободы	Компоненты дисперсии
G	$g-1$	$\sigma_0^2+s\sigma_{gt}^2+ts\sigma_g^2$
T	$t-1$	$\sigma_0^2+s\sigma_{gt}^2+sg\sigma_t^2$
S	$s-1$	$\sigma_0^2+gt\sigma_s^2$
$G\times T$	$(g-1)(t-1)$	$\sigma_0^2+s\sigma_{gt}^2$
Остаточная	$(gt-1)(s-1)$	σ_0^2
Итого ...	$gts-1$	

Для проверки существования основного эффекта S требуется установить наличие σ_s^2; это может быть строго сделано сравнением $\sigma_0^2+gt\sigma_s^2$ с σ_0^2.

Для основного эффекта G требуется установить наличие σ_g^2; для этого нужно показать, что $\sigma_0^2+s\sigma_{gt}^2+ts\sigma_g^2$ превосходит $\sigma_0^2+s\sigma_{gt}^2$; другими словами, основной эффект G сравнивается с взаимодействием $G\times T$. Таким же путем проверяется основной эффект T.

(III) Два взаимодействия значимы, например $G\times T$ и $T\times S$. Табл. 11.2 превращается в табл. 11.5.

ТАБЛИЦА 11.5

Источники дисперсии	Числа степеней свободы	Компоненты дисперсии
G	$g-1$	$\sigma_0^2+s\sigma_{gt}^2+ts\sigma_g^2$
T	$t-1$	$\sigma_0^2+g\sigma_{ts}^2+s\sigma_{gt}^2+sg\sigma_t^2$
S	$s-1$	$\sigma_0^2+g\sigma_{st}^2+gt\sigma_s^2$
$G\times T$	$(g-1)(t-1)$	$\sigma_0^2+s\sigma_{gt}^2$
$T\times S$	$(t-1)(s-1)$	$\sigma_0^2+t\sigma_{st}^2$
Остаточная	$t(g-1)(s-1)$	σ_0^2
Итого ...	$gts-1$	

Рассматривая компоненты дисперсии, мы видим, что значимость основных эффектов G и S может быть установлена путем сравнения их средних квадратов со средними квадратов соответственно для взаимодействий $G\times T$ и $T\times S$. Но в этом случае нельзя строго установить наличие основного эффекта T, так как хотя и можно было бы доказать, что он больше взаимодействия $G\times T$, но этот результат мог бы получиться как вследствие существования σ_{ts}^2, так в равной мере и вследствие существования σ_t^2.

Это затруднение практически не очень существенно, так как весь анализ можно разбить на отдельные анализы с двумя факторами $G\times S$ для каждого уровня T. Этот прием показан в разделе (Е) настоящей главы.

Рассмотрим теперь данные о процессе очистки некоторого продукта, приведенные в таблице. Данные, содержащиеся в таблице (зависимая переменная x), относятся к чистоте продукта, измеренной определенным способом. Фактор T с тремя уровнями есть время кипячения. Фактор A представляет примененный растворитель; W_1 — окончательный холодный смыв продукта, W_2 — окончательный горячий смыв.

		T_1	T_2	T_3
A_1	W_1	3,1	2,0	1,3
	W_2	2,1	1,6	1,9
A_2	W_1	4,3	2,8	4,1
	W_2	6,1	3,9	2,6

В анализе была принята произвольная нулевая точка 3,0. Этим уменьшены величины всех данных, а следовательно, и их квадратов. Затем, чтобы избежать десятичных дробей, числа были умножены на 10. Суммы квадратов стали поэтому в 100 раз большими, но это не отражается на анализе, поскольку при проверке значимости отдельные члены сравниваются друг с другом и перемена масштаба не играет роли. Если же, однако, пожелают рассчитать величину какой-либо дисперсии, необходимо ввести поправку на изменение масштаба, разделив полученные величины на 100.

Результаты преобразования приняли вид:

		T_1	T_2	T_3
A_1	W_1	1	—10	—17
	W_2	—9	—14	—11
A_2	W_1	13	—2	11
	W_2	31	9	—4

Следует помнить, что данные вполне симметричны относительно трех переменных A, W и T. Это обстоятельство было бы нагляднее, если бы мы были в состоянии расположить данные в трех измерениях и поместить нижнюю половину таблицы (для A_2) над верхней половиной таблицы (для A_1). Тогда горизонтальная ось поперек страницы представила бы различие между продолжительностями кипячения (T), вертикальная ось (сверху вниз в плоскости страницы) — тип смыва (W) и ось, перпендикулярная плоскости страницы, — тип растворителя (A).

В том виде, как она приведена, таблица по существу содержит две отдельные таблицы W и T для каждого из значений A, это — совершенно произвольная комбинация; с равным успехом можно представить себе как бы таблицы W и A для каждого из значений T, или таблицы A и T для каждого из значений W.

Прежде всего составляем три таблицы „с двумя входами", производя суммирование по каждой из трех переменных поочередно. Суммируя по W при каждом значении A и T, получаем таблицу:

	T_1	T_2	T_3	Итого
A_1	— 8	—24	—28	—60
A_2	44	7	7	58
Итого . . .	36	—17	—21	— 2

Суммируя по A при каждом значении T и W, получим такую таблицу:

	T_1	T_2	T_3	Итого
W_1	14	—12	— 6	—4
W_2	22	— 5	—15	+2
Итого . . .	36	—17	—21	—2

Третья таблица получается путем суммирования по T при каждом значении A и W:

	A_1	A_2	Итого
W_1	—26	22	— 4
W_2	—34	36	+ 2
Итого . . .	—60	58	— 2

Анализируем первую из составленных таблиц (для A и T). Для этого нам понадобятся следующие величины:

(1) Сумма квадратов индивидуумов данной таблицы, деленная на количество первоначальных индивидуумов, просуммированных для получения индивидуумов рассматриваемой таблицы с двумя входами. В нашем случае каждый индивидуум является суммой двух первоначальных индивидуумов. Таким образом, получаем:

$$\frac{(-8)^2+(-24)^2+(-28)^2+44^2+7^2+7^2}{2}=1729.$$

(2) Сумма квадратов итогов по строкам, деленная на число первоначальных индивидуумов, просуммированных для получения итогов по строкам. В данном случае каждый итог по строке является суммой 6 первоначальных индивидуумов (каждый итог является суммой 3 индивидуумов таблицы с двумя входами, каждый из которых, в свою очередь, является суммой 2 первоначальных индивидуумов). Мы имеем:

$$\frac{(-60)^2+58^2}{6}=1160{,}67.$$

(3) То же для итогов по столбцам. В данном случае каждый итог по столбцу является суммой 4 первоначальных индивидуумов; поэтому делитель равен 4:

$$\frac{36^2+(-17)^2+(-21)^2}{4}=506{,}5.$$

(4) Квадрат общего итога, деленный на общее число первоначальных индивидуумов:

$$\frac{(-2)^2}{12}=0{,}33.$$

Заметим, что эти операции подобны анализу по строкам и столбцам, рассмотренному ранее; отличие лишь в том, что каждая сумма квадратов делится не на число индивидуумов в таблице с двумя входами, относящихся к каждой сумме квадратов, а на число первоначальных индивидуумов, относящихся к каждой сумме квадратов.

Составляем теперь таблицу дисперсионного анализа:

Источники дисперсии	Числа степеней свободы	Суммы квадратов	Средние квадратов
Между A	$a-1=2-1=1$	$(2)-(4)=1160{,}33$	1160,33
Между T.	$t-1=3-1=2$	$(3)-(4)=\ 506{,}17$	253,08
Взаимодействие $A\times T$.	$(a-1)(t-1)=2$	62,17	31,08
Итого	$at-1=5$	$(1)-(4)=1728{,}67$	

В этой таблице a — число уровней величины A, t — число уровней величины T.

Эта таблица очень похожа на рассмотренную ранее таблицу анализа по строкам и столбцам; разница в том, что в строке остаточной дисперсии теперь указано „взаимодействие $A\times T$“. Причина этого будет объяснена ниже. Сумму квадратов, соответствующую этой строке, лучше всего взять как разность между итогом и суммой значений для строк A и T, т. е.

$$1728{,}67-(1160{,}33+506{,}17)=62{,}17.$$

Совершенно такие же операции выполняются и по другим таблицам с двумя входами.

Наконец, общая дисперсия получается путем возведения в квадрат всех первоначальных индивидуумов, сложения этих квадратов и вычитания из этой суммы возведенного в квадрат общего итога, деленного на полное число первоначальных индивидуумов:

$$1^2+(-10)^2+(-17)^2+(-9)^2+\ldots+(-4)^2-\frac{(-2)^2}{12}=2139{,}67.$$

Все результаты сводим в табл. 11.6:

ТАБЛИЦА 11.6

Источники дисперсии	Числа степеней свободы	Сумма квадратов	Средние квадратов	Компоненты дисперсии
Между A	$a-1=1$	1160,33	1160,33	$wt\sigma_a^2+w\sigma_{at}^2+$ $+t\sigma_{aw}^2+\sigma_0^2$
Между W	$w-1=1$	3,00	3,00	$at\sigma_w^2+a\sigma_{wt}^2+$ $+t\sigma_{wa}^2+\sigma_0^2$
Между T	$t-1=2$	506,17	253,08	$wa\sigma_t^2+w\sigma_{at}^2+$ $+a\sigma_{wt}^2+\sigma_0^2$
Взаимодействие $A\times W$	$(a-1)(w-1)=1$	40,33	40,33	$t\sigma_{wa}^2+\sigma_0^2$
Взаимодействие $W\times T$	$(w-1)(t-1)=2$	45,50	22,75	$a\sigma_{wt}^2+\sigma_0^2$
Взаимодействие $T\times A$.	$(t-1)(a-1)=2$	62,17	31,08	$w\sigma_{ta}^2+\sigma_0^2$
Остаточная	$(a-1)(t-1)(w-1)=2$	322,17	161,08	σ_0^2
Итого	$awt-1=11$	2139,67		

В этой таблице:

a, w, t — количества уровней величин A, W и T,

σ_a^2, σ_w^2, σ_t^2 — значения дисперсий вследствие общих различий соответственно между уровнями A, уровнями W и уровнями T,

σ_{aw}^2, σ_{wt}^2 и σ_{at}^2 — значения дисперсий вследствие взаимодействий соответственно между A и W, между W и T и между A и T,

σ_0^2 — остаточная дисперсия, не приписываемая ни одному из вышеуказанных источников.

Сумма квадратов, соответствующая остатку, получена как разность между общей суммой квадратов и суммами квадратов по шести другим источникам.

Смысл этих величин:

σ_a^2 — изменчивость, которая имела бы место, если бы факторы W и T поддерживались постоянными; аналогично, σ_w^2 — изменчивость, которая имела бы место, если бы A и T поддерживались постоянными; σ_t^2 — изменчивость при постоянных A и W.

Взаимодействие $A \times W$, σ_{aw}^2, служит мерой того, насколько влияние величины A зависит от значения W и, наоборот, насколько влияние величины W зависит от значения A. Аналогичен смысл и для других взаимодействий.

Остаток σ_0^2 является той частью изменчивости, которая не может быть приписана другим шести источникам. Эта величина может считаться мерой ошибки.

В настоящем эксперименте могло бы иметь место взаимодействие второго порядка $A \times W \times T$; если бы оно имело место, то соответствующая величина вошла бы в состав σ_0^2. Взаимодействие второго порядка $A \times W \times T$ выражает, насколько взаимодействие $A \times W$ зависит от уровня T, или, аналогично, насколько взаимодействие $A \times T$ зависит от уровня W, и, наконец, насколько взаимодействие $T \times W$ зависит от уровня A. Если взаимодействие $A \times W \times T$ действительно имеет место, то оно вызовет преувеличенную оценку ошибки, так что эффекты, возможно, не достигнут значимости. Подобные обстоятельства рассмотрены ниже, в главе XII, (E). Пока же примем, что взаимодействия $A \times W \times T$ не существует или по крайней мере оно невелико.

Теперь следует применить проверку значимости существования указанных величин. Если, например, σ_{ta}^2 должна существовать, то $w\sigma_{ta}^2 + \sigma_0^2$ должно значимо превосходить σ_0^2; другими словами, значение среднего квадратов взаимодействия $T \times A$ должно быть значимо больше остаточного значения среднего квадратов.

В нашей таблице значения средних квадратов для W, $A \times W$, $W \times T$ и $T \times A$ меньше остаточного значения среднего квадратов;

поэтому σ_w^2, σ_{aw}^2, σ_{wt}^2 и σ_{ta}^2 равны нулю. Для более точного расчета эти значения средних квадратов могут быть использованы как оценки для σ_0^2.

Объединяем их суммы квадратов и числа степеней свободы:

$$\frac{(3{,}00 + 40{,}33 + 45{,}50 + 62{,}17 + 322{,}17)}{1 + 1 + 2 + 2 + 2} = 59{,}1.$$

Эта новая остаточная величина имеет 8 степеней свободы.

Рассматривая строку для T, заметим, что среднее квадратов теперь оценивает $wa\sigma_t^2 + \sigma_0^2$, поскольку σ_{at}^2 и σ_{wt}^2, как было показано, равны нулю. Для установления существования σ_t^2 сравнивают значение среднего квадратов T с новым остаточным значением. Отношение дисперсий равно $\frac{253{,}08}{59{,}1} = 4{,}3$ при числах степеней свободы $n_1 = 2$, $n_2 = 8$. Это значение близко к соответствующему уровню значимости в 5%, так что можно принять, что T действительно оказывает влияние.

Аналогично, отношение дисперсий для члена A равно $\frac{1160{,}33}{59{,}1} = 19{,}6$ при числах степеней свободы $n_1 = 1$, $n_2 = 8$, что значительно превышает значение, соответствующее уровню значимости в 1%; поэтому мы принимаем, что A оказывает влияние.

Теперь можно взять средние для каждого уровня A и T и притти к заключению, что среднее значение для A_1 равно 2,00 (переходя обратно к зависимости от первоначальной переменной); для A_2 среднее значение равно 3,97 (см. исходную таблицу). Средняя для T_1 равна 3,90, для T_2 — 2,58 и для T_3 — 2,48. Ясно, что разница между T_2 и T_3 настолько невелика, что не может считаться значимой, но разница между T_2 и T_3, с одной стороны, и T_1, с другой, довольно значительна. Можно также вывести, что разница между двумя уровнями W совершенно незначима.

Каждая из средних для трех уровней T выведена из четырех наблюдений. В единицах, в которых был проведен дисперсионный анализ, остаточная дисперсия равна 59,15; стандартное отклонение было 7,691, или, в первоначальных единицах, 0,7691. Соответственно, стандартное отклонение среднего четырех результатов будет $\frac{0{,}7691}{\sqrt{4}} = 0{,}3845$. Это — стандартное отклонение для каждого из трех вышеуказанных средних, взятых для трех уровней T. Стандартное отклонение разности между любой парой равно $0{,}3845 \cdot \sqrt{2} = 0{,}5429$.

Рассматривая таким же путем средние для двух уровней A выводим, что каждое из них имеет стандартное отклонение

$\frac{0{,}7691}{\sqrt{6}}=0{,}314$; стандартная ошибка разности между этими двумя средними, равной $3{,}97-2{,}0=1{,}97$, составляет $0{,}314\cdot\sqrt{2}=0{,}444$. Остаточная дисперсия, на которой основана эта стандартная ошибка, имеет 8 степеней свободы; величина t при 5-процентном уровне значимости равна 2,31 (см. табл. I в приложении). В соответствии с этим, 95-процентные доверительные пределы для разности между A_1 и A_2 составляют $1{,}97 \pm 2{,}31 \times 0{,}444$, т. е. 3,00 и 0,94.

В этом примере, установив значимость определенных эффектов, мы приступили к расчету средних для каждого уровня этих эффектов с целью показать характер их воздействия. Каждый фактор и его уровни имели точный смысл; сравнивая A_1 с A_2, мы знали, что сравнивался один растворитель с другим. Однако в некоторых типах анализа разница между отдельными уровнями того или иного фактора может не быть наглядной. Так, фактором может быть рабочая смена с тремя уровнями; или это может быть календарный день работы при любом числе уровней; в другом случае это может быть станок или группа считаемых подобными станков. При наличии фактора такого типа, когда различие между его уровнями довольно неопределенно, часто предпочтительнее рассчитать действительную величину дисперсии, принадлежащей каждому из членов.

(E) АНАЛИЗ С ЧЕТЫРЬМЯ ФАКТОРАМИ

(I) В предыдущем разделе рассматривался пример, в котором было g сортов полимера, получавшихся от s поставщиков и обрабатывавшихся при каждой из t температур. Если предположить, что каждый из поставщиков полимера получает сырье из r источников, возникает четвертый основной фактор R. Теперь имеется всего $gtsr$ индивидуумов; каждый из них однозначно определяется заданными значениями факторов. Важно при этом, чтобы каждый поставщик получал сырье из каждого источника.

Компоненты дисперсии и соответствующие числа степеней свободы приведены в табл. 11.7.

Проверки значимости могут быть произведены по компонентам дисперсии.

(II) Ниже приводятся данные эксперимента с четырьмя факторами; каждый фактор имел два уровня. Эксперимент относился к процессу очистки продукта; бо́льшие значения зависимой переменной соответствуют худшему качеству продукта. Процесс заключался в том, что сырой материал мог подвергаться холодной или горячей промывке (H_1 или H_2), затем мог подвергаться или не подвергаться кипячению в течение известного периода

ТАБЛИЦА 11.7

Источники дисперсии	Количества степеней свободы	Компоненты дисперсии
G	$g-1$	$\sigma_0^2+r\sigma_{gts}^2+t\sigma_{srg}^2+s\sigma_{rgt}^2+sr\sigma_{gt}^2+$ $+ts\sigma_{rg}^2+rt\sigma_{gs}^2+tsr\sigma_g^2$
T	$t-1$	$\sigma_0^2+r\sigma_{gts}^2+g\sigma_{tsr}^2+s\sigma_{rgt}^2+sr\sigma_{gt}^2+$ $+rg\sigma_{st}^2+gs\sigma_{rt}^2+gsr\sigma_t^2$
S	$s-1$	$\sigma_0^2+r\sigma_{gts}^2+g\sigma_{tsr}^2+t\sigma_{srg}^2+rg\sigma_{ts}^2+$ $+gt\sigma_{sr}^2+rt\sigma_{gs}^2+gtr\sigma_s^2$
R	$r-1$	$\sigma_0^2+g\sigma_{tsr}^2+s\sigma_{rgt}^2+t\sigma_{srg}^2+gt\sigma_{sr}^2+$ $+ts\sigma_{rg}^2+gs\sigma_{rt}^2+sgt\sigma_r^2$
$G\times T$	$(g-1)(t-1)$	$\sigma_0^2+s\sigma_{rgt}^2+r\sigma_{gts}^2+sr\sigma_{gt}^2$
$T\times S$	$(t-1)(s-1)$	$\sigma_0^2+r\sigma_{gts}^2+g\sigma_{tsr}^2+rg\sigma_{ts}^2$
$S\times R$	$(s-1)(r-1)$	$\sigma_0^2+g\sigma_{tsr}^2+t\sigma_{srg}^2+gt\sigma_{sr}^2$
$R\times G$	$(r-1)(g-1)$	$\sigma_0^2+t\sigma_{srg}^2+s\sigma_{rgt}^2+ts\sigma_{rg}^2$
$G\times S$	$(g-1)(s-1)$	$\sigma_0^2+r\sigma_{gts}^2+t\sigma_{srg}^2+rt\sigma_{gs}^2$
$T\times R$	$(t-1)(r-1)$	$\sigma_0^2+g\sigma_{tsr}^2+s\sigma_{rgt}^2+gs\sigma_{rt}^2$
$G\times T\times S$	$(g-1)(t-1)(s-1)$	$\sigma_0^2+r\sigma_{gts}^2$
$T\times S\times R$	$(t-1)(s-1)(r-1)$	$\sigma_0^2+g\sigma_{tsr}^2$
$S\times R\times G$	$(s-1)(r-1)(g-1)$	$\sigma_0^2+t\sigma_{srg}^2$
$R\times G\times T$	$(r-1)(g-1)(t-1)$	$\sigma_0^2+s\sigma_{rgt}^2$
Остаточная	$(g-1)(t-1)(s-1)(r-1)$	σ_0^2
Итого . .	$gtsr-1$	

времени (B_2 или B_1); далее материал растворялся, фильтровался и осаждался из одного, из двух растворителей (S_1 или S_2) и, наконец, подвергался опять холодной или горячей промывке (W_1 или W_2).

S_1								S_2							
W_1				W_2				W_1				W_2			
B_1		B_2		B_1		B_2		B_1		B_2		B_1		B_2	
H_1	H_2	H_1	H_2	H_1	H_2	H_1	H_2	H_1	H_2	H_1	H_2	H_1	H_2	H_1	H_2
4,2	2,7	1,4	1,3	3,3	3,5	1,3	1,9	2,2	3,1	4,0	4,1	5,0	3,0	2,2	2,6
17	2	—11	—12	8	10	—12	—6	—3	6	15	16	25	5	—3	1

При проведении дисперсионного анализа из всех данных было вычтено 2,5 в целях упрощения вычислений; для того чтобы освободиться от дробей, результаты были умножены на 10 и получили вид, записанный в нижней строке таблицы. В результате этого преобразования все суммы квадратов оказались в $10^2 = 100$ раз больше истинных.

Далее поступаем так*:

(1) Составляем по очереди четыре таблицы, каждая из которых получается путем суммирования по одной из четырех переменных.

Таблица, полученная путем суммирования по S, имеет вид:

W_1				W_2			
B_1		B_2		B_1		B_2	
H_1	H_2	H_1	H_2	H_1	H_2	H_1	H_2
14	8	4	4	33	15	—15	—5

Таблица, полученная путем суммирования по W:

S_1				S_2			
B_1		B_2		B_1		B_2	
H_1	H_2	H_1	H_2	H_1	H_2	H_1	H_2
25	12	—23	—18	22	11	12	17

Остальные две таблицы „с тремя входами" получаются аналогичным способом, путем суммирования по B и H.

* Иной метод расчетов дисперсионного анализа для эксперимента „типа 2^n" см. гл. XIV, (F).

(2) Составляем шесть таблиц путем суммирования каждый раз по двум переменным (таблицы с двумя входами):

СУММИРОВАНИЕ ПО H и S:

	B_1	B_2	Итого
W_1	22	8	30
W_2	48	—20	28
Итого . . .	70	—12	58

СУММИРОВАНИЕ ПО S и B:

	H_1	H_2	Итого
W_1	18	12	30
W_2	18	10	28
Итого . . .	36	22	58

СУММИРОВАНИЕ ПО H и W:

	B_1	B_2	Итого
S_1	37	—41	— 4
S_2	33	29	62
Итого . . .	70	—12	58

СУММИРОВАНИЕ ПО B и W:

	H_1	H_2	Итого
S_1	2	— 6	— 4
S_2	34	28	62
Итого . . .	36	22	58

СУММИРОВАНИЕ ПО B и H:

	W_1	W_2	Итого
S_1	— 4	0	— 4
S_2	34	23	62
Итого . . .	30	28	58

СУММИРОВАНИЕ ПО S и W:

	H_1	H_2	Итого
B_1	47	23	70
B_2	—11	— 1	—12
Итого . . .	36	22	58

(3) Возводим общий итог в квадрат и делим на общее число наблюдений:

$$\frac{58^2}{16} = 210{,}25.$$

Этим числом придется много раз пользоваться; назовем его „корректирующим фактором", поскольку оно корректирует все суммы квадратов, являющиеся отклонениями от нуля, а не от среднего.

(4) Для получения суммы квадратов для основного эффекта фактора S возводим в квадрат итоги по S, складываем результаты и делим сумму на число первоначальных индивидуумов, входящих в каждый итог по S; из частного вычитаем корректирующий фактор:

$$\frac{(-4)^2+62^2}{8}-210{,}25=272{,}25.$$

Аналогичным способом получаются и другие основные эффекты:

Для W: $$\frac{30^2+28^2}{8}-210{,}25=0{,}25.$$

Для B: $$\frac{70^2+(-12)^2}{8}-210{,}25=420{,}25.$$

Для H: $$\frac{36^2+22^2}{8}-210{,}25=12{,}25.$$

(5) Взаимодействия первого порядка являются мерой того, насколько влияние (эффект) одного фактора зависит от значения другого. Например, существуют три взаимодействия первого порядка, включающие величину S, а именно: $S\times W$, $S\times B$ и $S\times H$.

Сумма квадратов, соответствующая первому из указанных взаимодействий $S\times W$, получается путем возведения в квадрат индивидуумов в таблице с двумя входами для S и W, сложения этих квадратов и последующего деления суммы на число первоначальных индивидуумов, производящих индивидуумы данной таблицы; из частного вычитаются корректирующий фактор и суммы квадратов для S и W:

$$\frac{(-4)^2+0^2+34^2+28^2}{4}-210{,}25-272{,}25-0{,}25=6{,}25.$$

Аналогично, сумма квадратов для взаимодействия $B\times H$ равна:

$$\frac{47^2+23^2+(-11)^2+(-1)^2}{4}-210{,}25-420{,}25-12{,}25=72{,}25.$$

Таким же путем выводятся и другие взаимодействия первого порядка.

(6) Очевидно возможно, что, например, эффект взаимодействия $B\times H$ зависит от значения S или от значения W. В первом случае это будет взаимодействие $S\times B\times H$, во втором случае — $W\times B\times H$. Такие взаимодействия называются взаимодействиями второго порядка.

Суммы квадратов их выводятся из таблиц с тремя входами.

Так, для вывода суммы квадратов для взаимодействия $S\times B\times H$ следует сперва возвести в квадрат индивидуумы в таблице $S\times B\times H$, сложить эти квадраты и разделить сумму на число первоначальных индивидуумов, производящих каждый индивидуум этой таблицы; из частного вычитается корректирующий фактор:

$$\frac{25^2+12^2+(-23)^2+(-18)^2+22^2+11^2+12^2+17^2}{2}-210{,}25=1119{,}75.$$

Из этого числа затем вычитаются суммы квадратов всех основных эффектов и всех взаимодействий первого порядка, не включающих оставшуюся переменную. Так, в случае определения суммы квадратов для взаимодействия $S\times B\times H$ оставшейся переменной является W; поэтому из числа 1119,75 вычитаются суммы квадратов для S, B, H, $S\times B$, $B\times H$ и $H\times S$:

$$1119{,}75-272{,}25-420{,}25-12{,}25-342{,}25-72{,}25-0{,}25=0{,}25.$$

Таким же способом получаются прочие взаимодействия второго порядка.

(7) Общая сумма квадратов является разностью между суммой квадратов первоначальных индивидуумов и корректирующим фактором:

$$17^2+2^2+(-11)^2+\ldots+(-3)^2+1^2-210{,}25=1877{,}75.$$

(8) Остаточная сумма квадратов получается путем вычитания из общей суммы квадратов сумм квадратов четырех основных эффектов, шести взаимодействий первого порядка и четырех взаимодействий второго порядка.

Таким путем получается полное исследование, приведенное в табл. 11.8. Числа степеней свободы определяются так: если обозначить строчными буквами количества уровней, которые имеют факторы, обозначенные соответствующими прописными буквами, то числа степеней свободы будут таковы:

(a) для основных эффектов, например, для B: $b-1$,

(b) для взаимодействий первого порядка, например, для $H\times B$: $(h-1)(b-1)$,

(c) для взаимодействий второго порядка, например, для $H\times B\times W$: $(h-1)(b-1)(w-1)$,

(d) для остатка: $(h-1)(b-1)(w-1)\ (s-1)$,

(e) для итога: $hbws-1$.

В данном примере все числа степеней свободы, кроме итогового, равны 1*. Поэтому графа для чисел степеней свободы в таблице опущена. Поскольку по той же причине суммы квадратов оказываются равными соответственным значениям средних квадратов, графа для последних также опущена.

ТАБЛИЦА 11.8

Источник дисперсии	Суммы квадратов
B	420,25
H	12,25
S	272,25
W	0,25
$B\times H$	72,25
$H\times S$	0,25
$S\times W$	6,25
$W\times B$	182,25
$W\times H$	0,25
$B\times S$	342,25
$B\times H\times S$	0,25
$H\times S\times W$	156,25
$S\times W\times B$	272,25
$W\times B\times H$	30,25
Остаточная	110,25
Итого	1877,75

Из рассмотрения данных этой таблицы ясно, что взаимодействий $W\times B\times H$ и $B\times H\times S$ не существует. Поэтому, объединяя относящиеся к ним данные с остатком, получим новую величину последнего $\frac{140{,}75}{3}=46{,}92$ с тремя степенями свободы. Взаимодействие $H\times S\times W$ незначимо; присоединяя также его к остатку, получаем новый остаток, равный $\frac{297}{4}=74{,}25$ с четырьмя

* В этом отношении настоящий пример представляет собой особый случай. Применение этих методов расчета к более общим случаям совершенно аналогично. Формулы для подсчета числа степеней свободы только-что были приведены. Что же касается сумм квадратов, то если мы имеем эксперимент с четырьмя факторами A, B, C и D соответственно при 2, 3, 4 и 5 уровнях, то будет два итога по A, каждый итог по $3\times4\times5=60$ наблюдениям; поэтому при расчете суммы квадратов A сумма квадратов итогов по A делится на 60 до вычитания корректирующего фактора. Затем мы будем иметь пять итогов по D, каждый итог по $2\times3\times4=24$ наблюдениям, поэтому делителем для суммы квадратов по D является число 24. Что касается взаимодействий, то, например, таблица $B\times C$ имеет $3\times4=12$ входов, каждый из которых является итогом $2\times5=10$ наблюдений; соответственно, для получения суммы квадратов $B\times C$ сумма квадратов по указанным 12 входам делится на 10 до вычитания корректирующего фактора и сумм квадратов B и C.

степенями свободы. Сличенное с этим остатком взаимодействие $S\times W\times B$ не достигает уровня значимости в 5%; тем не менее оно подозрительно велико; поэтому желательно разбить эксперимент на два эксперимента с тремя факторами с целью проверки отсутствия взаимодействия факторов.

При наличии одного значимого взаимодействия второго порядка (например, $S\times W\times B$) можно сделать разбивку эксперимента тремя путями: $H\times W\times B$ для каждого уровня S, $H\times S\times W$ для каждого уровня B или же $H\times S\times B$ для каждого уровня W.

Выбор пути произволен; иногда полезно в смысле получения большей информации использовать более чем один из них. В данном случае решительно напрашивается один из путей разбивки — соответственно примененному растворителю (фактор S).

Проводим поэтому два анализа с тремя факторами — один для S_1 и другой для S_2. Результаты сведены в табл. 11.9.

ТАБЛИЦА 11.9

Источник дисперсии	Средние квадратов	
	эксперимент при $S=S_1$	эксперимент при $S=S_2$
B	760,5	2,0
H	8,0	4,5
W	2,0	4,5
$B\times H$	40,5	32,0
$H\times W$	72,0	84,5
$W\times B$	4,5	450,0
Остаточная ..	12,5	128,0
Итого	900,0	705,0

Числа степеней свободы для каждого члена, кроме итога, равны 1, так что средние квадратов равны соответствующим суммам квадратов; поэтому графы для чисел степеней свободы и сумм квадратов опущены.

Рассмотрим сперва эксперимент с S_1. Взаимодействие $W\times B$ незначимо, так же как и взаимодействие $B\times H$. Выше установлено, что взаимодействие $H\times W$ также незначимо. Ясно, что и основные эффекты H и W незначимы. Присоединяя соответствующие члены к остатку, получаем новое значение последнего, 23,2 с шестью степенями свободы. Остался лишь основной эффект B с отношением дисперсий $\frac{760{,}5}{23{,}2}=32{,}8$ при числах степеней свободы $n_1=1$, $n_2=6$. Это соответствует уровню значимости примерно в 0,1%.

Можно притти к заключению, что фактор B оказывает влияние на качество продукта; B_1 дает среднее 3,47, а B_2 — среднее 1,47. Значимость других факторов W и H не установлена; средние значения для H_1 и H_2 равны соответственно 2,55 и 2,35; средние значения для W_1 и W_2 равны соответственно 2,4 и 2,5.

Таким образом, повидимому, безразлично, какой из двух уровней H или W использовать; однако очень важно избрать правильный уровень B.

Теперь рассмотрим эксперимент с S_2. Здесь большой остаток. Взаимодействия $B \times H$ и $H \times W$ не могут быть значимыми; после их присоединения к остатку получается новый остаток, 81,5, с тремя степенями свободы. Основной эффект H не содержит члена σ^2_{wb} среди его компонентов дисперсии, но этот член входит в число компонентов дисперсии выдающегося по величине взаимодействия $W \times B$; член H незначим и также может быть присоединен к остатку; значение последнего станет теперь 62,25 при четырех степенях свободы. Сличая с этой цифрой взаимодействие $W \times B$, получается отношение дисперсий $\frac{450{,}00}{62{,}25} = 7{,}2$ при числах степеней свободы $n_1 = 1$; $n_2 = 4$. Это отношение близко к соответствующему уровню значимости в 5%.

Наличие заметного взаимодействия $W \times B$ для S_2 и его несущественность для S_1 указывает на существование взаимодействия $W \times B \times S$, которые мы подозревали в анализе с четырьмя факторами. Незначительность основного эффекта B для S_2 по сравнению с большим его значением для S_1 указывает на существование взаимодействия $S \times B$; анализ с четырьмя факторами также показал его наличие, но в усложненном, замаскированном виде — в виде взаимодействия $W \times B \times S$.

В данном случае мы имеем анализ с тремя факторами и со значимым взаимодействием первого порядка $W \times B$. Соответственно этому требуется разбить эксперимент с тремя факторами на два с двумя факторами — либо на эксперименты $W \times H$ для B_1 и для B_2, либо на эксперименты $B \times H$ для W_1 и для W_2.

ТАБЛИЦА 11.10

Источник дисперсии	Суммы квадратов	
	$B = B_1$	$B = B_2$
H	30,25	6,25
W	182,25	272,25
Остаток	210,25	2,25
Итого	422,75	280,75

Результаты разбивки первым способом сведены в табл. 11.10. Здесь опять-таки числа степеней свободы для каждого члена, кроме итога, равны 1; соответствующая графа, а также и графа значений средних квадратов поэтому опущены.

Рассматривая результаты эксперимента для B_2, видим, что член H незначим и может быть присоединен к остатку; значение остатка тогда станет $\frac{8{,}5}{2} = 4{,}25$ при двух степенях свободы. Отношение дисперсий для члена W будет равно $\frac{272{,}25}{4{,}25} = 64$ при $n_1 = 1$, $n_2 = 2$; оно более значимо, чем соответствующее уровню значимости 5%.

Таким образом, можно притти к заключению, что при применении растворителя S_2 и операции B_2 (кипячение) уровень величины H безразличен; но важно избрать правильный уровень величины W, так как средние значения W_1 и W_2 равны соответственно 4,05 и 2,40.

Обращаясь теперь к эксперименту для B_1, мы видим наличие большого остатка, что указывает на существование взаимодействия $H \times W$. Из того факта, что его явно не существует для $B = B_2$, вытекает, что взаимодействие $H \times W$ маскируется в виде взаимодействия $B \times H \times W$. Это не было обнаружено в анализе с четырьмя факторами, но полученное здесь свидетельство весьма веско. При столь большой величине остатка члены H и W будут незначимы; однако мы можем получить представление о поведении системы путем составления таблицы $H \times W$ (табл. 11.11).

ТАБЛИЦА 11.11

$S = S_2$; $B = B_1$

	W_1	W_2
H_1	2,2	5,0
H_2	3,1	3,0

Эта таблица дает указание, что при S_2 и B_1 следует применять уровень W_1; возможно, что применение H_1 также имеет некоторое преимущество. Однако эти соображения нельзя рассматривать как определенные утверждения.

Настоящим завершается все, что может быть сказано о первом пути разбивки эксперимента. Обратимся теперь ко второму пути; результаты анализа сведены в табл. 11.12. Числа степеней свободы для каждого члена, кроме итога, равны 1, так что соответствующая графа, а также графа значений средних квадратов опущены.

ТАБЛИЦА 11.12

Источник дисперсии	Суммы квадратов	
	$W=W_1$	$W=W_2$
H	25	64
B	196	256
Остаточная . .	16	144
Итого . . .	237	464

Рассматривая эксперимент для W_1, видим, что член H незначим. Присоединяя его к остатку, получаем новый остаток: $\frac{41}{2}=20{,}5$ при двух степенях свободы. Отношение дисперсий для члена B равно $\frac{196}{20{,}5}=9{,}55$ при $n_1=1$, $n_2=2$. Это более значимо, чем требуется для уровня в 20%, но менее значимо, чем требуется для уровня в 5%. Однако мы можем предварительно считать этот эффект реальным.

Таким образом, можно притти к заключению, что, применяя растворитель S_2 и операцию W_1, уровень H безразличен; но важно избрать правильный уровень B, поскольку средние значения B_1 и B_2 равны соответственно 2,65 и 4,05.

В эксперименте для W_2 имеется большой остаток; поэтому нельзя ожидать, чтобы члены H и B были значимы. Тем не менее можно получить представление о поведении системы с помощью таблицы $B\times H$:

$S=S_2$; $W=W_2$

	B_1	B_2
H_1	5,0	2,2
H_2	3,0	2,6

Эта таблица указывает, что при S_2 и W_2 следует применять процесс B_2, и возможно, что лучше применять H_1.

Посмотрим теперь, не приводят ли оба пути разбивки эксперимента к противоречивым заключениям. Разбивка по B_1 и B_2 указывает, что для получения низкой цифры примесей следует совместно с операцией B_2 применять операцию W_2; но с операцией B_1 следует применять W_1. Разбивка по W_1 и W_2 указала, что совместно с W_1 следует применять B_1, а с W_2 применять B_2. Что же касается операции H, то разбивка по B_1 и B_2 не принесла достоверных результатов, но разбивка по W_1 и W_2 указала (качественно, без точной проверки значимости), что для получения низких цифр примесей следует применять H_1. Таким образом, оба пути разбивки приводят к одинаковым заключениям.

В табл. 11.13 показан эффект взаимодействия $W\times B$ при $S=S_2$.

ТАБЛИЦА 11.13

$S=S_2$ (при среднем значении H)

	B_1	B_2
W_1	2,65	4,05
W_2	4,0	2,4

Эта таблица ясно показывает, что для получения низкой цифры примесей следует применять совместно с обработкой B_1 обработку W_1; если же применяют B_2, то следует совместно с этим ввести обработку W_2. Нет „лучшего" уровня ни для B, ни для W; „лучшие" уровни зависят от уровней другого фактора. Точно так же растворитель S_2 ведет себя в этом отношении отлично от S_1, для которого обработка B_2 была всегда лучшей, независимо от значений H или W. Указанное выше взаимодействие было типа $B\times W$; поскольку оно имеет место для S_2, но не для S_1, можно сказать, что существует взаимодействие $S\times W\times B$.

Интересно отметить ошибку, к которой можно было бы притти при проведении этого испытания не факторным, а классическим приемом. В классическом приеме мы исследовали бы каждый фактор при одном уровне. В неудачном случае мы исследовали бы B при W_2 и W при B_1. Из первого испытания мы решили бы, что для получения низкой цифры примесей следует применять B_2, а из второго испытания — что следует применять W_1; в результате мы избрали бы комбинацию W_1B_2, что было бы совершенно ошибочным выводом.

Итак, резюмируем кратко выводы проведенного длительного исследования:

(1) При применении растворителя S_1:

(a) безразлично, применять ли H_1 или H_2,

(b) безразлично, применять ли W_1 или W_2,

(c) желательно применять процесс B_2, дающий среднее примесей 1,47 по сравнению с 3,42 для процесса B_1.

(2) При применении растворителя S_2:

(a) безразлично, применять ли H_1 или H_2,

(b) применяя B_1, следует применять W_1; тогда получим для содержания примесей цифру 2,65,

(c) применяя B_2, следует применять W_2; в этом случае для содержания примесей имеем цифру 2,4.

Эти заключения получены с достаточной строгостью, без догадок; они дают полную картину поведения системы.

Заметим, что данное испытание было усложнено тем, что имелись значимые взаимодействия второго порядка. Испытание было бы проще, если бы значимыми были только основные эффекты или взаимодействия первого порядка. При наличии взаимодействий второго порядка поведение системы в целом усложняется; неизбежно должен усложниться и анализ.

Заметим также, что при значительных взаимодействиях факторный прием испытания теряет свое превосходство над классическим в отношении продуктивности; фактически он сводится к серии классических испытаний. Но дело в том, что, если не применить факторного приема, эти взаимодействия не могли бы быть обнаружены.

Эти вопросы рассмотрены далее, в главе XII, (D); пока же следует иметь в виду, что таблицы 11.10 и 11.12 должны считаться неудовлетворительными из-за недостаточности чисел степеней свободы. Когда взаимодействия высших порядков значимы, требуется больше экспериментов, чем было сделано в данном случае.

(F) АНАЛИЗ С ПЯТЬЮ ФАКТОРАМИ

(I) В предыдущем разделе рассматривался пример с четырьмя факторами: источник сырья (R), поставщик полимера (S), сорт полимера (G) и температура обработки (T). Допустим теперь, что имеется еще пятый фактор — давление при обработке (P).

Процедура анализа с пятью факторами может быть легко обобщена на основании уже рассмотренных случаев анализа с двумя, тремя и четырьмя факторами:

(1) Суммируют по каждой из переменной поочередно. Получится пять таблиц с четырьмя переменными каждая.

(2) Суммируют по двум переменным каждый раз. Получится десять таблиц с тремя переменными каждая.

(3) Суммируют по трем переменным каждый раз. Получится десять таблиц с двумя переменными каждая.

(4) Подсчитывают итоги по строкам в таблицах (3). Эта операция эквивалентна суммированию по четырем переменным каждый раз.

1332

(5) Суммируют по пяти переменным сразу, т. е. получают общий итог.

(6) Суммируют квадраты всех цифр в таблицах с (1) до (5); каждая сумма квадратов делится на число первоначальных индивидуумов, производивших индивидуумы соответствующей таблицы. Например, в таблице класса (3), образованной путем суммирования по G, T и S, т. е. в таблице для R и P делителем будет gts.

(7) Основные эффекты выражаются как результаты вычитания (5) из каждой суммы квадратов в (4) (оба числа должны быть разделены на соответствующие делители).

(8) Для определения взаимодействия, например, $G\times T\times S$ берем таблицу для G, T и S (т. е. таблицу класса (2), суммированную по R и P) и определяем ее сумму квадратов, разделенную на соответствующий делитель (в данном случае rp); отсюда вычитаются: (a) корректирующий фактор (5), разделенный на его делитель $gtsrp$, (b) суммы квадратов для G, T, S, $S\times G$, $G\times T$ и $T\times S$ (уже выведенные ранее).

Рассматривая компоненты дисперсии в анализе с четырьмя факторами (табл. 11.7) и сравнивая их с табл. 11.2 для трех факторов и с табл. 11.1 для двух факторов, легко усмотреть, как будут образовываться компоненты дисперсии для анализа с пятью факторами:

(a) Взаимодействие высшего порядка, рассматриваемое как остаток, σ_0^2, не имеет коэффициентов.

(b) Взаимодействие следующего по старшинству порядка содержит величину σ_0^2 и компоненту, соответствующую взаимодействию; эта компонента имеет в качестве коэффициента буквенные символы, *не* входящие в индекс взаимодействия. Например, компоненты для $G\times T\times S\times R$ таковы: $p\sigma_{gtsr}^2+\sigma_0^2$.

(c) Взаимодействие следующего по старшинству порядка содержит σ_0^2, компоненту, соответствующую данному взаимодействию и всем промежуточным взаимодействиям, индексы которых включают индекс рассматриваемого взаимодействия. Общее правило для коэффициентов остается в силе повсюду: коэффициентами являются буквенные символы, не входящие в индексы. Так, компоненты $G\times T\times S$ имеют вид:

$$rp\sigma_{gts}^2+r\sigma_{gtsp}^2+p\sigma_{gtsr}^2+\sigma_0^2.$$

(d) Взаимодействия первого порядка имеют компоненты дисперсии, составленные по тем же общим правилам. Так, компоненты $S\times R$ имеют вид:

$$gtp\sigma_{sr}^2+gt\sigma_{psr}^2+tp\sigma_{gsr}^2+pg\sigma_{srt}^2+g\sigma_{psrt}^2+$$
$$+t\sigma_{gsrp}^2+p\sigma_{gsrt}^2+\sigma_0^2.$$

9 Зак. 1691

(e) Выражения для основных эффектов выводятся аналогичным путем. Например, выражение для основного эффекта S включает σ_s^2, взаимодействия первого порядка от S: σ_{sg}^2, σ_{sr}^2, σ_{sp}^2 и σ_{st}^2; взаимодействия второго порядка от S: σ_{sgr}^2, σ_{sgp}^2, σ_{spr}^2, σ_{stp}^2, σ_{str}^2 и σ_{stg}^2; взаимодействия третьего порядка от S: σ_{rgst}^2, σ_{strp}^2, σ_{srpg}^2, σ_{spgt}^2 и остаток σ_0^2. Все перечисленные компоненты имеют коэффициенты, составленные согласно вышеуказанному общему правилу; так, σ_s^2 получает коэффициент $prgt$ и т. д.

(II) Испытание с пятью факторами проводилось в процессе очистки некоторого кристаллического продукта; все факторы имели по два уровня. Испытывалось прежде всего количество вещества на цикл (T); один из уровней в действительности соответствовал двойной концентрации по сравнению с другим. Далее, испытывалась продолжительность растворения вещества (D) в растворителе; при этом одно из значений последней в четыре раза превосходило другое. Время имело значение, так как растворение делалось в горячем растворе и, поскольку образование примесей являлось автокаталитическим, продолжительное воздействие высокой температуры могло оказать губительный эффект. Третий фактор W представлял две немного отличающиеся друг от друга жидкости, применяемые в следующей стадии процесса; в последней были две разные скорости размешивания (S) и два периода времени (G), в течение которых делалось осаждение чистых кристаллов путем дестилляции паром.

Первоначальные данные приведены в табл. 11.14. Десятичные знаки не учитывались; большие числа соответствуют худшему качеству продукта.

ТАБЛИЦА 11.14

Первоначальные данные испытания с пятью факторами процесса очистки кристаллов

		W_1				W_2			
		G_1		G_2		G_1		G_2	
		D_1	D_2	D_1	D_2	D_1	D_2	D_1	D_2
T_1	S_1	138	108	113	95	103	137	106	116
	S_2	116	119	106	114	122	122	128	148
T_2	S_1	155	170	131	172	146	152	135	144
	S_2	133	172	138	181	150	156	124	146

Поскольку все факторы имеют по два уровня, то в дисперсионном анализе числа степеней свободы для всех членов равны единице; следовательно, значения средних квадратов равны соответствующим суммам квадратов. В табл. 11.15, содержащей дисперсионный анализ, взаимодействия четвертого порядка не включены; они объединены с остатком, число степеней свободы которого равно поэтому $1+5=6$. Строго говоря, эти взаимодействия следовало бы подсчитать; однако на практике трудно считать их сколько-нибудь важными; практический интерес представляют эффекты, приведенные в таблице. Цифры в скобках (31) и (6) в табл. 11.15, поставленные после итога и остатка, являются соответствующими числами степеней свободы.

ТАБЛИЦА 11.15

Дисперсионный анализ для испытания с пятью факторами процесса очистки кристаллов

Источники дисперсии	Суммы квадратов
S	91,1
G	325,1
W	21,1
T	8256,1
D	1352,0
$S\times G$	264,5
$G\times W$	12,6
$W\times T$	924,5
$T\times D$	741,2
$D\times S$	171,1
$G\times T$	18,0
$W\times D$	1,2
$T\times S$	128,1
$D\times G$	120,1
$W\times S$	112,5
$S\times G\times W$	21,1
$G\times W\times T$	406,1
$W\times T\times D$	1200,4
$T\times D\times S$	0,4
$D\times S\times G$	24,6
$S\times W\times T$	91,1
$S\times G\times T$	91,1
$T\times D\times G$	40,5
$G\times W\times D$	31,9
$D\times W\times S$	288,0
Остаток (6)	661,6
Итого (31) . . .	15396,0

Видно, что все взаимодействия второго порядка, кроме $W\times T\times D$, незначимы, но взаимодействие $W\times T\times D$ велико.

При проверке $W\times T\times D$ следует иметь в виду, что его компоненты дисперсии включены в число компонент дисперсии величин $W\times T$, $T\times D$, $D\times W$, W, T и D; поэтому ни один из перечисленных членов не может быть менее, чем $W\times T\times D$. Однако этому условию удовлетворяют только основные эффекты T и D.

На этом основании мы объединяем другие члены с $W\times T\times D$ и получаем новую оценку последнего, $\frac{2884{,}5}{5}=577{,}7$, при пяти степенях свободы. Сравнивая эту величину с остатком $\frac{1656{,}4}{15}=110{,}4$, имеющим 15 степеней свободы, получаем отношение дисперсий, равное 5,2 при $n_1=5$ и $n_2=15$ и соответствующее вероятности в 1%. Поскольку T является наиболее заметным из основных эффектов, казалось бы наиболее целесообразным разбить анализ с пятью факторами на два анализа с четырьмя факторами, каждый из которых относится к определенному уровню T. Это сделано в табл. 11.16.

ТАБЛИЦА 11.16

Дисперсионный анализ для двух испытаний с четырьмя факторами процесса очистки кристаллов

Источники дисперсии	Суммы квадратов	
	T_1	T_2
W	333,06	612,56
S	217,56	1,56
G	95,06	248,06
D	45,56	2047,56
$W \times S$	203,07	0,57
$S \times G$	333,07	22,57
$G \times D$	10,57	150,07
$D \times S$	76,57	95,07
$W \times D$	637,57	564,07
$G \times W$	280,57	138,07
$W \times S \times G$	45,55	175,55
$G \times S \times D$	85,55	5,08
$S \times D \times W$	430,55	10,55
$W \times G \times D$	27,55	7,55
Остаточная . .	162,58	76,55
Итого	2984,44	4155,44

Рассматривая члены для T_1, начиная с остатка и идя по таблице вверх, видим, что ни один из них незначим. T_1 соответствует применению разведенного раствора. В этих условиях мы можем выбирать для четырех факторов любые уровни, удобные практически, зная, что это не окажет вредного влияния на чистоту продукта.

Возвращаясь к табл. 11.15, замечаем, что основной эффект T, сравненный с $W \times T \times D$ (все взаимодействия первого порядка, включающие T, незначимы), значим при уровне 5%, даже ближе к уровню 1%. Величина среднего для T_1 ниже, чем для T_2; поэтому для получения более чистого продукта следует избрать T_1. С другой стороны, загрузка производственного оборудования будет при T_1 составлять лишь половину той, которая возможна при T_2; пропускная способность завода окажется сниженной в два раза; таким образом, может оказаться необходимым применить T_2, если это вообще возможно.

Обращаясь к табл. 11.16 и рассматривая анализ для T_2, находим, что взаимодействие $W \times D$ имеет уровень значимости 5%; вследствие этого проведены анализы для каждого уровня D в отдельности (табл. 11.17).

В анализе для D_1 все члены незначимы; возможным исключением является G; средние значения G_1 и G_2 равны соответственно 146 и 132. В анализе для D_2, с другой стороны, велико только W, средние значения которого W_1 и W_2 равны соответственно 173,8 и 149,5.

ТАБЛИЦА 11.17

Источники дисперсии	Суммы квадратов	
	D_1	D_2
G	392,0	6,15
W	0,5	1176,15
S	60,5	36,15
$G \times W$	40,5	105,10
$W \times S$	8,0	3,10
$S \times G$	24,5	3,10
Остаточная . . .	242,0	10,15
Итого	768,0	1339,90

Вспомним теперь, что при применении T_1 и любых значений D, W, S и G получаемое среднее равно 118,2, так что, если вопрос ставится о чистоте продукта, предпочтительнее применять T_1, если же применяется T_2, следует предпочесть продолжительность растворения D_1; при этом будет применяться любое значение G, W и S, и чистота продукта выразится цифрой 139. В случае необходимости применения D_2 можно использовать любое значение G и S, но очень важно применять W_2, что даст значение среднего 149,5.

Очевидно, что столь сложное поведение системы трудно было бы проследить без применения факторного анализа. Кроме того, дисперсионный анализ был абсолютно необходим для нахождения значимых эффектов и взаимодействий. Это было бы совершенно невозможно сделать путем качественного просмотра данных в табл. 11.14 *.

(G) НЕПОЛНЫЙ АНАЛИЗ С ДВУМЯ ФАКТОРАМИ. ОДИН ИЗ ФАКТОРОВ С ПОВТОРЕНИЯМИ

В примере анализа с двумя факторами (в разделе (C)) рассматривались g сортов материала, обрабатываемого при t различных значениях температуры. Предположим, что из каждого g сортов материала делается t партий продукта при обработке на одной и той же температуре. При этих условиях первая одиночная партия 1-го сорта имеет не больше соответствия с первой партией 2-го сорта, чем со второй партией 2-го сорта, и т. д.

* Пример промышленного эксперимента с семью факторами приведен в статье R. L. Cunningham and J. Ansel Anderson, *Cereal Chemistry*, XX (1943), 482.

Анализ с двумя факторами теперь становится анализом с одним фактором (сорта) и с t повторениями. Это влечет за собой объединение числа степеней свободы и суммы квадратов главного эффекта T в табл. 11.1 с остаточными значениями этих величин. Результат получается в виде табл. 11.18:

ТАБЛИЦА 11.18:

Источники дисперсии	Числа степеней свободы	Компоненты дисперсии
G	$g-1$	$t\sigma_g^2+\sigma_0^2$
Остаток (в пределах значений G)	$(t-1)+(t-1)(g-1)=$ $=g(t-1)$	σ_0^2
Итого . . .	$gt-1$	

Этот род анализа рассмотрен подробно в главе VII, (С).

(H) НЕПОЛНЫЙ АНАЛИЗ С ТРЕМЯ ФАКТОРАМИ. ДВА ФАКТОРА С ПОВТОРЕНИЯМИ

В разобранном ранее примере анализа с тремя факторами было S поставщиков, каждый из которых поставлял g сортов полимера; каждый сорт обрабатывался при одной из t различных температур. Допустим теперь, что t партий обрабатывались при одной и той же температуре; другими словами, теперь налицо два фактора G и S с t повторениями. Пусть из табл. 11.2 вытекает, что взаимодействия $G\times T$ и $T\times S$, а также основной эффект T отсутствуют. Причина отсутствия указанных взаимодействий может быть, например, в том, что основной эффект G, по определению, один и тот же для всех повторений, и те или иные отклонения в его значении могут быть отождествлены с погрешностями, влияние которых учитывается в остаточных значениях. Тогда получается табл. 11.19.

Рассмотрим задачу, в которой требуется исследовать один из факторов в некотором производственном процессе; нужно, например, определить эффект этого фактора при трех его уровнях: F_1, F_2 и F_3. Допустим, что в одной из стадий производства материал подвергается специальному процессу обработки, и известно (или подозревается), что это может вызвать увеличение изменчивости. Далее будет необходимо исключить погрешности, возникающие из-за этого добавочного фактора.

ТАБЛИЦА 11.19

Источники дисперсии	Числа степеней свободы	Компоненты дисперсии
G	$g-1$	$\sigma_0^2+t\sigma_{gs}^2+ts\sigma_g^2$
S	$s-1$	$\sigma_0^2+t\sigma_{gs}^2+tg\sigma_s^2$
$G\times S$	$(g-1)(s-1)$	$\sigma_0^2+t\sigma_{gs}^2$
Остаточная	$(g-1)(s-1)(t-1)+(t-1)+$ $+(g-1)(t-1)+(s-1)(t-1)=$ $=gs(t-1)$	σ_0^2
Итого . . .	$gst-1$	

Изложенная ниже постановка исследования удовлетворяет этим требованиям. Сравнение средних величин F имеет силу, поскольку нарушающее закономерность влияние станков M_1, M_2 и M_3 исключается. Тот же эксперимент дает нам и имеющее силу сравнение трех станков M_1, M_2 и M_3. Данные экспериментов исследуются с помощью дисперсионного анализа по строкам и столбцам, рассмотренного ранее в главе VII, (E), и в разделе (С) настоящей главы.

	Станок		
	M_1	M_2	M_3
F_1			
F_2			
F_3			

Это исследование являлось средством исключения погрешности из-за влияния заданного фактора; конечно, порядок исследования был бы точно таким же, как и для исследования с двумя факторами, иначе говоря, мы могли быть одинаково заинтересованы в обоих факторах.

В представленном виде исследования нельзя определить, имеется ли какое-либо взаимодействие между отдельными членами, хотя мы и могли бы сравнить, например, M_1-M_2 для $F=F_1$ с M_1-M_2 для $F=F_3$. Это сравнение было бы бесцельно, потому что у нас нет оценки погрешности каждого из членов, а следовательно, и их разности. Без оценки погрешности нельзя оценить значимость какого бы то ни было видимого различия.

Мы получили бы оценку погрешности, если бы каждый эксперимент был повторен. Тогда исследование могло бы называться анализом по строкам и столбцам или анализом с двумя факторами с повторениями.

Рассмотрим данные нижеследующей таблицы:

	N_1		N_2		Итого
D_1	−12	−5	27	27	37
D_2	−29	−22	34	24	7
Итого . .	−68		112		44

Данные относятся к эксперименту с двумя факторами $N \times D$; каждый фактор имел два уровня; все наблюдения производились дважды. Проводим анализ следующим путем:

(1) Возводим в квадрат итог по каждой ячейке (по двум наблюдениям), складываем эти квадраты и делим результат на число первоначальных индивидуумов в ячейке:

$$\frac{(-12-5)^2+(-29-22)^2+(27+27)^2+(34+24)^2}{2}=4585.$$

(2) Подсчитываем итог по каждому уровню D, возводим эти итоги в квадрат и складываем; результат делим на квадрат числа индивидуумов в итоговой графе:

$$\frac{37^2+7^2}{4}=354{,}5.$$

(3) Та же операция выполняется для N:

$$\frac{(-68)^2+112^2}{4}=4292.$$

(4) Возводим в квадрат общий итог и делим на общее количество индивидуумов:

$$\frac{44^2}{8}=242.$$

(5) Возводим каждый индивидуум в квадрат и складываем эти квадраты:

$$(-12)^2+(-5)^2+27^2+\ldots+24^2=4684.$$

Составляем таблицу дисперсионного анализа:

ТАБЛИЦА 11.20

Источник дисперсии	Суммы квадратов	Числа степеней свободы	Средние квадратов	Компоненты дисперсии
Между строками (D)	(2)—(4) = 112,5	$n_1-1=1$	112,5	$n_2n_3\sigma_1^2+n_3\sigma_{12}^2+\sigma_0^2$
Между столбцами (N)	(3)—(4) = 4050	$n_2-1=1$	4050	$n_1n_3\sigma_2^2+n_3\sigma_{12}^2+\sigma_0^2$
Взаимодействие $D \times N$	180,5	$(n_1-1)(n_2-1)=1$	180,5	$n_3\sigma_{12}^2+\sigma_0^2$
Остаточная	(5)—(1) = 99	$n_1n_2(n_3-1)=4$	24,75	σ_0^2
Итого	(5)—(4) = 4442	$n_1n_2n_3-1=7$		

В этой таблице:

n_1 — число строк, n_2 — число столбцов, n_3 — число повторений;

σ_1^2 — дисперсия вследствие различий между строками;

σ_2^2 — дисперсия вследствие различий между столбцами;

σ_{12}^2 — взаимодействие между строками и столбцами;

σ_0^2 — остаточная погрешность.

Сумма квадратов, относящихся к взаимодействию между строками и столбцами, получена путем вычитания сумм квадратов других трех членов из общей суммы квадратов: (5) — (4).

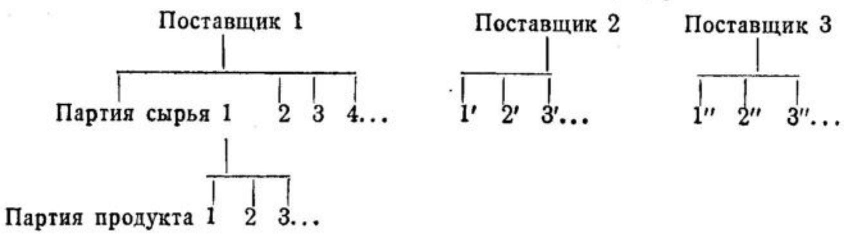
В общем случае при проверке значений средних квадратов на значимость значения средних квадратов между строками и между столбцами сравниваются со значением для взаимодействия между строками и столбцами, так как, если первые значимо превосходят последнее, единственной причиной этого может быть только, что σ_1^2 и σ_2^2 соответственно больше нуля. Аналогично, существование взаимодействия можно проверить путем сравнения его среднего квадрата с остаточным значением. В данном случае отношение дисперсий составляет $\frac{180{,}5}{24{,}75}=7{,}3$ при числах степеней свободы $n_1=1$, $n_2=4$; это близко к уровню значимости в 5%.

На основании этого следует притти к выводу, что между D и N есть взаимодействие, и следует рассмотреть данные в разбивке на эксперименты с одним фактором. При каждом уровне N можно сравнить D_1 с D_2, применяя для сравнения двух средних значений D критерий t Стьюдента. Сделав это, найдем, что различие между D_1 и D_2 незначимо при N_2, но довольно значимо при N_1. Таким же образом можно сравнить различие N_1 и N_2 при каждом уровне D; найдем, что различие между двумя уровнями N больше при D_2, чем при D_1.

Заметим, что этот анализ с двумя факторами с повторениями является на самом деле анализом с тремя факторами, содержащим все члены, характеризующие третий фактор R (основной эффект R и взаимодействия $R \times N$ и $R \times D$), объединенным с остаточными значениями.

(I) ДВАЖДЫ НЕПОЛНЫЙ АНАЛИЗ С ТРЕМЯ ФАКТОРАМИ. ОДИН ИЗ ФАКТОРОВ С ПОВТОРЕНИЯМИ ДВОЙНОГО ПОРЯДКА

(I) В предыдущем примере было s поставщиков, поставляющих g различных сортов полимера каждый; из каждого материала было изготовлено по t партий продукта, обработанных при одной и той же температуре. Предположим теперь, что те же s поставщиков вместо поставки g различных сортов полимера, поставили g номинально одинаковых партий полимера одного и того же сорта. Система иллюстрируется теперь схемой:



Партии продукта делались из всех партий сырья; на схеме показаны только сделанные из партии сырья 1 от поставщика 1. Поскольку все партии сырья номинально одного и того же сорта, нет смысла сравнивать средние всех первых партий $(1, 1', 1'', \ldots)$ со средними всех вторых партий $(2, 2', 2'', \ldots)$ и т. д. Отдельные партии сырья, поступившего от поставщика 1, могут отличаться друг от друга; то же возможно и в партиях сырья других поставщиков. Можно использовать состав партий продукта, изготовленных из каждой партии сырья, как меру погрешности, с которой могли бы сравниваться различия между партиями сырья. Заметим, что если бы первая партия продукта во всех случаях обрабатывалась при температуре t_1, вторая партия во всех случаях — при другой температуре t_2 и т. д., то получился бы неполный анализ с двумя факторами, сходный с рассмотренным выше; факторами были бы поставщик (S) и температура (R).

Дисперсия может быть разложена в данном случае на компоненты так, как показано в табл. 11.21. Табл. 11.21 выведена из табл. 11.19 путем объединения сумм квадратов и чисел степеней свободы членов для G и $G \times S$. Компоненты дисперсии, соответствующие эффекту G в табл. 11.21, по своей природе лучше отражают взаимодействие, чем основной эффект; будем ставить в качестве индекса компоненты дисперсии символ G (а не g), чтобы отметить ее несколько особый характер.

(II) Один из заводов по концентрации слабой кислоты был подвержен авариям вследствие коррозии. Данные табл. 11.22 представляют количество кислоты, полученное от ряда производственных установок до выхода их из строя. Из массы данных было

ТАБЛИЦА 11.21

Источники дисперсии	Числа степеней свободы	Компоненты дисперсии
Поставщики (S)	$s-1$	$\sigma_0^2+t\sigma_G^2+tg\sigma_s^2$
Между партиями сырья (G)	$(g-1)+(g-1)(s-1)=s(g-1)$	$\sigma_0^2+t\sigma_G^2$
По партиям сырья (остаток)	$gs(t-1)$	σ_0^2
Итого	$gst-1$	

отобрано по 7 отчетных цифр от каждого из 9 цехов; эти 9 цехов разбиты на три группы соответственно типу слабой кислоты, которая ими концентрировалась.

ТАБЛИЦА 11.22

Тип слабой кислоты	Цех	Выпуск продукции							Итого
A	1	40	49	25	30	32	25	22	223
	2	40	46	41	30	50	39	31	277
	3	38	38	32	33	42	45	44	272
B	4	16	50	40	54	80	64	59	363
	5	49	35	70	64	23	70	88	399
	6	68	20	66	36	34	45	40	309
C	7	90	90	104	62	78	88	93	605
	8	84	93	73	80	210	137	89	766
	9	152	70	122	143	72	122	113	794

Для анализа данных подсчитываются следующие величины:

(1) Сумма квадратов всех индивидуумов:

$$40^2+49^2+25^2+\ldots+122^2+113^2=339\,900.$$

(2) Сумма квадратов итогов по блокам (цехам), деленная на количество индивидуумов, относящееся к каждому цеху:

$$\frac{223^2+277^2+\ldots+794^2}{7}=310016.$$

(3) Сумма квадратов итогов по суперблокам (типам слабой кислоты), деленная на количество индивидуумов, соответствующее каждому типу слабой кислоты:

$$\frac{(223+277+272)^2+(363+399+309)^2+(605+766+794)^2}{21}=306\,202.$$

(4) Квадрат общего итога, деленный на общее число индивидуумов:

$$\frac{(223+277+272+363+\ldots+794)^2}{63}=254\ 985.$$

По ним составляется таблица дисперсионного анализа (табл. 11.23):

ТАБЛИЦА 11.23

Источники дисперсии	Числа степеней свободы	Суммы квадратов	Средние квадратов	Компоненты дисперсии
Между суперблоками (типы слабой кислоты)	$n_1-1=2$	(3)—(4)=51 217	25 609	$n_2n_3\sigma_1^2+n_3\sigma_2^2+\sigma_3^2$
Между блоками (цехами) в пределах суперблоков	$n_1(n_2-1)=6$	(2)—(3)=3 814	636	$n_3\sigma_2^2+\sigma_3^2$
Остаток, по блокам (цехам) . .	$n_1n_2(n_3-1)=54$	(1)—(2)=29 884	553	$n\sigma_3^2$
Итого . . .	$n_1n_2n_3-1=62$	(1)—(4)=84 915		

В этой таблице:

n_1 — число суперблоков;
n_2 — число блоков в суперблоке;
n_3 — число индивидуумов в блоке;
σ_1^2 — дисперсия вследствие различий между суперблоками;
σ_2^2 — дисперсия вследствие различий между блоками в пределах суперблоков;
σ_3^2 — дисперсия вследствие различий между индивидуумами по блокам.

Из последнего столбца „Компоненты дисперсии“ видно, что σ_2^2 может быть больше нуля только в случае, если $n_3\sigma_2^2+\sigma_3^2$ значимо превосходит σ_3^2. Этим мы сравниваем средние квадратов между блоками и по блокам. Аналогичным путем в общем случае сравнивают средние квадратов для суперблоков и между блоками (в пределах суперблоков); если первое из них значимо больше второго, то только потому, что σ_1^2 больше нуля.

В данном случае мы видим, что значение среднего квадратов между блоками незначимо превосходит значение для остатка; таким образом, можно принять, что $\sigma_2^2=0$ и объединить два комплекта сумм квадратов и чисел степеней свободы, чтобы получить новый остаток $\frac{33\ 698}{60}=561{,}6$ с 60 степенями свободы.

Сравнивая с этим остатком среднее квадратов между суперблоками, получаем отношение дисперсий $\frac{25\ 609}{561{,}6}=45{,}6$ при числах степеней свободы $n_1=2$, $n_2=60$. При этих числах степеней свободы и при уровне значимости в 0,1% отношение дисперсий

должно быть не ниже 7,8. Отсюда следует, что значение среднего квадрата между суперблоками гораздо более значимо, чем соответствующее уровню в 0,1%. Имея в виду, что $\sigma_2^2=0$, решаем уравнения $3\cdot 7\sigma_1^2+\sigma_3^2=25\ 609$; $\sigma_3^2=562$. Величина σ_1^2 равна тогда 1192. Таким образом: $\sigma_1^2=1192$, $\sigma_2^2=0$, $\sigma_3^2=562$.

Ценность этого анализа в том, что он показывает, что различия между цехами, обрабатывающими одну и ту же группу кислоты, не обнаруживаются и что бóльшая часть изменчивости в выпуске продукции объясняется различиями между группами кислоты.

Польза первого из этих выводов в том, что он позволяет объединить данные всех цехов, обрабатывающих одну и ту же группу кислоты, с целью какого-либо детального сравнения, например для сравнения между литейными (изготовляющими тигли), так как в некоторых случаях может быть 10 тиглей литейной A и 1 тигель литейной B в одном цехе, 1 тигель литейной A и 10 тиглей литейной C во втором цехе и т. д. Пока мы ограничиваемся сравнениями в пределах цеха, оценки различий между литейными будут очень неточны. Если же доказано, что можно объединить данные по всем цехам, обрабатывающим данную группу кислоты, сравнения окажутся гораздо более полезными.

Польза второго из вышеприведенных выводов в том, что он подчеркивает важность типа кислоты, подвергаемого концентрации. Дальнейшее развитие вывода, конечно, может делаться только на основании рассмотрения химических закономерностей.

(J) НЕПОЛНЫЙ АНАЛИЗ С ЧЕТЫРЬМЯ ФАКТОРАМИ. ТРИ ФАКТОРА С ПОВТОРЕНИЯМИ

(I) В полном анализе с четырьмя факторами было s поставщиков, получавших в свою очередь снабжение сырьем из r источников и поставлявших g сортов полимера, которые затем обрабатывались при каждой из t температур.

Если, в изменение этого, t партий продукта обрабатывались бы при одной и той же температуре, мы имели бы анализ с тремя факторами и повторениями; факторами были бы сорт полимера (G), поставщик (S) и источник сырья (R).

Дисперсионный анализ в этом случае выводится из полного анализа с четырьмя факторами, приведенного в табл. 11.8. Основного эффекта T, взаимодействий первого порядка, включающих T, и взаимодействий второго порядка, включающих T, не существует. Поэтому можно объединить их числа степеней свободы и суммы квадратов с остаточными значениями. Тогда получим табл. 11.24.

ТАБЛИЦА 11.24

Источники дисперсии	Числа степеней свободы	Компоненты дисперсии
G	$g-1$	$\sigma_0^2+t\sigma_{srg}^2+ts\sigma_{rg}^2+rt\sigma_{gs}^2+tsr\sigma_g^2$
S	$s-1$	$\sigma_0^2+t\sigma_{srg}^2+gt\sigma_{sr}^2+rt\sigma_{gs}^2+gtr\sigma_s^2$
R	$r-1$	$\sigma_0^2+t\sigma_{srg}^2+gt\sigma_{sr}^2+ts\sigma_{rg}^2+gts\sigma_r^2$
$G\times S$	$(g-1)(s-1)$	$\sigma_0^2+t\sigma_{srg}^2+rt\sigma_{gs}^2$
$S\times R$	$(s-1)(r-1)$	$\sigma_0^2+t\sigma_{srg}^2+gt\sigma_{sr}^2$
$R\times G$	$(r-1)(g-1)$	$\sigma_0^2+t\sigma_{srg}^2+ts\sigma_{rg}^2$
$G\times S\times R$	$(g-1)(s-1)(r-1)$	$\sigma_0^2+t\sigma_{srg}^2$
Остаточная	$(g-1)(t-1)(r-1)(s-1)+(r-1)(g-1)(t-1)+(t-1)(s-1)(r-1)+(g-1)(t-1)(s-1)+(t-1)(r-1)+(t-1)(s-1)+(g-1)(t-1)+(t-1)=gsr(t-1)$	σ_0^2
Итого . . .	$gsrt-1$	

(II) Допустим, что проводится эксперимент с тремя факторами, каждый этап эксперимента повторяется. Это может рассматриваться как эксперимент с четырьмя факторами; четвертым фактором служит порядок повторений; полный анализ дисперсии получается путем объединения с остаточными величинами значений всех факторов, включающих A; именно, если N, D и R являются тремя основными факторами, объединению с остатком подлежат A, $A\times N$, $A\times D$, $A\times R$, $A\times N\times D$, $A\times D\times R$ и $A\times R\times N$.

Арифметические выкладки, однако, облегчаются, если провести суммирование по всем уровням A, т. е. просуммировать по всем повторениям для каждого значения N, D и R; таким образом, получается таблица типа анализа с тремя факторами, каждое число в которой является суммой повторений.

Рассмотрим данные нижеследующей таблицы:

	N_1				N_2			
	D_1		D_2		D_1		D_2	
	R_1	R_2	R_1	R_2	R_1	R_2	R_1	R_2
A_1	−3	−1	−8	−6	8	8	10	7
A_2	−7	−6	−3	−5	−6	1	0	−3
$S(A)$	−10	−7	−11	−11	2	9	10	4

Строка $S(A)$ содержит сумму значений A_1 и A_2 для каждого значения N, D и R. Данные этой строки подчиняются описанному выше порядку анализа с тремя факторами со следующими отличиями:

(1) Во всех случаях, когда складываются квадраты, вводится дополнительный делитель — число повторений (в данном случае 2).

Таким образом, одна из трех таблиц с двумя входами имеет вид:

	D_1	D_2	Итого
N_1	−17	−22	−39
N_2	11	14	25
Итого	−6	−8	−14

Корректирующий фактор равен $\frac{(-14)^2}{8\cdot 2}=12{,}25$.

Сумма квадратов для N равна:

$$\frac{(-39)^2+25^2}{4\cdot 2}-12{,}25=256{,}0.$$

Сумма квадратов для D равна:

$$\frac{(-6)^2+(-8)^2}{4\cdot 2}-12{,}25=0{,}25.$$

Взаимодействие $N\times D$ выражается величиной:

$$\frac{(-17)^2+(-22)^2+11^2+14^2}{2\cdot 2}-12{,}25-256{,}00-0{,}25=4{,}00.$$

Аналогичным путем получаются и суммы квадратов для эффектов R, $R\times D$ и $R\times N$.

(2) Независимая оценка остатка получается как разность между суммой квадратов всех первоначальных индивидуумов и суммой квадратов величин $S(A)$, деленной на число повторений (в данном случае 2):

$$(-3)^2+(-1)^2+\ldots+(-3)^2-\frac{(-10)^2+(-7)^2+\ldots+4^2}{2}=256.$$

(3) Общая сумма квадратов определяется как разность между суммой квадратов первоначальных индивидуумов и корректирующим фактором:

$$(-3)^2+(-1)^2+\ldots+(-3)^2-12{,}25=539{,}75.$$

(4) Сумма квадратов взаимодействия $N\times D\times R$ получается как разность между общей суммой квадратов и суммой для всех остальных членов.

Результат анализа получает следующий вид:

Источники дисперсии	Суммы квадратов	Числа степеней свободы	Средние квадратов
N	256,00	$n-1=1$	256,00
D	0,25	$d-1=1$	0,25
R	1,00	$r-1=1$	1,00
$N\times D$	4,00	$(n-1)(d-1)=1$	4,00
$D\times R$	16,00	$(d-1)(r-1)=1$	16,00
$R\times N$	0,25	$(r-1)(n-1)=1$	0,25
$N\times D\times R$	6,25	$(n-1)(d-1)(r-1)=1$	6,25
Остаточная	256,00	$ndr(a-1)=8$	32,00
Итого	539,75	$rnda-1=15$	

Взаимодействие $N\times D\times R$ сравнивается с остатком и, очевидно, является незначимым. Фактически все члены, кроме основного эффекта N, незначимы и могут быть объединены с остатком; новое значение остатка будет $\frac{283{,}75}{14}=20{,}2$; основной эффект N имеет тогда величину отношения дисперсий $\frac{256{,}00}{20{,}2}=12{,}7$ при числах степеней свободы $n_1=1$, $n_2=14$; это отношение более значимо, чем соответствующее уровню значимости в 1%. Затем можно подсчитать средние значения N_1 и N_2; они равны соответственно 4,875 и 3,125.

В этом примере могло бы случиться, что повторения (члены серии R_2) оказались на другом уровне, чем первые наблюдения (члены серии R_1); причинами этого могли бы быть регулировка оборудования между двумя сериями наблюдений, поступление новой партии сырья и т. п. Дисперсия из-за этого влияния включается в остаточное значение; если она велика, то получится преувеличенное значение остатка и снижение точности эксперимента.

В связи с этим можно изъять из остатка сумму квадратов, соответствующую повторениям; для этого подсчитываются итоги для рядов: $A_1=15$; $A_2=-29$ и определяется величина:

$$\frac{15^2+(-29)^2}{8}-12{,}25=121{,}00.$$

Теперь сумма квадратов для остатка равна $256{,}00-121{,}00=135{,}00$; число его степеней свободы теперь равно $8-1=7$; значение среднего квадратов стало 19,29.

В данном случае путем изъятия суммы квадратов для рядов мы получили значительный выигрыш; весьма существенно повысилась возможность обнаружения значимости членов.

(К) ДВАЖДЫ НЕПОЛНЫЙ АНАЛИЗ С ЧЕТЫРЬМЯ ФАКТОРАМИ. ДВА ФАКТОРА С ПОВТОРЕНИЯМИ ДВОЙНОГО ПОРЯДКА

(1) Предположим теперь, что поставщики предыдущего примера вместо g различных сортов материала выработали g номинально одинаковых сортов. Как и прежде, t партий из каждого из этих сортов подвергаются обработке при одной и той же температуре.

Дисперсия может быть теперь разложена на составляющие, вызываемые различием сырья (R), различием между поставщиками (S) и взаимодействием $R\times S$; еще две составляющих показаны в табл. 11.25. Число степеней свободы и сумма квадратов для нового члена G получены путем объединения с прежними значениями для G величин $G\times S$, $S\times R$ и $G\times S\times R$. Остаток тот же, что и в табл. 11.24.

ТАБЛИЦА 11.25

Источники дисперсии	Числа степеней свободы	Компоненты дисперсии
R	$r-1$	$\sigma_0^2+t\sigma_G^2+gt\sigma_{rs}^2+gts\sigma_r^2$
S	$s-1$	$\sigma_0^2+t\sigma_G^2+gt\sigma_{rs}^2+gtr\sigma_s^2$
$R\times S$	$(r-1)(s-1)$	$\sigma_0^2+t\sigma_G^2+gt\sigma_{rs}^2$
Между G в пределах R и S	$(g-1)+(g-1)(r-1)+(g-1)(s-1)+(g-1)(r-1)(s-1)=rs(g-1)$	$\sigma_0^2+t\sigma_G^2$
В пределах G (остаток)	$gsr(t-1)$	σ_0^2
Итого	$gsrt-1$	

(II) Целью нижеследующего эксперимента являлось определение того, влияет ли положение свитка нитрированной бумаги в стабилизирующем баке на содержание в ней азота. Два свитка (S) были укреплены в двух различных положениях (P); образцы отбирались с каждого свитка сверху и снизу (F). По каждому из этих восьми образцов были произведены дублированные определения содержания азота.

Положение свитка	Номер свитка	Часть свитка	Процент азота	
P_1	S_1	F_1	12,20	12,20
		F_2	12,26	12,26
	S_2	F_1	12,10	12,10
		F_2	12,39	12,38
P_2	S'_1	F_1	12,31	12,30
		F_2	12,44	12,44
	S'_2	F_1	12,23	12,21
		F_2	12,37	12,35

С первого взгляда этот пример как будто попадает в класс анализа с тремя факторами ($P\times S\times F$) с повторениями, рассмотренный в предыдущем разделе.

Однако краткое размышление покажет, что в действительности нет взаимнооднозначного соответствия между S_1 и S'_1 или же между S_2 и S'_2. S_1 имеет не больше связи с S'_1, чем с S'_2. Ясно, что фактор S в действительности является формой повторения.

Кроме того, дублирование анализа не представляет собой подлинного повторения, поскольку весь эксперимент полностью не повторялся (полное повторение эксперимента должно было бы состоять в обследовании различных свитков и в различных баках).

Для проведения дисперсионного анализа мы могли бы провести полное разложение дисперсии по компонентам, затем объединить, согласно вышеуказанным соображениям, соответствующие числа степеней свободы и суммы квадратов. Однако путем непосредственной оценки соответствующих членов работу можно сократить.

В последующем расчете за начало отсчета принята величина 12,20; масштаб был затем увеличен в 100 раз.

(1) Квадрат общего итога делим на число индивидуумов (получаем обычный корректирующий фактор):

$$\frac{134^2}{16}=1122{,}25.$$

(2) Составляем таблицу с двумя входами для P и F.

	F_1	F_2	Итого
P_1	−20	49	29
P_2	25	80	105
Итого	5	129	134

Из этой таблицы определяем обычным порядком суммы квадратов P, F и $P\times F$.

$$P=\frac{29^2+105^2}{8}-1122{,}25=361{,}00,$$

$$F=\frac{5^2+129^2}{8}-1122{,}25=961{,}00,$$

$$P\times F=\frac{20^2+49^2+25^2+80^2}{4}-1122{,}25-361{,}00-961{,}00=12{,}25.$$

(3) Берем сумму квадратов между разложениями:

$$(0^2+0^2+6^2+\ldots+15^2)-\frac{[(0+0)^2+(6+6)^2+\ldots]}{2}=5{,}00.$$

(4) Берем общую сумму квадратов:

$$(0^2+0^2+6^2+\ldots+15^2)-1122{,}25=1731{,}75.$$

(5) Суммы квадратов для погрешности определяем вычитанием. Результирующая таблица имеет вид:

Источники дисперсии	Суммы квадратов	Числа степеней свободы	Средние квадратов
Между положениями в баке (P)	361,00	1	361,00
Между частями свитка (F)	961,00	1	961,00
$F\times P$	12,25	1	12,25
Погрешность	392,50	4	98,12
Между разложениями	5,00	8	0,62
Итого	1731,75	15	

Взаимодействие $F\times P$ после сравнения с погрешностью явно незначимо; его можно отнести к погрешности. Затем можно сравнить с этой новой величиной ошибки основной эффект P, являющийся

*

главным объектом эксперимента. Основной эффект F хотя и не представляет прямого интереса, но может быть сравнен с погрешностью. Следовало, впрочем, включить в эксперимент эффект F, так как возможно, что P будет с ним взаимодействовать.

(L) ТРИЖДЫ НЕПОЛНЫЙ АНАЛИЗ С ЧЕТЫРЬМЯ ФАКТОРАМИ. ОДИН ИЗ ФАКТОРОВ С ПОВТОРЕНИЯМИ ТРОЙНОГО ПОРЯДКА

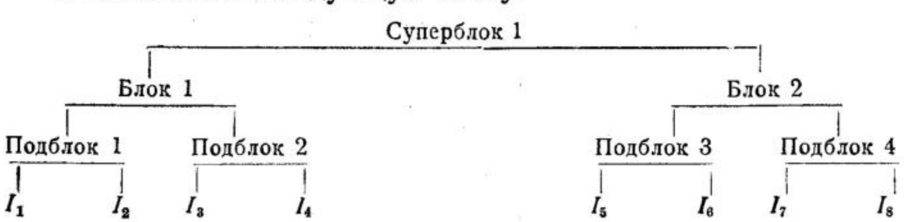
(I) В предыдущем примере фактор R представлял собой источник сырья. Предположим теперь, что этот фактор представляет собой рабочего в цехе. Очевидно, не имеет смысла сравнивать первых рабочих во всех цехах со вторыми во всех цехах. В этом отношении представит интерес лишь различие между рабочими в пределах цеха.

Сумма квадратов и число степеней свободы для нового члена R получаются путем объединения значений для R и $R \times S$ в табл. 11.25. Получаем табл. 11.26:

ТАБЛИЦА 11.26

Источники дисперсии	Числа степеней свободы	Компоненты дисперсии
S	$s-1$	$\sigma_0^2+t\sigma_G^2+gt\sigma_R^2+gtr\sigma_S^2$
Между R, в пределах S	$(r-1)+(r-1)(s-1)=$ $=s(r-1)$	$\sigma_0^2+t\sigma_G^2+gt\sigma_R^2$
Между G, в пределах R. .	$rs(g-1)$	$\sigma_0^2+t\sigma_G^2$
В пределах G (остаток). .	$gsr(t-1)$	σ_0^2
Итого	$gsrt-1$	

Система имеет следующую схему:



Аналогичные схемы и для других суперблоков. Число подразделений в каждой из единиц схемы может быть любым, начиная с 2.

(II) Рассмотрим нижеприведенную таблицу. Имеется предположительно однородный замес материала, содержащийся в шести

баках $V_1, V_2, \ldots, V_6$. Содержимое каждого бака высушивается центрифугой и пакуется в мешки. Отбираются по два мешка B_1 и B_2 из материала от каждого бака; из каждого мешка берутся две выборки; по каждой выборке проводятся дублированные анализы (A_1 и A_2) для выяснения содержания одного из ингредиентов.

	V_1				V_2				V_3				V_4				V_5				V_6			
	B_1		B_2		B_1		B_2		B_1		B_2		B_1		B_2		B_1		B_2		B_1		B_2	
	S_1	S_2	S_1	S_2	S_1	S_2	S_1	S_2	S_1	S_2	S_1	S_2	S_1	S_2	S_1	S_2	S_1	S_2	S_1	S_2	S_1	S_2	S_1	S_2
A_1	29	28	29	27	29	27	26	24	32	29	25	30	29	30	28	30	30	27	25	26	29	31	29	29
A_2	29	27	29	28	29	28	27	25	30	30	27	31	29	31	28	28	29	27	28	26	31	32	30	31

Для облегчения арифметических выкладок при дисперсионном анализе принимается новое начало отсчета, равное 20. Это число вычитается из результата каждого наблюдения.

Подсчитываются следующие величины:

(1) Сумма квадратов всех индивидуумов:

$$9^2+8^2+9^2+7^2+\ldots+10^2+11^2=3636.$$

(2) Сумма квадратов итогов по каждому подблоку, деленная на число индивидуумов в подблоке:

$$\frac{(9+9)^2+(8+7)^2+(9+9)^2+\ldots+(9+10)^2+(9+11)^2}{2}=3616.$$

(3) Сумма квадратов итогов по каждому блоку, деленная на число индивидуумов в блоке:

$$\frac{(9+9+8+7)^2+(9+7+9+8)^2+\ldots+(9+10+9+11)^2}{4}=3570{,}5.$$

(4) Сумма квадратов итогов по каждому суперблоку, деленная на число индивидуумов в суперблоке:

$$\frac{66^2+55^2+74^2+\ldots+82^2}{8}=3534{,}25.$$

(5) Частное от деления квадрата общего итога по всем индивидуумам на общее число индивидуумов:

$$\frac{408^2}{48}=3468.$$

Составляем таблицу дисперсионного анализа (табл. 11.27):

ТАБЛИЦА 11.27

Источники дисперсии	Числа степеней свободы	Суммы квадратов	Средние квадратов	Компоненты дисперсии
Между суперблоками (баки) . . .	$n_1 - 1 = 5$	$(4) - (5) = 66{,}25$	13,25	$n_2 n_3 n_4 \sigma_1^2 + n_3 n_4 \sigma_2^2 + n_4 \sigma_3^2 + \sigma_4^2$
Между блоками в пределах суперблоков (мешки из одного и того же бака)	$n_1 (n_2 - 1) = 6$	$(3) - (4) = 36{,}25$	6,04	$n_3 n_4 \sigma_2^2 + n_4 \sigma_3^2 + \sigma_4^2$
Между подблоками в пределах блоков (выборки из одного и того же мешка)	$n_1 n_2 (n_3 - 1) = 12$	$(2) - (3) = 45{,}50$	3,79	$n_4 \sigma_3^2 + \sigma_4^2$
Между индивидуумами в пределах подблоков (анализы по одной и той же выборке)	$n_1 n_2 n_3 (n_4 - 1) = 24$	$(1) - (2) = 20{,}00$	0,833	σ_4^2
Итого	$n_1 n_2 n_3 n_4 - 1 = 47$	$(1) - (5) = 168$		

В этой таблице:

n_1 — число суперблоков, равное 6,

n_2 — число блоков в пределах суперблока, равное 2,

n_3 — число подблоков в пределах блока, равное 2,

n_4 — число индивидуумов в пределах подблока, равное 2,

σ_1^2, σ_2^2, σ_3^2 и σ_4^2 — величины дисперсии, определяемые соответственно четырьмя членами, перечисленными в первом столбце „Источники дисперсии“.

Приступая к анализу табл. 11.27, сравниваем среднее квадратов между подблоками в пределах мешков со средним квадратов между индивидуумами в пределах подблоков (с целью проверки существования σ_3^2); отношение дисперсий составляет $\frac{3{,}79}{0{,}833} = 4{,}6$ при числах степеней свободы $n_1 = 12$, $n_2 = 24$; оно более значимо, чем требуется для уровня значимости в 1%.

Для проверки существования σ_2^2 сравниваем значение среднего квадратов между блоками в пределах суперблоков со средним квадратов между подблоками. Это отношение дисперсий составляет $\frac{6{,}04}{3{,}79} = 1{,}6$ при числах степеней свободы $n_1 = 6$, $n_2 = 12$. Оно менее значимо, чем требуется для уровня значимости в 20%; можно считать, что σ_2^2 отсутствует. Объединяем данные рассматриваемых двух членов и получаем новую оценку величины $n_4 \sigma_3^2 + \sigma_4^2$, равную $\frac{36{,}25 + 45{,}50}{6 + 12} = 4{,}53$, с 18 степенями свободы.

Для проверки существования σ_1^2 сравниваются значения средних квадратов двух верхних строчек табл. 11.27. Ранее обнаружено, что σ_2^2 отсутствует, так что сравниваются выражения $n_2 n_3 n_4 \sigma_1^2 + n_4 \sigma_3^2 + \sigma_4^2$ и $n_4 \sigma_3^2 + \sigma_4^2$. Это отношение дисперсий равно $\frac{13{,}25}{4{,}53} = 2{,}93$ при числах степеней свободы $n_1 = 5$, $n_2 = 18$. Оно более значимо, чем требуется уровнем значимости в 5%.

Теперь мы можем вычислить

$$\sigma_1^2 = \frac{13{,}25 - 4{,}53}{8} = 1{,}09 \quad \text{и} \quad \sigma_3^2 = \frac{4{,}53 - 0{,}83}{2} = 1{,}85.$$

Наше окончательное заключение сводится к тому, что изменчивость может определяться следующими величинами:

$$\sigma_1^2 = 1{,}09, \quad \sigma_2^2 = 0, \quad \sigma_3^2 = 1{,}85, \quad \sigma_4^2 = 0{,}83.$$

Таким образом, имеется заметная изменчивость между баками; изменчивость между мешками, заполненными продуктом из данного бака, недостаточно велика для того, чтобы быть обнаруженной; изменчивость между выборками, взятыми из данного мешка, велика — фактически она является самой важной компонентой; изменчивость между дублированными анализами одного и того же образца также достаточно заметна. Общая величина дисперсии одиночного анализа равна:

$$1{,}09 + 0 + 1{,}85 + 0{,}83 = 3{,}77.$$

Из этого анализа ясно, что предполагавшаяся однородной партия материала в значительной мере неоднородна; бо̀льшая часть неоднородности возникает при упаковке материала в мешки ($\sigma_3^2 = 1{,}85$). Каждый бак, повидимому, однороден ($\sigma_2^2 = 0$), но имеются заметные различия между баками ($\sigma_1^2 = 1{,}09$).

(M) НЕПОЛНЫЙ АНАЛИЗ С ПЯТЬЮ ФАКТОРАМИ

При наличии пяти факторов может быть значительное количество разных типов неполного анализа. Здесь будет рассмотрен только один тип, но с такой точки зрения, что общий метод обработки станет, как мы надеемся, ясен.

Нижеприводимая таблица содержит данные о процессе обработки одного из производных продуктов целлюлозы. Окончательный продукт оказался неустойчивым в отношении определенного свойства; требуется найти, на какой стадии возникает эта изменчивость. Рассматриваемый производный продукт целлюлозы поступает на обработку в виде больших количеств однородных смесей; ряд партий продукта здесь замешивается. Обычно новый замес полностью накладывается в одиночную вагонетку и затем высушивается; после этого возможно сделать пробу качества продукта.

Номера смесей	Номера пар	Номера замесов							Итого по замесам	Итого по парам	Итого по смесям
1	1	1	11	16	27	17	16	33	60		
		2	0	0	0	14	13	27	27		
					27			60		87	
	2	3	12	12	24	10	8	18	42		
		4	6	6	12	1	3	4	16		
					36			22		58	145
2	3	5	9	9	18	19	18	37	55		
		6	3	1	4	6	6	12	16		
					22			49		71	
	4	7	8	9	17	14	13	27	44		
		8	6	8	14	6	7	13	27		
					31			40		71	142
3	5	9	15	17	32	22	21	43	75		
		10	21	20	41	13	13	26	67		
					73			69		142	
	6	11	18	18	36	19	21	40	76		
		12	6	6	12	6	4	10	22		
					48			50		98	240
4	7	13	11	10	21	3	2	5	26		
		14	5	6	11	1	0	1	12		
					32			6		38	
	8	15	8	8	16	6	6	12	28		
		16	4	4	8	7	8	15	23		
					24			27		51	89

При таком порядке процесса невозможно отделить друг от друга эффекты замеса и сушки; для разделения этих эффектов нужно, чтобы вагонетка загружалась частями, разными замесами (для сравнения замесов независимо от их сушки) и чтобы каждый из замесов высушивался в нескольких вагонетках (для сравнения сушки независимо от состава замеса в вагонетке). По практическим причинам невозможно разделить один и тот же замес более чем на две части, поэтому замес приготовляется и делится на две части, погружаемые в две вагонетки.

Следующий замес также делим на две части и ими заполняем прежние две вагонетки. Аналогичным способом испытываем вторую пару замесов, сделанных из той же смеси производного продукта целлюлозы. Весь эксперимент повторяется на трех последующих смесях продукта. По каждому индивидууму (т. е. по материалу из заданной вагонетки и замеса от заданной пары из данной смеси) делаются повторные определения качества.

Первые две цифры в таблице (11 и 16) выражают результаты дублированных испытаний качества продукта по замесу 1, паре 1, смеси 1, первой вагонетке. Цифра 27 является итогом 11 и 16. Следующая пара цифр 17 и 16 и их итог 33 относятся ко второй вагонетке замеса 1. Последняя цифра в строке (60) — итог по замесу. Вторая строка относится ко второму замесу и составлена совершенно так же, как первая. В третьей строке проставлены три итога: 27 — для первой вагонетки, 60 — для второй вагонетки и 87 — для пары вагонеток.

Аналогичные записи в последующих трех строках относятся ко второй паре вагонеток, заполненных из смеси 1; в конце проставлен итог по смеси — 145.

С первого взгляда это — анализ с пятью факторами: смеси (B), пары вагонеток (P), замесы (M), вагонетки (T) и определения качества (A). Однако сразу становится очевидным, что этот анализ не является полным, так как A является простым повторением: нет взаимных различий между первыми и вторыми определениями качества. Далее, нет никаких специальных различий между первыми и вторыми парами вагонеток, так что отсутствует основной эффект пар P, и эффект между парами в пределах смесей получится путем объединения сумм квадратов и чисел степеней свободы членов P и $B \times P$.

Начнем с подсчета этого члена и основного эффекта B. Корректирующий фактор равен квадрату общего итога, деленному на общее число индивидуумов: $\frac{616^2}{64} = 5929.$

Основной эффект по смесям определяется из их итогов:

$$\frac{145^2 + 142^2 + 240^2 + 89^2}{16} - 5929 = 740{,}37.$$

Член, соответствующий взаимодействию между парами в пределах смесей, получается непосредственно путем суммирования квадратов итогов по парам и деления суммы на число индивидуумов, входящих в пару; из результата вычитается корректирующий фактор и сумма квадратов по смесям:

$$\frac{87^2+58^2+\ldots+51^2}{8}-5929-740{,}37=184{,}13.$$

Тот же самый результат получился бы, конечно, путем составления таблицы $B\times P$, подсчета B, P и $B\times P$ (остатка) и объединения данных P с $B\times P$.

Рассмотрим теперь эффект M. В действительности основного эффекта M не может быть, поскольку первые замесы во всех парах не имеют систематических отличий от вторых замесов; поэтому $\sigma^2_m=0$. Далее, эффект M представляет интерес только в пределах данной пары, находящейся в пределах заданной смеси; этот член равен σ^2_{mbp}. Члены же σ^2_{mb} и σ^2_{mp} равны нулю. Если вычеркнуть σ_m, σ^2_{mb} и σ^2_{mp} из числа компонент дисперсии для M, $M\times B$ и $M\times P$, то в каждом случае останутся лишь следующие компоненты дисперсии:

$$\sigma^2_a+a\sigma^2_{mtbp}+at\sigma^2_{mbp}.$$

Эти же компоненты входят и в среднее квадрата для $M\times B\times P$. Поэтому объединяем четыре члена M, $M\times B$, $M\times P$, $M\times B\times P$ и результирующему члену присвоим название „между замесами в пределах пар и смесей"; объединенное число степеней свободы равно:

$$(m-1)+(m-1)(b-1)+(m-1)(p-1)+\\+(m-1)(p-1)(n-1)=bp(m-1).$$

Это выражение можно получить также следующим образом: каждая пара замесов вносит $m-1$ степеней свободы; всего имеется bp пар, следовательно, общее число степеней свободы равно $bp(m-1)$.

Для подсчета объединенного члена $M\times B\times P$ требуется таблица $M\times B\times P$, т. е. таблица, получаемая путем суммирования по A и T. Эта таблица уже имеется в виде графы „Итого по замесам" (60, 27, 42, 16, 55 и т. д.). Общая сумма квадратов для этой таблицы равна:

$$\frac{(60^2+27^2+42^2+\ldots+23^2)}{4}-5929=1771{,}5.$$

Для получения объединенного члена $M\times B\times P$ можно было бы полностью вычислить члены для M, $M\times B$, $B\times P$ и $M\times B\times P$ и затем объединить их. Можно видеть, однако, что итог таблицы $M\times B\times P$ составляется без этих членов и членов для B, P

и $B\times P$. Эти последние уже известны, так что объединенный член $M\times B\times P$ легко вычисляется так:

$$1771{,}5-740{,}37-184{,}13=847{,}00.$$

Его число степеней свободы: $bp(m-1)=4\cdot2(2-1)=8$. Член T получается точно таким же путем, как и член M. Требуется найти член, образуемый путем объединения T, $T\times B$, $T\times P$ и $T\times B\times P$. Начинаем с таблицы $T\times B\times P$, составляемой путем суммирования по A и M. Общая сумма квадратов для этой таблицы равна:

$$\frac{27^2+60^2+36^2+22^2+22^2+\ldots+27^2}{4}-5929{,}0=1274{,}5.$$

Из этого числа вычитаются, как и ранее, суммы квадратов B, P и $B\times P$; остаток будет суммой квадратов члена $B\times P\times T$:

$$1274{,}5-740{,}37-184{,}13=350{,}00.$$

Число степеней свободы равно: $bp(t-1)=4\cdot2(2-1)=8$.

Теперь переходим к взаимодействию $M\times T$. Аналогично рассуждениям для M и T, устанавливаем, что взаимодействие $M\times T$ существует только в пределах пар и смесей; поэтому объединяем $M\times T$ с $M\times T\times B$, $M\times T\times P$ и $M\times T\times B\times P$. Исходим из таблицы $M\times T\times B\times P$, т. е. из таблицы, образованной путем суммирования по A. Общая сумма квадратов для этой таблицы равна:

$$\frac{27^2+33^2+0^2+27^2+24^2+\ldots+15^2}{2}-5929{,}00=2316{,}00.$$

Для получения суммы квадратов для объединенного члена $M\times T\times B\times P$ можем из этого итога вычесть все другие члены таблицы $M\times T\times B\times P$, а именно:

(a) B и $B\times P$; они составляют $740{,}37+184{,}13$.

(b) M, $M\times B$, $M\times P$ и $M\times B\times P$; подсчитанная ранее сумма их равна 847,00.

(c) T, $T\times B$, $T\times P$ и $T\times B\times P$; подсчитанная ранее сумма их равна 350,00.

Искомая сумма квадратов тогда равна:

$$2316{,}00-740{,}37-184{,}13-847{,}00-350{,}00=194{,}50.$$

Число степеней свободы равно:

$$(m-1)(t-1)+(m-1)(t-1)(b-1)+(m-1)(t-1)(p-1)+\\+(m-1)(t-1)(p-1)(b-1)=bp(m-1)(t-1)=\\=4\cdot2(2-1)(2-1)=8.$$

Источники дисперсии	Объединяемые значения средних квадратов из полного анализа	Суммы квадратов	Числа степеней свободы	Средние квадраты	Компоненты дисперсии
Между смесями	B	740,37	3	246,79	$\sigma_a^2 + a\sigma_{mtbp}^2 + am\sigma_{tbp}^2 + at\sigma_{mbp}^2 + amt\sigma_{bp}^2 + amtp\sigma_b^2$
Между парами в пределах смесей	P, $P \times B$	184,13	4	46,03	$\sigma_a^2 + a\sigma_{mtbp}^2 + am\sigma_{tbp}^2 + at\sigma_{mbp}^2 + amt\sigma_{bp}^2$
Между вагонетками в пределах пар и смесей	T, $T \times B$, $T \times P$, $T \times B \times P$	350,00	8	43,75	$\sigma_a^2 + a\sigma_{mtbp}^2 + am\sigma_{tbp}^2$
Между замесами в пределах пар и смесей	M, $M \times B$, $M \times P$, $M \times B \times P$	847,00	8	105,87	$\sigma_a^2 + a\sigma_{mtbp}^2 + at\sigma_{mbp}^2$
Взаимодействие между замесами и вагонетками в пределах пар и смесей	$M \times T$, $M \times T \times B$, $M \times T \times P$, $M \times T \times B \times P$	194,50	8	24,31	$\sigma_a^2 + a\sigma_{mtbp}^2$
Между анализами		33,00	32	1,03	σ_a^2
Итого		2349,00	63		

Последним искомым числом будет сумма квадратов между анализами, равная

$$(11^2 + 16^2 + 17^2 + 16^2 + 0^2 + \ldots + 8^2) - \frac{27^2 + 33^2 + 0^2 + \ldots + 15^2}{2} = \\ = 8278 - 8245 = 33{,}00.$$

Общая сумма квадратов подсчитывается обычным путем:

$$(11^2 + 16^2 + \ldots + 8^2) - 5929 = 2349{,}00.$$

Вносим все полученные члены в таблицу дисперсионного анализа на стр. 156.

Соответствующие проверки значимости могут быть сделаны по компонентам дисперсии. Следует, однако, иметь в виду, что если члены „между замесами“ и „между вагонетками“ окажутся значимыми при сравнении с взаимодействием „между замесами и вагонетками“, т. е. если обе величины σ_{mbp}^2 и σ_{tbp}^2 больше нуля, то в этом случае не будет действительной проверки для значения средних квадратов „между парами“. Это обстоятельство сходно с возникшим в табл. 11.5 (глава XI, (D)).

Среднее квадратов „между вагонетками“ незначимо, но, возможно, все же стоит подсчитать вносимую им компоненту дисперсии. В результате получим:

$$\sigma_a^2 = 1{,}03, \quad \sigma_{mt}^2 = 11{,}64, \quad \sigma_m^2 = 15{,}40,$$
$$\sigma_t^2 = 4{,}86, \quad \sigma_p^2 = 0{,}00, \quad \sigma_b^2 = 10{,}05.$$

Отсюда видно, что источники изменчивости довольно широко разбросаны; в большей части изменчивости повинны замешивание и его взаимодействие с вагонетками (сушка) и смеси.

ГЛАВА XII

РАЗНЫЕ ВОПРОСЫ, СВЯЗАННЫЕ С ДИСПЕРСИОННЫМ АНАЛИЗОМ

(A) ВВЕДЕНИЕ

В этой главе рассматривается ряд разных вопросов, связанных с применением дисперсионного анализа. Четыре следующих раздела охватывают материал, который следует иметь в виду при планировании и анализе результатов факторных испытаний; в разделе (G) описан полезный для экспериментирования прием (так называемый „латинский квадрат)“; в последних разделах рассматриваются различные применения дисперсионного анализа.

(B) ПРИМЕНЕНИЕ КОМПОНЕНТ ДИСПЕРСИИ

В систематическом изложении дисперсионного анализа в главе XI для каждого типа анализа были даны компоненты дисперсии, соответствующие каждому из значений средних квадратов.

Следует иметь в виду, что средние квадратов являются только оценками соответствующих членов; их нельзя рассматривать как алгебраически точные выражения. Например, в полном анализе с тремя факторами (табл. 11.6)

31,08 служило оценкой для $w\sigma_{ta}^2+\sigma_0^2$,

161,08 " " " σ_0^2.

Очевидно, что σ_{ta}^2 равно нулю; таким образом, имеются две оценки для σ_0^2, которые можно объединить обычным путем, с учетом чисел степеней свободы. (Простейший путь — сложить суммы квадратов и числа степеней свободы и рассчитать новое значение среднего квадрата, разделив первую величину на вторую.)

Это объединение, конечно, имеет силу только в случае, когда объединяемые оценки окажутся совместимыми при проверке по методу отношения дисперсий. Если $w\sigma_{ta}^2+\sigma_0^2$ значимо превосходит σ_0^2, то, конечно, объединение не делается. Возможно, что $w\sigma_{ta}^2+\sigma_0^2$ окажется значимо меньшим, чем σ_0^2. Это весьма неправдоподобно; если же так случилось, то смысл этого весьма неясен: либо это случайность (которая может встретиться однажды в 20 случаях) и ее можно не принимать во внимание, либо здесь замешан субъективный фактор, каким-то образом исказивший результаты.

Компоненты дисперсии полезны для указания последовательности проверок на значимость. Кроме того, они полезны для оценки действительной величины изменчивости, вызываемой каждой стадией процесса и для получения указания относительно стадии процесса, в которой должен быть улучшен контроль в целях повышения однородности продукции. Это применение было рассмотрено в главах VII, (C), и XI, (M).

Коэффициенты компонент дисперсии, приведенные в различных таблицах в главе XI, строго говоря, имеют силу только в случае, когда предположенные совокупности бесконечны. Если совокупность состоит из партий сырья, вполне можно допустить, что имеется бесконечная совокупность их, даже если имеют дело только с несколькими дюжинами или несколькими сотнями. Если же совокупность состоит из цехов или отдельных заводов, то обычно их бывает всего два или три.

В таких условиях несколько затруднительно рассматривать совокупность бесконечной, и, строго говоря, коэффициенты компонент дисперсии оказываются измененными (см. H. E. Daniels, Supp. Journ. Roy. Stat. Soc., VI, № 2 (1939), 188).

На практике этими изменениями можно пренебречь, так как: (a) последовательность проверок значимости неизменна, поэтому применение неизмененных коэффициентов не вызовет ошибок; (b) действительные оценки компонент дисперсии с простыми коэффициентами должны умножаться на коэффициент $\frac{k-1}{k}$, где k — объем совокупности. Для $k=2$ или 3 величина коэффициента заметно меньше единицы и составляет соответственно 0,5 и 0,66. При больших значениях k разность между коэффициентом и единицей обычно несущественна, если учесть достаточно большую величину стандартной ошибки, конечно, присущую этим компонентам дисперсии.

В случае, когда фактором являются произвольно выбираемые различные уровни непрерывной переменной, совокупность может с достаточным основанием считаться бесконечной; но, вероятно, больше интереса представят действительные средние значения, так что вычислять компоненты дисперсии в этом случае может быть и не придется.

C) РАЗЛОЖЕНИЕ СУММЫ КВАДРАТОВ НА ЛИНЕЙНЫЕ, КВАДРАТИЧНЫЕ И Т. Д. КОМПОНЕНТЫ

Часто встречаются случаи, когда фактор имеет 3, 4, 5 или более уровней; сам по себе фактор выражается численно, уровни выбраны на известных интервалах непрерывного изменения переменной (например, температуры). Требуется выяснить, может ли соотношение между независимой и зависимой переменными рассматриваться как линейное, или же кривая, изображающая эту зависимость, имеет заметную кривизну. Если уровни выбраны равноотстоящими, например в случае температуры в 20, 25, 30, 35°C, ответ на поставленный вопрос получить довольно легко.

Рассмотрим данные в таблице 12.1. Четыре T представляет собою четыре равноотстоящих значения температуры при нитрации

ТАБЛИЦА 12.1

	T_1	T_2	T_3	T_4	Итого
W_1	6	32	45	63	146
W_2	24	27	45	62	158
W_3	6	24	44	45	119
W_4	1	22	23	39	85
Итого . .	37	105	157	209	508

целлюлозы в нитрирующей кислотной смеси; последняя, в свою очередь, может иметь четыре уровня W по содержанию воды; эти уровни тоже выбраны с равными промежутками друг от друга. Зависимой переменной, значения которой приведены в таблице, является вязкость получаемой нитроцеллюлозы.

Непосредственный дисперсионный анализ с двумя факторами приведен в табл. 12.2.

ТАБЛИЦА 12.2

Источник дисперсии	Число степеней свободы	Суммы квадратов	Средние квадратов
Температура	3	4052,0	1350,67
Содержание воды . . .	3	787,5	262,50
Остаточная	9	367,5	40,83
Итого	15	5207,0	

Очевидно, что оба эффекта высоко значимы. Теперь можно приступить к разложению каждой из двух сумм квадратов на три компоненты, с одной степенью свободы для линейного члена, с одной степенью свободы для квадратичного члена и с одной степенью свободы для кубичного члена.

Линейная компонента суммы квадратов определяется по формуле:

$$\frac{(3T_4+T_3-T_2-3T_1)^2}{N\Sigma k^2}.$$

T представляют собою итоги по N наблюдениям; Σk^2 — сумма квадратов коэффициентов в числителе. В данном случае

$$\Sigma k^2 = 3^2+1^2+(-1)^2+(-3)^2=20.$$

Подставляя цифры, получаем, что линейная компонента суммы квадратов по температурам равна:

$$\frac{(3\cdot 37+105-157-3\cdot 209)^2}{4\cdot 20}=4032{,}8.$$

Квадратичная компонента суммы квадратов равна:

$$\frac{(T_4-T_3-T_2+T_1)^2}{N\Sigma k^2}=\frac{(37-105-157+209)^2}{4\cdot 4}=16.$$

Кубичная компонента равна:

$$\frac{(T_4-3T_3+3T_2-T_1)^2}{N\Sigma k^2}=\frac{(37-3\cdot 105+3\cdot 157-209)^2}{4\cdot 20}=3{,}2.$$

Проверка правильности арифметических вычислений: сумма всех трех компонент должна равняться итогу в таблице 12.2.

Таким же путем производится разложение на компоненты суммы квадратов по содержанию воды.

Результаты сведены в таблицу 12.3.

ТАБЛИЦА 12.3

Источник дисперсии	Число степеней свободы	Суммы квадратов	Средние квадратов
Температура:			
Линейная компонента	1	4032,8	4032,8
Квадратичная »	1	16	16
Кубичная »	1	3,2	3,2
Содержание воды:			
Линейная компонента	1	616,06	616,06
Квадратичная »	1	132,25	132,25
Кубичная »	1	39,2	39,2
Остаток	9	367,5	40,83
Итого	15	5207,0	

Линейные компоненты в обоих случаях очевидно значимы; значимость же квадратичных и кубичных компонент можно проверить обычным путем, сравнивая их с остатком. Интересно заметить, что кажущийся перегиб зависимости от W над точкой W_2 в сторону уменьшения вязкости по направлению к W_1 не является значимым.

В случае трех уровней имеются 2 степени свободы: одна может быть отнесена к линейной компоненте, другая — к квадратичной. Эти компоненты сумм квадратов соответственно равны

$$\frac{(T_1-T_3)^2}{N\Sigma k^2};\quad \Sigma k^2=2,$$

$$\frac{(2T_2-T_1-T_3)^2}{N\Sigma k^2};\quad \Sigma k^2=6.$$

В книге Фишера и Йэйтса „Статистические таблицы для биологических, сельскохозяйственных и медицинских исследований“ (Fisher and Yates, Statistical Tables for Biological, Agricultural and Medical Research) приведены соответствующие коэффициенты для различных случаев, вплоть до 52 уровней.

В случае, когда два фактора являются числовыми переменными, изменяющимися равными интервалами, их взаимодействие можно разложить таким же образом. Рассмотрим таблицу 3×3 (таблица 12.4).

Для разложения двух основных эффектов на линейные (X_L, Y_L) и квадратичные (X_Q, Y_Q) компоненты можно использовать выражения:

$$\frac{(T_1-T_3)^2}{N\Sigma k^2}\ \text{и}\ \frac{(-T_1+2T_2-T_3)^2}{N\Sigma k^2}.$$

ТАБЛИЦА 12.4

	X_1	X_2	X_3	Итого
Y_1	17	20	24	61
Y_2	14	19	25	53
Y_3	10	18	30	58
Итого ..	41	57	79	177

Можно также распределить 4 степени свободы для взаимодействия по компонентам, относящимся соответственно к $X_L Y_L$, $X_Q Y_L$, $X_L Y_Q$ и $X_Q Y_Q$. Для этого выписываем таблицу типа 12.5.

ТАБЛИЦА 12.5

X_L	1	0	—1		
X_Q	—1	2	—1	Y_L	Y_Q
				1	—1
				0	2
				—1	—1

Чтобы образовать таблицу коэффициентов для получения $X_L Y_L$, берем коэффициенты для X_L и умножаем их на первый коэффициент для Y_L; получаем первую строку. Умножая коэффициенты для X_L на второй коэффициент для Y_L получаем вторую строку; аналогичная операция проделывается и для получения третьей строки. Для получения $X_Q Y_L$ умножаем тем же путем коэффициенты для X_Q на коэффициенты для Y_L.

В результате получаем таблицу коэффициентов (табл. 12.6).

ТАБЛИЦА 12.6

$X_L Y_L$	$X_Q Y_L$	$X_L Y_Q$	$X_Q Y_Q$
1 0 —1	—1 2 —1	—1 0 1	1 —2 1
0 0 0	0 0 0	2 0 —2	—2 4 —2
—1 0 1	1 —2 0	—1 0 1	1 —2 1
$\sum k^2 = 4$	$\sum k^2 = 12$	$\sum k^2 = 12$	$\sum k^2 = 36$

$\sum k^2$ есть сумма квадратов соответствующих групп коэффициентов. Для получения, например, суммы квадратов для $X_Q Y_L$ умножаем цифры таблицы 12.4 на соответственные коэффициенты таблицы 12.6, суммируем результат, возводим сумму в квадрат

и делим на $N\sum k^2$, где N — число наблюдений по каждой цифре таблицы 12.4, которая умножалась на коэффициент из таблицы 12.6. В настоящем случае $N=1$; если бы цифры в таблице 12.4 являлись итогами N_1 наблюдений, то $N=N_1$.

В соответствии с указанным, сумма квадратов для $X_Q Y_L$ равна:

$$\frac{(-1\cdot 17+2\cdot 20-1\cdot 24+1\cdot 10-2\cdot 18+1\cdot 30)^2}{12}=0{,}750.$$

Теперь выпишем дисперсионный анализ полностью (табл. 12.7). Величины в столбце „Суммы квадратов“ получаются путем возведения в квадрат величины, стоящей во втором столбце и деления результата на число, стоящее в третьем столбце. Проверкой правильности вычислений послужит равенство суммы всех компонент с суммой, вычисленной непосредственно из первоначальных результатов наблюдений.

ТАБЛИЦА 12.7

Источник дисперсии	Произведения коэффициентов и данных	$N\sum k^2$	Суммы квадратов
X_L	—38	3×2	240,667
X_Q	— 6	3×6	2,0
Y_L	— 3	3×2	1,5
Y_Q	— 3	3×6	0,5
$X_L Y_L$	13	1×4	42,25
$X_Q Y_L$	3	1×12	0,75
$X_L Y_Q$	5	1×12	2,083
$X_Q Y_Q$	3	1×30	0,25
Итого ...			290,0

Если не делать этого детального анализа, а выполнить попросту анализ по строкам и столбцам, мы получили бы таблицу 12.8.

ТАБЛИЦА 12.8

Источник дисперсии	Числа степеней свободы	Суммы квадратов	Средние квадратов
X	2	242,667	121,333
Y	2	2,00	1,00
Остаток	4	45,333	11,333
Итого . .	8	290,0	

Последний вид анализа, хотя и дает кое-что, но далеко не столь продуктивен, как анализ, приведенный в таблице 12.7. Таблица 12.8 совершенно не может дать какой-либо информации о возможном взаимодействии XY. В таблице же 12.7 следует, разумеется, объединить малые члены X_Q, Y_Q, X_LY_Q, X_QY_L; сравнивая же с результатом X_LY_L, найдем, что этот член значим.

Описанный прием применим, вообще говоря, при любом числе уровней. Из таблицы типа 12.5 мы можем образовать таблицы типа 12.6; таким образом, можно будет исследовать смешанные уровни, например, 3×4 или 3×5 или 4×5. Кроме того, могут присутствовать и другие факторы; это обстоятельство учитывается путем выбора величины N в знаменателе $N\sum k^2$ так, чтобы она соответствовала числу первоначальных наблюдений, результаты которых были суммированы для получения таблицы типа 12.4, которой пользуются совместно с таблицей коэффициентов для вычисления компонент.

В случае четырех или более уровней для каждого фактора число компонент быстро возрастет; например, для испытания вида 4×4 остаток имеет 9 степеней свободы, значит, существуют 9 компонент; испытания вида 5×5 будут иметь 16 компонент. Расчет всех компонент потребует много времени и, в действительности, не нужен. Обычно, представляют интерес только компоненты X_LY_L, X_QY_L и X_LY_Q. Даже последние две компоненты, едва ли могут представлять интерес, если X_Q и Y_Q оказались малыми. В соответствии с этим, как правило достаточно вычислить обычным порядком итоговое взаимодействие, и получить остаток в виде разности между итогом и суммой упомянутых трех компонент. Разумеется, в этом случае нет простой проверки правильности арифметических вычислений, так что производимые действия нуждаются в тщательной проверке по мере их выполнения.

Следует подчеркнуть еще раз, что весь вышеописанный порядок работы основан на предположении, что для каждого фактора выбраны равноотстоящие уровни, например, для температуры взяты 10°, 15°, 20°, 25°. Если окажется удобным, уровни могут быть взяты равноотстоящими на логарифмической шкале, например, 1%, 2%, 4%, 8%.

В таком случае компонента X_L будет мерой линейности тоже в логарифмическом масштабе, а X_Q и высшие члены — отклонения от линейности в том же масштабе.

(D) ДОПУЩЕНИЕ, ЛЕЖАЩЕЕ В ОСНОВЕ ФАКТОРНОГО ИСПЫТАНИЯ

Факторное испытание гораздо полезнее классического тем, что интерпретация факторного испытания является самой простой и что его „производительность" в смысле объема получаемой информации при данном числе опытов достигает максимума, когда взаимодействия не являются значимыми.

Интересно заметить, что отсутствие взаимодействий заранее предполагает, что зависимая переменная y может быть выражена в виде суммы ряда функций независимых переменных x_1, x_2 и т. д., причем каждая из функций содержит одну и только одну переменную:

$$y = f(x_1) + f(x_2) + \ldots . \tag{1}$$

Эти функции, разумеется, могут быть какими угодно, без ограничений в отношении сложности.

Возьмем простой пример. Пусть

$$y = x_1 + 2x_2. \tag{2}$$

Если каждая из независимых переменных имеет два уровня 1 и 2, получим величины y, сведенные в таблице:

	$x_2 = 1$	$x_2 = 2$
$x_1 = 1$	3	5
$x_1 = 2$	4	6

Легко видеть отсутствие взаимодействия $x_1 \times x_2$. Так, например, изменение y от изменения x_2 при $x_1 = 1$ составляет $5 - 3 = 2$; то же изменение при $x_1 = 2$ равно также $6 - 4 = 2$; иначе говоря, изменение y при изменении x_2 не зависит от уровня x_1.

Рассмотрим теперь соотношение вида

$$y = f(x_1) f(x_2), \tag{3}$$

где функции $f(x_1)$ и $f(x_2)$ могут быть какие угодно.

Возьмем простейший пример; пусть

$$y = x_1 x_2^2. \tag{4}$$

Если каждая из независимых переменных имеет два уровня 1 и 2, мы имеем следующую таблицу значений y:

	$x_2=1$	$x_2=2$
$x_1=1$	1	4
$x_1=2$	2	8

Легко видеть, что здесь существует взаимодействие $x_1 \times x_2$. Изменение y при изменении x_2 от 1 до 2 при $x_1=1$ равно $4-1=3$, а при $x_1=2$ то же изменение составляет $8-2=6$; другими словами, изменение y при изменении x_2 зависит от уровня x_1.

Очевидно, что, прологарифмировав уравнение (3), получим:

$$\log y = \log f(x_1) + \log f(x_2), \tag{5}$$

т. е. уравнение такого же вида, что и (1); следовательно, в (5) между факторами в правых частях не будет взаимодействия.

(E) ПРИМЕНЕНИЕ ВЗАИМОДЕЙСТВИЙ В КАЧЕСТВЕ ОЦЕНОК ПОГРЕШНОСТИ

В нашем простейшем случае анализа с двумя факторами (строки и столбцы) то, что мы называли остатком, включало не только дисперсию из-за погрешности, но также из-за любого возможного взаимодействия между факторами. Если последнее существует и велико, мы получим для остатка большое значение среднего квадратов; следовательно, при этом ни один из основных эффектов, вероятно, не будет значимым; мы закончим эксперимент, зная немногим более о поведении системы, чем в начале работы.

В таких обстоятельствах можно действовать двояко. Можно сделать повторения и провести анализ так, как описано в разделе об анализе с двумя факторами и повторениями; это позволит получить оценку погрешности, не зависящую от взаимодействия.

Другой прием — повторить ряды наблюдений при наличии третьего фактора, поддерживаемого постоянным на определенном уровне в течение первой серии наблюдений, и на другом уровне — в течение второй серии. Теперь эксперимент будет иметь три фактора; мы положим, что взаимодействие второго порядка $P \times Q \times R$ мало или вообще отсутствует, чтобы не увеличивать остаток.

Однако может случиться, что это взаимодействие второго порядка окажется велико. Хороший пример был приведен в разделе (E) главы XI: рассмотренный там эксперимент с четырьмя факторами был разбит на два эксперимента с тремя факторами; S_2 имело большой остаток, почти наверное из-за наличия большого взаимодействия второго порядка.

В таких условиях также возможны два пути. Можно сделать повторения; в этом случае анализ проводится так, как указано в разделе об анализе с тремя факторами и с повторениями. Или же можно добавить четвертый фактор и использовать взаимодействие третьего порядка в качестве оценки погрешности. Может оказаться, что это взаимодействие третьего порядка велико, и мы получим преувеличенную оценку погрешности и все еще не окажемся в лучшем положении. В этом случае, однако, мы узнали, что это взаимодействие велико, так что включение дополнительного фактора не оказалось бесполезным.

Дальнейший пример приведен в эксперименте с пятью факторами, рассмотренном в главе XI, (F), где было очень велико взаимодействие $W \times T \times D$. Если бы эксперимент проводился только с тремя факторами W, T и D, это большое взаимодействие $W \times T \times D$ было бы включено в остаток, и нельзя было бы вывести никакого заключения. В действительности же эксперимент включал еще два фактора S и G, и, к счастью, ни один из них не имел взаимодействий, так что был получен достаточно малый остаток.

Таким образом, можно формулировать некоторые общие положения, касающиеся планирования факторных экспериментов. Прежде всего, нужно решить, какими основными факторами следует интересоваться, и избрать число уровней, при которых будет проводиться их испытание.

Затем нужно определить факторы, нормально изменяющиеся в процессе и наиболее вероятно взаимодействующие с избранными основными факторами; эти вторые факторы следует включить, по возможности, при двух только уровнях. Затем надо решить, даст ли комбинация введенных факторов удовлетворительную величину остатка, как бы вызванную некоторым несуществующим взаимодействием высшего порядка. Часто бывает трудно притти к такому решению. Некоторая предварительная информация об общем поведении системы позволяет сделать подобное заключение с некоторой степенью уверенности; однако иногда бывает совершенно невозможно что-либо предугадать. Можно лишь положиться на общее правило, что взаимодействия высоких порядков встречаются не очень часто, а если и встречаются, то обычно они не очень велики; кроме того, нужно быть готовыми повторить весь эксперимент, если в конце работы получится большой остаток. Если же абсолютно необходима уверенность в получении результатов в первом эксперименте, повторения должны быть включены в эксперимент сразу, быть может даже за счет исключения одного из факторов.

По всей вероятности, вообще, желательно включать партии сырья в качестве одного из факторов, так как может быть, что не все партии одинаково реагируют на заданный способ обработки. В опыте автора был случай, когда сырой химикалий подвергался процессу очистки; факторное испытание установило с уровнем значимости выше 1%, что уровень одного из факторов приводил к получению намного лучшей продукции, чем уровень другого. Этот эксперимент проводился над одной партией сырьевого материала; последующий опыт показал, что сделанное заключение в общем случае ошибочно. Если бы эксперимент был повторен еще на одной или более партий материала, такая ошибка не произошла бы. Если мы в состоянии проводить повторения, то их следует делать на различных партиях материала, а не в пределах той же партии; фактически это окажется введением дополнительного фактора — партии сырого материала.

(F) СВЕДЕНИЯ, ТРЕБУЮЩИЕСЯ В ОТЧЕТЕ

При представлении доклада или отчета об исследовании, проведенном факторным методом с дисперсионным анализом результатов, желательно указать:

(a) Факторы и уровни их, при которых велось исследование.

(b) Сам анализ. Целесообразно отмечать звездочками значимые члены; знак (*) — для уровня в 5% (**) — для уровня в 1% и (***) — для уровня $0{,}1\%$.

(c) Если мы заинтересованы в компонентах дисперсии, надлежит подсчитать те из них, которые окажутся значимыми. Если интересуются средними, их следует привести. В случае взаимодействия двух факторов необходимо составить таблицу с двумя входами, например типа табл. 11.13.

(d) Обычно нет надобности приводить исходные данные, особенно когда они громоздки; существенные факты содержатся в перечисленных статистических данных. Однако, если исходные данные негромоздки, не мешает их привести.

(G) ЛАТИНСКИЙ КВАДРАТ

Латинский квадрат — статистический прием, впервые введенный Фишером для преодоления в сельскохозяйственных экспериментах затруднений, возникающих из-за различий в плодородности почв.

Допустим, что надо исследовать влияние двух факторов P и Q на некоторый процесс; каждый фактор имел четыре уровня. P могло бы выражать давление, при котором прессуется кордит, а Q — угол, под которым установлена нажимная плита пресса,

Затем провели 16 экспериментов, охватывающих все комбинации пар факторов P_1Q_1, P_1Q_2, P_1Q_3, P_1Q_4, P_2Q_1 и т. д.

	Q_1	Q_2	Q_3	Q_4
P_1				
P_2				
P_3				
P_4				

В первом приближении (пренебрегая взаимодействиями) сравнения уровней величины Q независимы от сравнений уровней величины P; результаты такого эксперимента могли бы быть проверены на значимость путем дисперсионного анализа по строкам и столбцам.

Предположим, однако, что материал поступает на прессование с предыдущей стадии производственного процесса (смешивание) в однородных по составу партиях такого размера, что из одной такой партии можно сделать, например, только десять прессовок. Тогда было бы невозможно провести указанный эксперимент, требующий 16 прессовок; часть прессовок пришлось бы сделать из одной партии материала, часть — из другой; возможные различия между партиями безнадежно запутали бы сравнение эффектов величин P и Q.

Это затруднение может быть избегнуто путем применения латинского квадрата. Пусть A, B, C и D — четыре различные партии материала. Первый эксперимент над партией A был проведен при значении P, равном P_1, и значении Q, равном Q_1, второй эксперимент — при значениях P_2 и Q_4 и т. д. Аналогично для других партий

	Q_1	Q_2	Q_3	Q_4
P_1	A	B	C	D
P_2	B	C	D	A
P_3	C	D	A	B
P_4	D	A	B	C

Заметим, что каждая строка этой таблички содержит A, B, C и D по одному и только по одному разу. Поэтому, каковы бы ни были нарушающие влияния, вносимые возможными различиями между партиями, они скажутся в каждой строке; следовательно, они полностью будут исключены, если для сравнения величин P будут взяты средние значения по строкам.

Точно так же каждый столбец содержит A, B, C и D ровно по одному разу; все нарушающие влияния, вносимые различиями между партиями, исключаются путем сравнения средних Q по столбцам.

Результаты подобного эксперимента подлежат проверке на значимость и дисперсионному анализу следующим порядком:

(1) Определяется сумма квадратов всех наблюдений.

(2) Сумма квадратов итогов по каждой строке делится на число индивидуумов в строке.

(3) То же — по столбцам.

(4) То же — по партиям.

(5) Квадрат общего итога делится на общее число индивидуумов.

Затем составляется таблица дисперсионного анализа.

Источник дисперсии	Суммы квадратов	Числа степеней свободы	Средние квадратов
Строки	(2) — (5)	$n-1$	
Столбцы	(3) — (5)	$n-1$	
Партии	(4) — (5)	$n-1$	
Остаток		$(n-1)(n-2)$	
Итого	(1) — (5)	n^2-1	

Здесь n — размер квадрата (число индивидуумов, составляющих его сторону). Остаточная сумма квадратов является, как обычно, разностью между итоговой суммой квадратов и суммами квадратов всех других строк.

Различные значения средних квадратов проверяются на значимость с помощью отношения дисперсий по Фишеру. В данном случае главный интерес представляют эффекты по строкам и столбцам, соответствующим рассматриваемым факторам P и Q. Если какой-либо из эффектов не оказывается значимым, можно считать, что нет значимой разницы соответственно между влияниями величин P (или Q). Если же этот эффект окажется значимым, можно подсчитать средние значения для каждого уровня соответствующей величины P (или Q). Для детального сравнения любой пары уровней P (или Q) вычисляется стандартная ошибка среднего значения, равная частному от деления квадратного корня значения среднего квадрата остатка на квадратный корень из n.

Здесь рассмотрен латинский квадрат 4×4; вообще говоря, размер квадрата неограничен. Разработаны расчеты для квадратов до 12×12.

Одним из допущений, лежащих в основе применения латинского квадрата, является то, что взаимодействие между факторами не должно иметь места. Поскольку в промышленных химических процессах типично наличие взаимодействий, латинский квадрат не может иметь широкого применения в этой области.

Альтернативный метод работы при ограниченных размерах партий материала, когда требуется провести заданное число экспериментов для факторного испытания, не имея достаточных по размеру партий сырьевого материала, известен под названием „смешивания“ (confounding) (см. гл. XIV).

(Н) ГРЕКО-ЛАТИНСКИЙ КВАДРАТ И КВАДРАТЫ ВЫСШИХ ПОРЯДКОВ

Рассмотрим латинский квадрат табл. 12.10. Можно построить другой латинский квадрат из греческих букв (табл. 12.11) так, что, по наложении квадратов друг на друга, каждая латинская буква встречается по одному разу с каждой греческой и наоборот (табл. 12.12). Такой квадрат называется „греко-латинским“.

ТАБЛИЦА 12.10

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

ТАБЛИЦА 12.11

α	β	γ	δ	ε
δ	ε	α	β	γ
β	γ	δ	ε	α
ε	α	β	γ	δ
γ	δ	ε	α	β

ТАБЛИЦА 12.12

$A\alpha$	$B\beta$	$C\gamma$	$D\delta$	$E\varepsilon$
$B\delta$	$C\varepsilon$	$D\alpha$	$E\beta$	$A\gamma$
$C\beta$	$D\gamma$	$E\delta$	$A\varepsilon$	$B\alpha$
$D\varepsilon$	$E\alpha$	$A\beta$	$B\gamma$	$C\delta$
$E\gamma$	$A\delta$	$B\varepsilon$	$C\alpha$	$D\beta$

Ясно, какую роль может сыграть такой квадрат в планировании проведения испытаний. Если, как это было в предыдущем примере, различные столбцы представляли различные углы, строки — давления, а латинские буквы — различные партии материала, то греческие буквы могут представлять любой дополнительный фактор, например, скорость прессования.

Дисперсионный анализ проводится точно так же, как описано в предыдущем разделе, за исключением того, что теперь существует четвертый член, с $n-1$ степенями свободы, числовые значения которого получаются таким же путем, что и значения первых трех членов в табл. 12.9. Сумма квадратов для остатка, получаемая, как и ранее, в виде разности между итоговой суммой квадратов и суммой сумм квадратов всех других компонент, имеет $(n-1)(n-3)$ степеней свободы.

Можно наложить на табл. 12.12 еще пятый фактор так, чтобы он удовлетворял соответствующим условиям; число степеней свободы для остатка станет равным $(n-1)(n-4)$, или, в данном случае, 4.

Вообще говоря, кроме случая $n=6$, на любой квадрат со стороною n можно наложить n дополнительных факторов, оставляя еще $n-1$ степеней свободы для остатка. Эти приемы могут быть чрезвычайно полезны для приближенных, грубых, исследований; например, в серии из 25 наблюдений можно получить хорошие оценки 5 факторов, при 5 уровнях для каждого. Недостатком приемов такого рода является отсутствие оценки взаимодействий; тем не менее, такие приемы ценны для быстрого проведения исследования в какой-нибудь новой области, где прежде всего интересуются основными эффектами и намереваются исследовать позднее более подробно те из них, которые окажутся практически важными.

В таблицах Фишера и Йэйтса (табл. XVI) дано построение таких квадратов для значений n от 3 до 9. Замечательно, что для $n=6$ кроме обычного латинского квадрата нельзя построить никакого квадрата высшего порядка.

(I) ТЕОРИЯ ВЫБОРА ХИМИЧЕСКИХ ОБРАЗЦОВ *

(I) Рассмотренные выше способы анализа могут служить основой научного изучения техники извлечения выборок. Эта область еще не привлекла к себе достаточно внимания в химической промышленности.

Рассмотрим, например, дисперсии, выведенные в задаче в главе XI, (F).

Между баками:

$$\sigma_1^2 = 1{,}09.$$

* С точки зрения теории выбора интересны следующие статьи: D. J. Finney, *Biometrics Bull.*, **2**, № 1 (1946), 1; C. S. Pierse, Journ. of Royal Stat. Soc., CVII (1944), 117; F. Yates, там же, CIX (1946), 12. См. также G. W. Snedecor, Statistical Methods, 4th ed., ch. 17.

Между мешками, в пределах, взятых из одних и тех же баков:

$$\sigma_2 = 0{,}00.$$

Между образцами, в пределах, взятых из одних и тех же мешков:

$$\sigma_3^2 = 1{,}85.$$

Между отдельными анализами одних и тех же образцов:

$$\sigma_4^2 = 0{,}83.$$

Пусть имеется анализ по одной из выборок, взятой из какого-то мешка, относящегося к какому-то баку. Ее дисперсия будет составлять $1{,}09+0{,}00+1{,}85+0{,}83=3{,}77$; стандартное отклонение будет равно 1,94. Мы получим 95% шансов за то, что результат анализа будет отличаться от истинного среднего значения менее чем на $1{,}94 \cdot 1{,}96 = 3{,}8$. (Величина 1,96 является значением t для неограниченного числа степеней свободы и при уровне значимости в 5%.)

Если эта оценка качества партии слишком неточна, погрешность могла бы быть уменьшена путем проведения дублированных анализов. В таком случае дисперсия стала бы равной $1{,}09+0{,}00+1{,}85+\frac{0{,}83}{2}=3{,}35$, а стандартное отклонение — 1,83. Очевидно, что полученное улучшение точности, по сравнению с первоначальной оценкой, 1,94, основанной на одиночном анализе, очень невелико.

Можно было бы взять из мешка пару выборок и подвергнуть их одиночному анализу. Дисперсия получила бы значение $1{,}09+\frac{1{,}85+0{,}83}{2}=2{,}43$, а стандартное отклонение 1,56. Это — уже определенное улучшение прежней оценки.

Или же можно было взять пару выборок из разных мешков, взятых из разных баков. Оценка среднего значения качества партии теперь имела бы дисперсию $\frac{1{,}09+0{,}00+1{,}85+0{,}83}{2}=1{,}88$. Стандартное отклонение было бы 1,37.

Ясно, что, если лаборатории дается задание делать два полных анализа, менее выгодно делать дублированные анализы одной и той же выборки или отдельные анализы двух выборок, взятых из одного и того же мешка; лучше поручить лаборатории сделать отдельные анализы выборок, взятых из разных мешков, относящихся, в свою очередь, к разным бакам.

Если имеют дело с химическим веществом, выборки которого можно смешать друг с другом (жидкости, порошки и т. п.), открывается еще одна возможность. Можно взять пару выборок из разных мешков, взятых из разных баков, сделать из них смесь и провести анализ смеси.

Оценка получит теперь дисперсию погрешности, равную $\frac{1,09+0,00+1,85}{2}+0,83=2,3$. Стандартное отклонение будет 1,52. Или же, если взять четыре выборки из четырех баков, сделать из них смесь и провести анализ смеси, оценка получит дисперсию погрешности, равную $\frac{1,09+0,00+1,85}{4}+0,83=1,56$. Стандартное отклонение станет равным 1,25. Таким образом, при небольшом усложнении техники выбора мы получили оценку среднего качества партии со стандартным отклонением погрешности 1,25, вместо 1,94, полученного от одиночной оценки одной выборки; при этом объем работы лаборатории остался прежним.

Именно на базе такого дисперсионного анализа можно вывести, в чем заключается ошибочность порядка взятия выборок; если ошибка слишком велика, дисперсионный анализ покажет наиболее экономичный метод улучшения точности.

(II) Ниже приводятся еще примеры приложения дисперсионного анализа к изучению техники выбора химических проб. Смесь, считающаяся однородной, упаковывалась в картузы. Для проверки однородности сделали три анализа по каждой из трех выборок, взятых из каждого из трех картузов. Материалом была нитроцеллюлоза; анализ делался на содержание азота и вязкость.

СОДЕРЖАНИЕ АЗОТА (в %)

Выборка \ Картуз	1	2	3
I	12,20 12,22 12,21	12,23 12,24 12,23	12,19 12,21 12,22
II	12,21 12,20 12,19	12,22 12,23 12,25	12,24 12,21 12,23
III	12,21 12,22 12,20	12,22 12,21 12,24	12,22 12,22 12,22

ВЯЗКОСТЬ (В ПУАЗАХ)

Выборка \ Картуз	1	2	3
I	116 112 114	124 130 126	109 128 119
II	128 127 132	134 114 127	114 127 119
III	116 118 117	126 129 132	119 119 130

Данные были подвергнуты анализу, описанному в главе XI, (I); результаты получились следующие:

Источники дисперсии	Число степеней свободы	Содержание азота: суммы квадратов	Содержание азота: средние квадратов	Вязкость: суммы квадратов	Вязкость: средние квадратов
Между картузами	2	24,51	12,25	267,56	133,78
Между выборками в пределах картузов	6	10,23	1,70	427,11	71,19
Между анализами в пределах выборок	18	25,33	1,40	614,00	34,11
Итого	26	60,07		1308,67	

Содержание азота — в сотых долях процента.

В части содержания азота среднее квадратов между выборками незначимо; но значение среднего квадратов между картузами значимо и соответствует уровню значимости в 1%. В части вязкости значение среднего квадратов между выборками незначимо, а значение среднего квадратов между картузами почти значимо, при уровне значимости в 5%.

Таким образом, имеется заметная неоднородность смеси в различных картузах; неоднородность в пределах картузов незначительна. Можно подсчитать компоненты дисперсии:

Источники дисперсии	Содержание азота	Вязкость
Между картузами	1,20	10,0
Между выборками	0,00	0,0
Остаточная	1,48	43,4
Итого	2,68	53,4

В строке „Итого“ указана дисперсия одиночного анализа выборки, взятой из какого-либо картуза.

Чтобы осуществить улучшение в содержании азота, нужно сосредоточить внимание на обоих действующих источниках погрешности. Недостаточно взять n выборок, смешать их и провести анализ смешанной выборки; например, если $n=4$, полученная дисперсия оценки была бы $1,48+\frac{1,20}{4}=1,78$. Это — некоторое улучшение, но сколько бы выборок ни входило в смесь, дисперсия никогда не будет менее 1,48, пока проводится только один анализ. Если же сделать два анализа смешанной выборки, составленной из содержимого четырех картузов, дисперсия среднего значения станет равной

$$\frac{1{,}20}{4}+\frac{1{,}48}{2}=1{,}04.$$

Обращаясь к определению вязкости, видим, что основное внимание должно быть уделено получению большего числа анализов. Если бы имелось четыре анализа одной выборки, дисперсия оценки средней была бы $10+\frac{43{,}4}{4}=20{,}8$; четыре анализа отдельных выборок дали бы величину дисперсии $\frac{10}{4}+\frac{43{,}4}{4}=13{,}3$. Поэтому выигрыш от взятия отдельных выборок получается заметным.

(J) ОДНОРОДНОСТЬ ДАННЫХ

В конце главы III, в которой рассматривалось сравнение средних, было указано на возникновение осложнений в расчете стандартных отклонений в случаях, когда наблюдения не являются взаимно независимыми.

В общем случае при проведении дисперсионного анализа между партиями материала и по партиям (один фактор с повторениями) и при наличии всего m партий из n индивидуумов каждая мы обозначим величину простой дисперсии, определяемой итоговым значением среднего квадратов, через σ_t^2. Это — величина дисперсии, которая получилась бы в расчете, если игнорировать неоднородность данных; ее можно было бы рассматривать относящейся к nm независимых индивидуумов. Обозначая компоненты дисперсии, соответствующие различиям между партиями и по партиям соответственно через σ_A^2 и σ_B^2, получим соотношение

$$\sigma_t^2=\frac{m(n-1)}{mn-1}\,\sigma_A^2+\sigma_B^2=\left[1-\frac{m-1}{mn-1}\right]\sigma_A^2+\sigma_B^2.$$

Очевидно, что σ_t^2 всегда меньше, чем $\sigma_T^2=\sigma_A^2+\sigma_B^2$; величина ошибки будет наибольшая, когда $\frac{\sigma_B^2}{\sigma_A^2}$ стремится к нулю, а также когда m велико, а $n=2$. В этом случае σ_t^2 равняется только половине правильного значения общей дисперсии σ_T^2.

Обратимся к группе из 30 наблюдений, рассматривавшейся в главе VII, (B) (анализ между столбцами и по столбцам). Допустим, желают подсчитать стандартное отклонение среднего значения этих наблюдений.

Если не принять во внимание того, что эти 30 наблюдений не являются независимыми друг от друга, а составляют в действительности 10 групп, каждая из которых состоит из трех наблюдений, то дисперсия должна подсчитываться, исходя из ее обычного определения; это соответствует разности величин

(1) — (3), вычисленных в гл. VII, (B), деленной на общее число степеней свободы: $\frac{203{,}87}{29}=7{,}03$; стандартное отклонение индивидуального наблюдения будет 2,65, а стандартное отклонение среднего было бы равным $\frac{2{,}65}{\sqrt{30}}=0{,}484$.

Однако это — неверный результат, так как наше среднее не является средним от 30 независимых индивидуумов, но только средним от 10 средних по столбцам.

Дисперсия каждого среднего по столбцу равна $\frac{\sigma_B^2}{3}+\sigma_A^2$; поскольку имеется 10 таких средних, то дисперсия общего среднего равна $\frac{\frac{1}{3}\sigma_B^2+\sigma_A^2}{10}=\frac{\frac{1}{3}2{,}17+5{,}22}{10}=0{,}594$. Стандартное отклонение среднего, таким образом, равно $\sqrt{0{,}594}=0{,}771$.

Эта истинная стандартная ошибка среднего 0,771 намного больше неверного значения 0,484, полученного без учета неоднородности данных. Если бы эту неверную стандартную ошибку использовали для сравнения среднего с некоторой другой величиной посредством критерия t Стьюдента, получились бы серьезные погрешности; дело в том, что не только правильное значение стандартной ошибки среднего намного больше неправильно подсчитанного, но и истинное число степеней свободы составляет только 9, а не 29. Именно поэтому весьма желательно вообще планировать эксперименты этого типа в форме, удобной для применения дисперсионного анализа.

Таким образом, если рассматриваемые 30 единиц (10 партий продукта) были бы сделаны из одной партии сырого материала и требовалось бы сравнить среднее другой партии сырого материала, то было бы удобнее всего сделать из последней тоже 10 партий продукта (30 единиц). При этом было бы удобно провести следующий дисперсионный анализ.

Источники дисперсии	Числа степеней свободы	Суммы квадратов	Средние квадратов
Между партиями сырьевого материала	$2-1=1$		
Между партиями продукта в пределах партий сырьевого материала	$2\cdot(10-1)=18$		
Между единицами продукта в пределах партий продукта	$2\cdot10\cdot(3-1)=40$		
Итого	$2\cdot10\cdot3-1=59$		

12 Зак. 1691

Заметим, что применение этого приема было бы невозможно, если бы эксперимент проводился наудачу, без планирования; единицы из некоторых партий могли бы отсутствовать, и количества партий продукта, изготовленных из разных партий сырья, были бы неравными.

(K) ПРИМЕНЕНИЕ ЛОГАРИФМОВ В ДИСПЕРСИОННОМ АНАЛИЗЕ

Ранее рассматривалась возможность использования вместо самой зависимой переменной ее логарифма; это относилось к случаям, когда отклонения имеют склонность быть пропорциональными среднему значению, т. е. постоянными при использовании логарифмов.

Такой пример был в случае анализа производства кислотного завода (глава XI, (I), табл. 11.22). Член дисперсии между цехами в пределах групп кислоты был незначим, поэтому можно было взять объединенную остаточную дисперсию с $3 \cdot 6 = 18$ степенями свободы для каждого типа слабой кислоты. Эти величины равны соответственно 57,19, 405,21 и 1197,84. Они совершенно несовместимы при проверке по методу отношения дисперсий, однако стандартные отклонения равны 7,56, 20,13 и 34,61. Если их выразить в виде долей соответствующих средних величин 26,7, 51,0 и 103,3, получатся значения 0,294, 0,396 и 0,334. Они приблизительно постоянны, что наталкивает на мысль попытаться анализировать логарифмы данных.

В качестве функций, наиболее удобных для вычисления, были выбраны натуральные логарифмы от одной десятой значения выпуска продукции. Как и прежде, для каждого из трех типов кислоты были подсчитаны остаточные дисперсии; каждая из них имела 18 степеней свободы; их значения оказались: 0,04489, 0,22450 и 0,08891. Хотя они и не вполне совместимы, но все же в гораздо большей степени, чем сами переменные. Дисперсионный анализ по логарифмам данных приведен в табл. 12.13.

ТАБЛИЦА 12.13

Дисперсионный анализ логарифма выпуска продукции кислотного завода

Источники дисперсии	Числа степеней свободы	Суммы квадратов	Средние квадратов
Тип слабой кислоты	2	11,4670	5,7335
Цехи	6	0,6769	0,1128
В пределах цехов	54	6,4494	0,1194
Итого	62	18,5933	

Выводы из этого анализа совпадают с заключениями, полученными из табл. 11.23, содержащей анализ простых данных.

Это дает дополнительную уверенность в применимости дисперсионного анализа: можно довольно смело не принимать во внимание предварительные допущения.

(L) ДРУГИЕ ПРЕОБРАЗОВАНИЯ В ДИСПЕРСИОННОМ АНАЛИЗЕ *

В дисперсионном анализе иногда случается, что зависимая переменная изменяется не непрерывно, а дискретно.

Например, в случае исследования пятен на производимых предметах, число пятен может меняться от нуля до очень больших, иногда, количеств. Технически, вероятно, получится распределение Пуассона. В таких условиях более удобно брать в качестве зависимой переменной квадратный корень из числа пятен **. Если число пятен во многих случаях меньше 10, несколько предпочтительнее использовать в качестве зависимой переменной величину $\sqrt{\frac{x+1}{2}}$, где x — число пятен.

Если число подлежащих исследованию случаев x имеет конечный верхний предел X и нижний предел 0, лучше всего брать в качестве зависимой переменной угол $\theta = \arcsin \sqrt{\frac{x}{X}}$. Например, если $x = 40$ и $X = 120$, то $\theta = 35{,}2°$.

Для облегчения преобразований в практике существуют удобные таблицы (см. Fisher and Yates, Statistical Tables for Biological, Agriculture, and Medical Research, 2-е изд., таблица XII, а также G. W. Snedecor, Statistical Methods, 4-е изд., стр. 449).

(M) УПУЩЕННЫЕ ЧИСЛОВЫЕ ЗНАЧЕНИЯ ***.

Случайно те или иные опыты могут не дать результата; испытуемый образец может быть утерян, или встретится какая-либо иная неожиданность. В случае испытания с двумя факторами, при r строках и c столбцах, одно упущенное числовое значение может быть рассчитано по формуле (см. F. Yates, *Empire Journ. of Experimental Agriculture*, 1, (1933), 129):

$$X = \frac{rR + cC - T}{(r-1)(c-1)},$$

* О различных преобразованиях в испытаниях $3 \times 3 \times 2 \times 2$ см. статью Thorp и др., *Phytopathology*, XXXI (1941), 26.

** M. S. Bartlett, *Suppl. to the Journ. of the Roy. Stat. Soc.* (1936).

*** Подробное изложение вопроса с библиографией см. R. L. Anderson, *Biometrics Bulletin*, 2, № 3, (1946) 41.

где r — число строк,
c — число столбцов,
R — итог по строке с упущенным членом,
C — итог по столбцу с упущенным членом,
T — общий итог.

Так, например, если бы в табл. 12.1 значение величины T_2W_2 (результат наблюдения был в действительности 27) было упущено, мы могли бы его оценить так:

$$X = \frac{4 \cdot 131 + 4 \cdot 78 - 481}{(4-1) \cdot (4-1)} = 39{,}4.$$

Для проведения дисперсионного анализа это число вписывается в таблицу, как будто она представляет собой результат действительного наблюдения; затем, дисперсионный анализ проводится обычным путем. Но числа степеней свободы для остатка и для итога уменьшаются на единицу.

Благоразумнее относиться к заключениям, выведенным из исследования с упущенным наблюдением, с большей осторожностью, чем в случае, когда исследование было полным. Мы хотим сказать, что лучше приписать значимость порядка 5% тем выводам, которые окажутся с уровнем значимости 1%.

Если упущены два результата наблюдений, их можно восстановить методом итерации. Вписываем некоторое предполагаемое значение вместо упущенного X_1, например, генеральное среднее; затем, рассчитываем ожидаемое значение X_2. Вписав это значение X_2, вычеркиваем предположенное ранее значение X_1 и рассчитываем его ожидаемое значение X_1'. Вписав это значение X_1', вычеркиваем прежде записанное значение X_2 и рассчитываем его новое значение X_2'. Вписав X_2', вычеркиваем X_1' и рассчитываем новое значение для X_1, X_1'', и т. д. После нескольких итераций новые значения перестанут отличаться от предыдущих. Дисперсионный анализ проводится затем при числах степеней свободы для остатка и итога, уменьшенных на 2.

Очевидно, что восстановленное значение всегда будет менее доброкачественно, чем результат действительного наблюдения; поэтому, следует приложить все усилия к тому, чтобы избежать работы с неполными данными. Если недостаток данных принял большой масштаб, то простейший случай может быть проанализирован строго, с небольшими осложнениями, как описано в главе VII, (D). Снедекор (G. W. Snedecor. „Statistical Methods applied to Experiments in Agriculture and Biology“. Iowa State College Press, 4-е изд., 1946) рассматривает ряд более сложных случаев. Однако, в практике можно часто получить довольно удовлетворительный результат, выбрасывая наудачу несколько данных из классов, содержащих большие количества данных, так что останутся только одно или два упущенных значения, которые можно восстановить с помощью описанных методов.

ГЛАВА XIII

СБАЛАНСИРОВАННЫЕ НЕПОЛНЫЕ БЛОКИ

(А) ВВЕДЕНИЕ

Общее описание метода проведения испытаний, известного под названием „метода сбалансированных неполных блоков“ было дано в главе I, (F). В настоящей главе будут даны способы проведения расчетов, связанных с применением метода, основанные на работе Йэйтса (*Ann. of Eugenics*, **7**, (1936), 121).

Йэйтс (*Ann. of Eugenics*, **10**, (1940), 317) разработал гораздо более чувствительный метод анализа таких результатов, но в тех случаях, когда число блоков не очень велико, выигрыш в точности незначителен, и воспроизводимый нами более старый метод обычно оказывается достаточным.

(В) ПОРЯДОК ПРОВЕДЕНИЯ РАСЧЕТОВ

В таблице 13.1 приведены результаты испытаний по сравнению 7 способов обработки в 7 блоках, содержащих по 3 единицы в каждом; таким образом, имеем 3 повторения по каждому способу обработки. Блок А, например, включает в себя три способа обработки T_1, T_3 и T_5; цифры, стоящие в столбцах A — G таблицы, выражают полученные результаты испытаний.

В обозначениях Йэйтса:

Количество сравниваемых способов обработки	$= t = 7$
Число единиц в блоке	$= k = 3$
Число повторений по каждому способу обработки	$= r = 3$
Число блоков	$= b = 7$
Общее число единиц	$= N = tr = bk = 21$
Итог по всем N наблюдениям	$= G = 1454$
Итог по всем r наблюдениям по первому способу обработки	$= T_1$
Итог по всем k наблюдениям по первому блоку	$= B_1$
Дисперсия ошибки одиночного наблюдения по блокам из k единиц	$= \sigma_k^2$

ТАБЛИЦА 13.1

Блоки / Способы обработки	A	B	C	D	E	F	G	Итого	Q	V	Корректированные средние по способам обработки
T_1	50	42	91					183	— 84	— 12,0	57,238
T_2			118	94	94			306	163	23,285	92,523
T_3	76			64		80		220	69	9,857	79,095
T_4			72			53	31	156	— 160	— 22,857	46,381
T_5	44				65		54	163	— 81	— 11,571	57,667
T_6		102			119	92		313	251	36,286	105,524
T_7		38		38			37	113	— 161	— 23,0	46,238
Итого	170	182	281	196	278	225	122	1454	0	0,0	

Для каждого из t способов обработки подсчитываем величину Q по формуле типа:

$$Q_1 = kT_1 - B_1 - B_2 - \ldots - B_r,$$

где величины B для расчета Q являются итогами по r блокам, включающим первый способ обработки; например,

$$Q_1 = 3 \cdot 183 - 170 - 182 - 281 = -84.$$

Сумма квадратов по способам обработки равна:

$$\frac{t-1}{Nk(k-1)} \sum Q^2 = \frac{7-1}{21 \cdot 3 \cdot (3-1)} \cdot (84^2 + 163^2 + \ldots + 161^2) = 7665{,}905.$$

Сумма квадратов по блокам получается обычным способом из итогов по блокам:

$$\frac{\Sigma B^2}{k} - \frac{G^2}{N} = \frac{1}{3} \cdot (170^2 + 182^2 + \ldots + 122^2) - \frac{1454^2}{21} = 6725{,}81.$$

Общая сумма квадратов находится также обычным путем:

$$50^2 + 42^2 + 91^2 + 118^2 + \ldots + 37^2 - \frac{1454^2}{21} = 15057{,}81.$$

Вносим вычисленные величины в таблицу дисперсионного анализа; определяем сумму квадратов для ошибки и число степеней свободы для нее путем вычитания из итоговых значений соответствующих величин по блокам и способам обработки.

ТАБЛИЦА 13.2

Источник дисперсии	Число степеней свободы	Суммы квадратов	Средние квадратов
Блоки	6	6725,810	1120,968
Способы обработки . .	6	7665,905	1277,651
Ошибка	8	666,095	83,261
Итого	20	15057,81	

Среднее квадратов для ошибки обозначается через σ_k^2.

Стандартная ошибка различия между любыми двумя средними по способам обработки равна:

$$\sqrt{2} \cdot \sqrt{\frac{k \cdot (t-1)}{N(k-1)} \cdot \sigma_k^2} = \sqrt{2} \cdot \sqrt{\frac{3(7-1)}{21(3-1)} \cdot 83{,}261} = 8{,}448.$$

Чтобы найти эти средние по способам обработки, скорректированные в отношении различий между блоками, вычисляем t величин следующего типа:

$$V_1 = \frac{t-1}{N(k-1)} Q_1 = \frac{7-1}{21(3-1)} \cdot (-84) = -12{,}0.$$

Эти величины V представляют собой отклонения каждого из t способов обработки от генерального среднего, которое равно $\frac{G}{N} = \frac{1454}{21} = 69{,}238$. Скорректированное среднее по первому способу обработки теперь равно:

$$V_1 + \frac{G}{N} = -12{,}0 + 69{,}238 = 57{,}238.$$

Если требуется определить также средние по блокам, то, при проведении испытаний такого типа, как в настоящем примере, т. е. когда каждая пара блоков включает одинаковое общее число способов обработки, можно оценить скорректированные средние по блокам путем подсчета величин Q', аналогичных ранее вычисленным Q. Например:

$$Q'_A = 3 \times 170 - 183 - 220 - 163 = -56$$

и

$$V'_A = Q_A \frac{t-1}{N(k-1)} = 56 \cdot \frac{(7-1)}{21(3-1)} = -8{,}0.$$

Отсюда, скорректированное среднее для блока A равно

$$69{,}238 - 8{,}0 = 61{,}238.$$

Более общим путем, здесь можно получить среднее для блока так:

$$\frac{1}{3} \cdot (170{,}0 - 57{,}238 - 79{,}095 - 57{,}667) + 69{,}238 = 61{,}238.$$

В данном частном случае, когда каждая пара блоков включает одинаковое общее число способов обработки, стандартная ошибка разности между скорректированными средними по блокам такова же, что и определенная выше стандартная ошибка для скорректированных средних по способам обработки.

В более общем случае, дисперсия, в среднем, равна

$$\overline{V}_b = \frac{2\sigma_k^2}{k} \cdot \left(1 + \frac{(t-1)(t-k)}{t(k-1)(b-1)}\right).$$

Подсчет по этой формуле в настоящем случае даст те же результаты, что были получены ранее.

(С) ВОЗМОЖНЫЕ ПЛАНЫ ПРОВЕДЕНИЯ ИСПЫТАНИЙ

Перечень возможных планов проведения испытаний дан в таблицах Фишера и Йэйтса. Вообще говоря, если число способов обработки и объем блока известны, то минимальное число повторений по каждому способу обработки тем самым уже определено. Если стремятся иметь лишь небольшое число повторений, непосредственное решение можно часто обойти. Например, при блоках объема 4 и сравнении 6 способов обработки, для сбалансирования требуется 10 повторений. Если считать, что это число повторений, 10, превышает необходимое для поставленной цели, можно взамен применить план проведения испытаний для 7 способов обработки, который потребует только 4 повторения; седьмой способ обработки может быть, по желанию, либо повторением одного из шести, либо каким-нибудь дополнительным.

Наиболее полезными являются планы проведения испытаний, показанные в таблицах 13.3, а также уже рассмотренный план с характеристиками: $k=3$, $b=7$, $t=7$, $r=3$ (т. е. 3 единицы в блоке, 7 блоков, 7 способов обработки, 3 повторения).

ТАБЛИЦА 13.3

$k=3;\ b=4;\ t=4;\ r=3$

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

$k=4;\ b=5;\ t=5;\ r=4$

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

$k=4;\quad b=7;\quad t=7;\quad r=4.$

A		*	*	*			*
B	*	*		*		*	
C	*		*			*	*
D		*	*		*	*	
E				*	*	*	*
F	*		*	*	*		
G	*	*			*		*

Таблицы Фишера и Йэйтса содержат ряд других интересных планов проведения испытаний. Например, 9 способов обработки можно подвергнуть сравнению с помощью блоков объема 3, с четырьмя повторениями; 13 способов обработки можно сравнивать с помощью блоков объема 4, с четырьмя повторениями; 16 способов обработки — с помощью блоков объема 4, с пятью повторениями; 21 способ обработки — с помощью блоков объема 5, с пятью повторениями.

(D) ДРУГИЕ ПРИМЕНЕНИЯ СИММЕТРИЧНЫХ НЕПОЛНЫХ БЛОКОВ

Хотя метод неполных блоков был введен в целях избежания затруднений, возникающих из-за ограниченности объема блоков, его можно использовать и для иных целей. Например, в случае испытания с двумя факторами, при 7 уровнях для каждого фактора, полный план потребовал бы проведения $7^2=49$ испытаний. Однако, можно получить удовлетворительные оценки двух факторов при помощи рассматриваемого метода неполных блоков, что потребует только 21 испытание, причем дисперсия разности между двумя скорректированными средними будет $\frac{6}{7}\sigma_k^2$. Проведение испытания полностью дало бы величину дисперсии $\frac{2}{7}\sigma_k^2$; хотя ошибка при применении метода неполных блоков получается больше, все же она может оказаться приемлемой для поставленных целей.

Альтернативным классическим решением этой проблемы было сравнение семи способов обработки $T_1, \dots, T_7$, при постоянном $B=B_A$ и последующее сравнение семи блоков $B_A, \dots, B_G$ при постоянном $T=T_1$; это потребовало бы 13 испытаний. Без повторений такой эксперимент был бы в высшей степени неудовлетворительным, так как нет оценки ошибки и отсутствуют данные, необходимые для проверки значимости. Если повторить наблюдения, сделав всего 26 испытаний, получится возможность сравнения средних с дисперсией σ_k^2. Сравнивая этот план проведения испытаний с планом по методу неполных блоков видно, что потребуется на 24% увеличить объем работы, а дисперсия за счет ошибки получится на 14% выше.

ГЛАВА XIV

СМЕШИВАНИЕ. ПРОБЛЕМА ОГРАНИЧЕННОСТИ ОБЪЕМА БЛОКА В ФАКТОРНЫХ ИСПЫТАНИЯХ*.

(A) АЛГЕБРАИЧЕСКИЕ ВЫРАЖЕНИЯ ДЛЯ ФАКТОРОВ

В главе I было указано, что количество наблюдений, требующихся для факторного испытания „типа 2^n" (n факторов, каждый с двумя уровнями), равно 2^n; при увеличении числа факторов до 4, 5 или 6, количество наблюдений быстро возрастает, достигая соответственно 16, 32 и 64. Эти числа могут быть слишком велики для включения в одиночный блок; в связи с этим возникает проблема „расщепления" эксперимента на более мелкие единицы.

* О вопросах смешивания см. R. A. Fisher, The Design of Experiments, ch. VI, и F. Yates, The Design and Analysis of Factorial Experiments.

Введем некоторые обозначения. Пусть строчные буквы p, q, r и т. д. указывают, что соответствующий фактор находится на его верхнем уровне. Например, сочетание rs будет обозначать, что p и q находятся на их нижних уровнях, а r и s — на верхних. Пусть прописная буква P указывает разность между суммой всех верхних уровней фактора P и суммой всех его нижних уровней. Аналогичные обозначения будут и для других факторов.

Для испытания с тремя факторами:

$$P=pqr+pq+pr+p-qr-q-r-1,$$

где символ 1 обозначает, что все факторы взяты на их нижних уровнях. Алгебраически это выражение эквивалентно следующему:

$$P=(p-1)(q+1)(r+1).$$

Аналогично

$$Q=(p+1)(q-1)(r+1),$$
$$R=(p+1)(q+1)(r-1).$$

Взаимодействие первого порядка PQ выражается разностью между P при нижнем уровне Q (равном $pr+p-r-1$) и P при верхнем уровне Q (в свою очередь, равном $pqr+pq-qr-q$). Таким образом,

$$PQ=pqr+pq-qr-q-pr-p+r+1.$$

Алгебраически это выражение эквивалентно следующему:

$$PQ=(p-1)(q-1)(r+1).$$

Заметим, что можно получить точно такой же результат, если рассматривать PQ как разность между Q при нижнем уровне P (равном $qr+q-r-1$) и между Q при верхнем уровне P (равном $pqr+pq-pr-p$). Действительно, эта разность равна:

$$PQ=pqr+pq-qr-q-pr-p+r+1=(p-1)(q-1)(r+1).$$

Аналогично выводим:

$$QR=(p+1)(q-1)(r-1)$$
$$RP=(p-1)(q+1)(r-1).$$

Вывод выражения для взаимодействия второго порядка PQR гораздо сложнее. Такое взаимодействие можно рассматривать как разность между значением PQ при нижнем уровне R и значением PQ при верхнем уровне R. Первое равно

$$(PQ)_{R_1}=pq-q-p+1,$$

второе —

$$(PQ)_{R_2}=pqr-qr-pr+r.$$

Разность их есть

$$PQR=pqr-qr-pr+r-pq+q+p-1=(p-1)(q-1)(r-1).$$

Взаимодействия второго порядка можно рассматривать так: например, PQR можно рассматривать как разность между значениями QR при P_1 и QR при P_2, или как разность между значениями RP при Q_1 и RP при Q_2.

(B) СМЕШИВАНИЕ С ТРЕМЯ ФАКТОРАМИ

Пусть требуется провести испытания с тремя факторами, которые потребуют $2^3 = 8$ наблюдений, но объем блока ограничен четырьмя наблюдениями; поэтому мы хотим рассмотреть два блока, по 4 испытания в каждом. Если нас интересуют главным образом основные эффекты и взаимодействия первого порядка и для достижения поставленной цели можно пожертвовать взаимодействием второго порядка, испытания можно распределить по двум блокам так, как показано в табл. 14.1.

ТАБЛИЦА 14.1

Блок 1	Блок 2
pqr	1
r	qr
p	pr
q	pq

Взаимодействие второго порядка теперь утеряно, так как оно тождественно с возможным различием между блоками; действительно,

$$pqr + r + p + q - 1 - qr - pr - pq$$

выражает как PQR, так и различие между блоками 1 и 2. Этим объясняется название плана проведения испытаний, поскольку здесь произошло *смешивание* эффекта с различием между блоками.

Теперь следует удостовериться в том, что возможные различия между блоками не повлияют на оценки основных эффектов и взаимодействий первого порядка. Предположим, что блок 2 имеет систематическое искажение, так что все содержащиеся в нем четыре результата оказались на величину d выше, чем они были бы в случае проведения этих же испытаний в блоке 1. Подсчитываем теперь величины P и PQ:

$$P = pqr + (pq + d) + (pr + d) + p - (qr + d) - q - r - (1 + d) = \\ = pqr + pq + pr + p - qr - q - r - 1,$$

$$PQ = pqr + (pq + d) - (qr + d) - q - (pr + d) - p + r + (1 + d) = \\ = pqr + pq - qr - q - pr - p + r + 1.$$

Эти выражения совпадают с выведенными ранее формулами; систематическое искажение d в обоих случаях исключено. Аналогичные результаты можно получить и для прочих эффектов.

(C) СМЕШИВАНИЕ С ЧЕТЫРЬМЯ ФАКТОРАМИ

При наличии четырех факторов число требующихся испытаний составляет $2^4 = 16$. Основные эффекты будут такого типа:

$$P = (p - 1)(q + 1)(r + 1)(s + 1).$$

Взаимодействия первого порядка будут типа:

$$PQ = (p - 1)(q - 1)(r + 1)(s + 1).$$

Взаимодействия второго порядка будут типа:

$$PQR = (p - 1)(q - 1)(r - 1)(s + 1).$$

Взаимодействия третьего порядка будут:

$$PQRS = (p - 1)(q - 1)(r - 1)(s - 1) = pqrs - pqr - pqs + pq - \\ - prs + pr + ps - p - qrs + qr + qs - q + rs - r - s + 1.$$

Если мы хотим произвести „смешивание“ в 2 блока по 8 испытаний в каждом и жертвуем при этом членом $PQRS$, то можно все положительные члены из выражения $PQRS$ сосредоточить в одном блоке, а все отрицательные — в другом, как показано на таблице 14.2.

ТАБЛИЦА 14.2

Блок 1	$pqrs + pq + pr + ps + qr + qs + rs + 1$
Блок 2	$pqr + pqs + prs + qrs + p + q + r + s.$

Чтобы произвести „смешивание“ в 4 блока по 4 испытания в каждом, мы выбираем 3 члена, которыми можно пожертвовать, в соответствии с 3 степенями свободы, „смешивающимися“ с различиями между блоками. Обнаружено, что два таких члена можно выбрать произвольно, в то время как третий уже оказывается однозначно определенным; он определяется по следующему правилу: надо взять алгебраическое произведение первых двух выбранных членов и заменить единицей каждый множитель, стоящий в квадрате.

Например, если двумя произвольно выбранными членами являются PQ и RS, то третьим членом будет $PQRS$; если двумя выбранными являются PQR и QRS, то третьим членом будет PS.

Легко вывести, что в случае смешивания в 4 блока, по 4 испытания в каждом, одно из взаимодействий первого порядка неизбежно окажется „смешанным“.

Один из методов распределения испытаний по блокам заключается в раскрытии скобок в выражениях для двух предназначенных к смешиванию взаимодействий, как показано в табл. 14.3.

ТАБЛИЦА 14.3

PQR QRS	$pqrs-pqs+pqr-pq-qrs+qs-qr+q$ $pqrs-pqs-pqr+pq+qrs-qs-qr+q$
$PQR*$ $QRS*$	$-prs+ps-pr+p+rs-s+r-1$ $-prs+ps+pr-p-rs+s+r-1$

[Значения со звездочками * выписываются с обратными знаками и с исключенным общим множителем, в данном случае q. — *Прим. перев.*].

Испытания, символы которых имеют в верхней и нижней паре выражений табл. 14.3 знаки $+$, помещаются в один и тот же блок. Испытания, символы которых характеризуются знаками $--$, $+-$, и $-+$, включаются соответственно во второй, третий и четвертый блоки (см. табл. 14.4).

ТАБЛИЦА 14.4

Блок 1 $++$	Блок 2 $--$	Блок 3 $+-$	Блок 4 $-+$
$pqrs$ q ps r	pqs qr prs 1	pqr qs p rs	pq qrs pr s

Существует альтернативный метод распределения по блокам, более удобный в сложных случаях. В данном случае мы избрали для смешивания три взаимодействия

$$PQR,\quad QRS \quad \text{и} \quad PS.$$

Включаем в первый блок те испытания, которые имеют четное число букв, одинаковых с буквами, входящих в символы трех выбранных членов. Нуль считается четным числом. Испытание (1) (т. е. при нижних уровнях всех факторов) удовлетворяет этому условию, так как имеет O общих букв с PQR, QRS и PS. Испытание qr также удовлетворяет этому условию, так как его символ имеет две общих буквы с PQR и QRS, и ни одной с PS. Двумя другими удовлетворяющими условию испытаниями будут prs и pqs. В результате получается блок, тождественный с блоком 2 в таблице 14.4.

Для определения состава следующего блока, выбираем какое-либо пока неиспользованное испытание, например p и умножаем на этот символ p каждый член, входящий в прежде составленный блок; при этом имеем в виду принятое ранее условие, что

квадрат любого символа считается равным единице. Таким образом, получаем:

$$\begin{aligned} pqs\times p &= p^2qs = qs,\\ qr\times p &= pqr,\\ prs\times p &= p^2rs = rs,\\ 1\times p &= p. \end{aligned}$$

Это — блок 3 таблицы 14.4. Состав следующего блока определяется опять-таки путем выбора какого-либо неиспользованного еще испытания, например q, и умножения на этот символ каждого члена первого из найденных блоков. Таким путем определяется блок 1 в таблице 14.4. Проведя аналогичную операцию еще с одним символом ранее неиспользованного испытания, например s, получаем состав последнего блока — блока 4 в таблице 14.4.

(D) СМЕШИВАНИЕ С ПЯТЬЮ ФАКТОРАМИ

При пяти факторах требуется $2^5=32$ испытания. Существует ряд возможностей смешивания, соответственно допускаемому объему блока.

Для разбивки на 2 блока, по 16 испытаний в каждом, очевидно, надо пожертвовать взаимодействием высшего порядка $PQRST$. Раскрывают скобки в его выражении

$$(p-1)(q-1)(r-1)(s-1)(t-1),$$

и включают все испытания, символы которых положительны, в один блок, а испытания, символы которых отрицательны — в другой блок. Тот же результат можно получить, включая в один блок все испытания, символы которых имеют четное число общих букв с символом $PQRST$; сюда войдут: (a) испытание (1), имеющее 0 общих букв с $PQRST$, (b) десять испытаний вида pq, qr, pr, ps, и т. д., (c) пять испытаний вида $pqrs$, $qrst$, и т. д.

Для разбивки на 4 блока, по 8 испытаний в каждом, лучше всего пожертвовать для смешивания взаимодействиями типа PQR, RST, $PQST$; при этом удается избежать потери взаимодействий первого порядка. Для получения распределения испытаний по четырем блокам раскрываем скобки в выражениях взаимодействий, подлежащих смешиванию, например PQR и RST; затем, распределяем испытания по блокам так, как было описано в предыдущем разделе: имеющие знаки $++$ в обоих выражениях — в один блок, имеющие знаки $+-$ в другой, $-+$ в третий, и $--$ в четвертый. Или же, включаем в первый блок все испытания, символы которых имеют четное число (или нуль) общих букв с символами взаимодействий, предназначенных к смешиванию. Сюда войдут испытания:

$$(1),\ pq,\ st,\ pqst,\ qrs,\ qrt,\ prs,\ prt.$$

Состав других блоков находится путем умножения этих символов на p (для второго блока), s (для третьего блока), r (для четвертого блока). Как и ранее, любой символ в квадрате должен считаться равным единице.

В случае разбивки на 8 блоков по 4 испытания в каждом, лучшим вариантом будет жертва двух взаимодействий первого порядка. Типичный комплект взаимодействий, предназначаемых для смешивания:

$$PQ,\ RS,\ PRT,\ QRT,\ PST,\ QST,\ PQRS.$$

Одним из способов распределения испытаний по 8 блокам будет раскрытие скобок в любых трех выражениях для написанных выше взаимодействий (выбор надо делать так, чтобы символ третьего взаимодействия не был произведением символов первых двух; например, если выбраны PQ и RS, третьим нельзя избрать $PQRS$). Затем, распределение по блокам делается в соответствии с расположением знаков $+$ и $-$ в трех выражениях, по комбинациям:

$$+++,\ \ ++-,\ \ +--,\ \ +-+,\ \ -++,\ \ -+-,$$
$$---\ \text{ и }\ --+.$$

Альтернативный способ: первый блок составят испытания

$$(1),\ \ pqt,\ \ rst,\ \ pqrs.$$

Состав других семи блоков определится путем умножения символов первого блока на символы, еще неиспользованные. Например, для получения состава второго блока выбирается множитель p; тогда во второй блок войдут испытания:

$$p,\ \ qt,\ \ prst,\ \ qrs.$$

Для третьего блока берется множитель q; его состав будет:

$$q,\ \ pt,\ \ qrst,\ \ prs.$$

Продолжая операции, находят состав всех прочих блоков.

(E) СМЕШИВАНИЕ С ШЕСТЬЮ ФАКТОРАМИ

При шести факторах потребуются $2^6=64$ испытания. При разбивке на 2 блока, по 32 испытания в каждом, очевидно следует пожертвовать для смешивания взаимодействием наивысшего порядка $PQRSTU$.

При разбивке на 4 блока, по 16 испытаний в каждом, наиболее удовлетворительным для смешивания комплектом взаимодействий будут: $PQRS$, $RSTU$, $PQTU$.

При 8 блоках, по 8 испытаний в каждом, подлежат смешиванию взаимодействия:

$$PQR,\ RST,\ PSU,\ QTU,\ PQST,\ QRSU,\ PRTU.$$

В перечисленных случаях удается избежать потери взаимодействий первого порядка.

Однако, для получения 16 блоков, по 4 испытания в каждом, неизбежна потеря трех взаимодействий первого порядка среди 15, предназначаемых к смешиванию: PQ, RS, PRT, QRT, PST, QST, $PQRS$, TU, $PQTU$, $RSTU$, PRU, QRU, PSU, QSU, $PQRSTU$.

Распределение испытаний по блокам делается тем же путем, что и ранее.

В первый блок входят испытания:

$$(1),\ \ pqrs,\ \ rstu,\ \ pqtu.$$

Состав последующих блоков находится путем умножения этих символов на символы еще неиспользованных испытаний.

(F) РАСЧЕТ РЕЗУЛЬТАТОВ ИСПЫТАНИЯ ПО МЕТОДУ СМЕШИВАНИЯ. ПРИМЕР

Расчет результатов испытания по методу смешивания производится точно так же, как будто смешивания не было, с единственным исключением, что взаимодействия, подвергшиеся смешиванию, не вычисляются. Вместо них берут итоги по блокам, возводят их в квадрат, делят на число способов обработки, входящих в каждый блок, и вычитают обычным образом подсчитанный корректирующий фактор (квадрат общего итога, деленный на общее число способов обработки). Число степеней свободы для этой суммы квадратов на единицу меньше числа блоков. Тождественный результат можно было бы получить, рассчитав сумму квадратов взаимодействий, подвергшихся смешиванию, и затем объединить их.

В качестве примера рассмотрим эксперимент по производственному процессу, связанному с применением смешанного катализатора, состоящего из трех компонент.

Компонента A катализатора имеет 4 уровня, компоненты C и D — по два уровня; выпуск продукции является зависимой переменной. Непосредственное факторное испытание потребовало бы $4\times2\times2=16$ наблюдений. Оказалось, однако, необходимым провести смешивание в 4 блока, по 4 наблюдения в каждом.

Первой задачей является смешивание с фактором, имеющим 4 уровня. По методу Йэйтса этот фактор может рассматриваться, как два фактора, по 2 уровня в каждом. Обозначая эти факторы через A' и A'' с индексами 1 и 2, указывающими их нижние и верхние уровни, и обозначая уровни исходного фактора с четырьмя уровнями через $a_1 - a_4$, получим требуемую разбивку в виде таблицы 14.5

ТАБЛИЦА 14.5

	A'_1	A'_2
A''_1	a_2	a_3
A''_2	a_1	a_4

При такой разбивке:

$$A' = a_4 + a_3 - a_2 - a_1,$$
$$A'' = a_4 - a_3 - a_2 + a_1.$$

Взаимодействие между A' и A'', обозначаемое через A''', будет определяться выражением:

$$A''' = a_4 - a_3 + a_2 - a_1.$$

Соответственно этому, линейной компонентой будет $2A' + A'''$, квадратичной компонентой A'', и кубичной компонентой $2A''' - A'$. При проведении смешивания допустим, что кубичная компонента незначительна, и подвергнем смешиванию A'''; это даст нам возможность оценки линейной и квадратичной компонент.

Согласно сказанному в разделе (*C*), при проведении смешивания при четырех факторах (каждый из них с двумя уровнями) одно из взаимодействий первого порядка должно быть потеряно; выберем для этой цели A''' (эквивалентное $A'A''$). Двумя другими наиболее подходящими для смешивания взаимодействиями являются $A'CD$ и $A''CD$.

Распределение способов обработки по четырем блокам производится обычным путем. Оно показано в таблице 14.6; там же приведены и цифры выпуска продукции.

ТАБЛИЦА 14.6

Блок 1		Блок 2		Блок 3		Блок 4	
(1)	24	a'	25	a''	56	c	48
cd	31	$a'cd$	23	$a''cd$	41	d	29
$a'a''d$	6	$a''d$	24	$a'd$	27	$a'a''cd$	23
$a'a''c$	14	$a''c$	36	$a'c$	40	$a'a''$	25
	75		108		164		125

Символические обозначения для способов обработки сделаны так, что, например, символ $a'a''$ указывает на то, что A' и A'' взяты на их верхних уровнях; таблица 14.5 показывает, что $A'_2A''_2$ соответствует a_4, т. е. высшему уровню фактора с четырьмя уровнями. Факторы C и D находятся, разумеется, на их нижних уровнях.

Символ $a''cd$ указывает на то, что C и D находятся на их верхних уровнях. Символ a'' обозначает $A'_1A''_2$ (т. е. что A' находится на нижнем уровне, а A'' на его верхнем уровне). По таблице 14.5 это соответствует a_1. В нижней строке таблицы 14.6 записаны итоги по блокам; сумма квадратов по блокам равна:

$$\frac{1}{4}(75^2 + 108^2 + 164^2 + 125^2) - \frac{(75 + 108 + 164 + 125)^2}{16} = 1028{,}5.$$

Эта сумма квадратов имеет 3 степени свободы. Остающаяся часть анализа проводится обычным путем.

Интересно заметить, что, если подсчитать три подвергшиеся смешиванию взаимодействия, их суммы квадратов будут равны:

$A'A''$	324,00
$A'CD$	2,25
$A''CD$	702,25

Сумма их равна подсчитанной выше величине суммы квадратов.

Для проверки правильности расчета можно найти взаимодействие наивысшего порядка $A'A''CD$ прямым путем, а не обычным способом вычитания. Раскрывая скобки, получим

$$(a'-1)(a''-1)(c-1)(d-1) = a'a''cd + a'c + a'd + a''c + $$
$$+ a''d + a'a'' + cd + 1 - a'cd - a''cd - a'a''c -$$
$$- a'a''d - a' - a'' - c - d.$$

Из таблицы результатов (таблица 14.6) получаем, что итог этого равен -12. Сумма квадратов для $A'A''CD$ равна квадрату этого результата, деленному на соответствующее число наблюдений, т. е. $\frac{(-12)^2}{16} = 9$. Эта величина совпадает с полученной путем вычитания из итога по сумме квадратов. В любом факторном испытании при 2 уровнях по каждому фактору, можно, по желанию, определить этим путем все суммы квадратов *.

Таблица дисперсионного анализа дана в табл. 14.7.

Интерпретация таблицы совершенно ясна. Остаток, представляющий собой $A'A''CD$, можно объединить с двумя взаимодействиями второго порядка $A'A''C$ и $A'A''D$; среднее квадратов будет 23,333, с тремя степенями свободы; эту величину используем, как характеристику ошибки. После этого найдем, что все взаимодействия первого порядка незначимы. Относительно основных эффектов вытекает, что D значимо при уровне в 1%, C незначимо, линейная компонента A значима при уровне $0{,}1\%$; квадратичная же компонента A незначима. Затем можно рассчитать соответствующие средние и определить их стандартные ошибки.

* Подобные методы вычисления систематически излагает F. Yates, Design and Analysis of Factorial Experiments.

ТАБЛИЦА 14.7

Источник дисперсии	Число степеней свободы	Суммы квадратов
Блоки	3	1028,5
A Линейная компонента A'	1	702,25
A Квадратичная „ A''	1	30,25
C	1	100,0
D	1	256,0
$A'C$	1	2,25
$A'D$	1	12,25
$A''C$	1	72,25
$A''D$	1	6,25
CD	1	36,0
$A'A''C$	1	36,0
$A'A''D$	1	25,0
Остаток	1	9,0
Итого	15	2316,0

Из таблицы 14.7 можно заметить, что среднее квадратов для блоков $\frac{1028}{3}=342,667$ является высоко значимым (при уровне в $0,1\%$). При применении метода смешивания оценка ошибки, полученная путем объединения всех незначимых взаимодействий первого порядка, составляет 18,78 (при 7 степенях свободы). Если бы смешивание не было применено, но взамен этого сравнивались бы случайно разбросанные 16 способов обработки, то сумма квадратов, относящаяся к различию между блоками, оказалась бы в объединенной ошибке остатка, которая была бы равна 115,575 (с десятью степенями свободы). Таким образом, благодаря применению смешивания получился очень большой выигрыш в точности, поскольку средний квадрат ошибки уменьшился до 18,178. Без этого выигрыша в точности ценность эксперимента была бы довольно малой.

Наконец, интересно посмотреть, как такого рода эксперимент был бы проведен традиционным путем. Лучшее, что можно было бы сделать, — включить фактор A с четырьмя уровнями в два блока, при C и D, поддерживаемых постоянными. Третий блок содержал бы C с двумя повторениями (два наблюдения по каждому из двух уровней), при постоянных A и D. Четвертый блок был бы предназначен для аналогичного испытания D. Такой эксперимент, по сравнению с факторным испытанием по методу смешивания, дал бы только половину точности в оценке A и одну четвертую часть точности в оценке C и D. Кроме того, не получилось бы вообще оценок возможного существования взаимодействий. Таким образом, проведение факторного испытания по методу смешивания дает значительный выигрыш.

(G) ЧАСТИЧНОЕ СМЕШИВАНИЕ

Ясно, что если нельзя допустить потери ни одного из взаимодействий первого порядка, а метод смешивания, к которому приходится прибегать из-за ограниченности объема блоков, требует потери одного из таких взаимодействий, то следует сделать некоторые повторения, выбранные по желанию так, чтобы подвергнуть смешиванию другие взаимодействия, а не те, которые оказались потерянными в первой серии испытаний. Такого рода метод проведения испытаний называется „частичным смешиванием“. Подробности такого метода можно найти в работах Йэйтса (Yates, Design and Analysis of Factorial Experiments. Sections 4 and 5) и Гулдена (Goulden, Methods of Statistical Analysis).

(H) ДРУГИЕ ВОЗМОЖНОСТИ МЕТОДА СМЕШИВАНИЯ

В указанной в предыдущем разделе работе Йэйтс указал на возможность применения смешивания в случаях, когда имеют дело с факторами, имеющими более двух уровней. Среди описанных случаев им приведены примеры смешивания для экспериментов $3\times3\times3$ и $3\times3\times3\times3$ с разбивкой соответственно на 3 или 9 блоков, объемом по 9 испытаний в каждом. Там же описаны смешанные типы $3\times2\times2$, $3\times2\times2\times2$, $3\times3\times2$ и $3\times3\times3\times2$ с разбивкой на блоки по 6 испытаний в каждом.

(I) ДВОИНОЕ СМЕШИВАНИЕ *

В факторном испытании на систему иногда накладывается двойное ограничение. Например, в испытаниях типа 2^5 (5 факторов, при двух уровнях каждый, что требует всего 32 отдельных испытания) может оказаться 4 реактива, которые, в свою очередь, могут быть различными. Соответственно, производится смешивание в 4 блока по 8 испытаний в каждом; на каждый из четырех реактивов придется по 8 испытаний. Однако, объем партий сырого материала может быть недостаточно велик для проведения всех испытаний. В таком случае требуется провести смешивание некоторым особым путем. Задача разрешима, если мы сделаем второе смешивание обратным по сравнению с первым типом смешивания. Так, если первое смешивание делалось в 4 блока, по 8 испытаний в каждом, то вторым должно быть смешивание в 8 блоков, по 4 испытания в каждом. Существует еще дополни-

* См. R. A. Fisher, The Design of Experiments, 3-е издание, стр. 116.

тельное ограничение в том, чтобы ни одно из взаимодействий, подвергшееся смешиванию во втором комплекте испытаний, не встречалось в первом, и наоборот.

Пусть при смешивании в 8 блоков по 4 испытания в каждом, мы избрали в качестве теряемых взаимодействия

$$PQ,\ RS,\ PRT,\ QST,\ QRT,\ PST,\ PQRS,$$

а при смешивании в 4 блока по 8 испытаний —

$$PQR,\quad RST,\quad PQST.$$

Ни одно из взаимодействий, вошедших во второй комплект, не встречается в первом.

Для проведения разбивки испытаний, сперва выпишем обычную таблицу разбивки в 8 блоков по 4 испытания в каждом (таблица 14.8).

ТАБЛИЦА 14.8

Партия 1	Партия 2	Партия 3	Партия 4	Партия 5	Партия 6	Партия 7	Партия 8
(1)	*p*	*q*	*r*	*s*	*t*	*pr*	*qr*
pqrs	*qrs*	*prs*	*pqs*	*pqr*	*pqrst*	*qs*	*ps*
rst	*prst*	*qrst*	*st*	*rt*	*rs*	*pst*	*qst*
pqt	*qt*	*pt*	*pqrt*	*pqst*	*pq*	*rqt*	*prt*

Делаем также обычную разбивку на 4 блока по 8 испытаний в каждом, как показано в таблице 14.9.

ТАБЛИЦА 14.9

Реактив 1	(1)	*pq*	*st*	*pqst*	*qrt*	*prt*	*prs*	*qrs*
Реактив 2	*p*	*q*	*pst*	*qst*	*pqrt*	*rt*	*rs*	*pqrs*
Реактив 3	*s*	*pqs*	*t*	*pqt*	*qrst*	*prst*	*pr*	*qr*
Реактив 4	*r*	*pqr*	*rst*	*pqrst*	*qt*	*pt*	*ps*	*qs*

Прикладываем таблицу 14.8 к таблице 14.9 сверху. Испытание (1) относится к партии 1 и реактиву 1; вносим его в соответствующую клетку таблицы 14.10. Испытание *pqrs* относится к партии I и реактиву 2; вносим его в соответствующую клетку таблицы 14.10. Испытание *rst* относится к партии 1 и реактиву 4 и т. д.

ТАБЛИЦА 14.10

	Партия 1	Партия 2	Партия 3	Партия 4	Партия 5	Партия 6	Партия 7	Партия 8
Реактив 1	(1)	*qrs*	*prs*	*st*	*pqst*	*pq*	*qrt*	*prt*
Реактив 2	*pqrs*	*p*	*q*	*pqrt*	*rt*	*rs*	*pst*	*qst*
Реактив 3	*pqt*	*prst*	*qrst*	*pqs*	*s*	*t*	*pr*	*qr*
Реактив 4	*rst*	*qt*	*pt*	*r*	*pqr*	*pqrst*	*qs*	*ps*

Таблица 14.10 соответствует условиям обеих таблиц 14.8 и 14.9. Поэтому, приведенная в таблице 14.10 разбивка испытаний является правильной. Далее, проводится обычный дисперсионный анализ, но подвергшиеся смешиванию взаимодействия не вписываются в таблицу дисперсионного анализа; их суммы квадратов и числа степеней свободы объединяются под рубрикой „блоки".

Заметим, что, при смешивании в 8 блоков, по 4 испытания в каждом, наложение второго ограничения не увеличивает числа теряемых взаимодействий первого порядка.

Двойное смешивание не особенно практично для испытаний типа 2^4; оно дает хорошие результаты в случае испытаний типа 2^5, описанных выше, а также в случае испытаний 2^6, где имеются две возможности смешивания: либо в 8 блоков, по 8 испытаний в каждом (одна комбинация), и в 8 блоков, по 8 испытаний (другая комбинация), либо в 16 блоков, по 4 испытания, и в 4 блока, по 16 испытаний.

В первом варианте одним из комплектов взаимодействий, теряемых при смешивании, являются:

$$PQRS,\ RSTU,\ PQTU,\ PRT,\ QST,\ PSU \text{ и } QRU;$$

вторым комплектом, не перекрывающим первый, являются:

$$PQRU,\ PRST,\ QSTU,\ PQS,\ RSU,\ RQT \text{ и } PTU.$$

Таким образом, двойное ограничение может быть наложено на систему вообще без потери каких-либо взаимодействий первого порядка.

При двойном смешивании в 16 блоков по 4 испытания и в 4 блока по 16 испытаний, комплектами взаимодействий, теряемых при смешивании, являются: первый комплект:

$$PQ,\ RU,\ PRT,\ QST,\ QRT,\ PST,\ PQRS,\ TU,$$
$$PQTU,\ RSTU,\ PRU,\ QSU,\ QRU,\ PSU,\ PQRSTU;$$

второй комплект (не перекрывающий первый):

$$PQRU,\ PRST,\ QSTU.$$

Таким образом, при смешивании в 16 блоков, по 4 испытания в каждом, теряются 3 взаимодействия первого порядка; наложение же второго ограничения не увеличивает числа этих потерянных взаимодействий первого порядка; поэтому, наложение второго ограничения можно проводить во всех случаях, когда это удобно или необходимо.

ГЛАВА XV

ОБЩИЕ ВЫВОДЫ

(А) ИССЛЕДОВАНИЕ СИСТЕМ С МНОГИМИ ПЕРЕМЕННЫМИ

Мы рассмотрели два различных общих метода исследования систем с многими переменными: множественная корреляция и факторные испытания.

Множественная корреляция удобна в случаях, когда имеется значительное количество данных об исследуемой системе, причем различные независимые переменные имеют случайный характер и находятся на всех уровнях. Такого рода данные накапливаются в процессе нормальной эксплоатации предприятия; в частности, эти данные очень легко собрать на современном предприятии, оборудованном автоматически-записывающими приборами, содержащимися в надлежащем рабочем порядке.

Факторные испытания удобны в случаях, когда производятся специальные исследования системы, и независимые переменные контролируются на уровнях, запроектированных по плану проведения факторных испытаний. Проведение таких испытаний требует вообще тщательного надзора и, поэтому, имеет тенденцию быть сравнительно дорогостоющим.

Метод множественной корреляции имеет недостаток, заключающийся в допущении, что зависимая переменная является линейной функцией независимых переменных; предполагается также, что последние действуют независимо друг от друга.

Однако, некоторые трудности, возникающие из-за нелинейности, можно обойти путем применения различных преобразований, указанных в главе IX, (В), и главе X, (В). Заметим, что удовлетворительность применения этих приемов обусловливается выбором соответствующего преобразования, основанном либо на знании физико-химических свойств и процессов системы, либо на разумном или удачном рассмотрении данных.

Если же пробовать сперва одно преобразование, потом другое, третье и т. д., то может потребоваться слишком много труда.

Затруднения, вызываемые зависимостью между аргументами, можно отчасти обойти путем добавления новых аргументов в виде произведений исходных. Например, если y — зависимая переменная, а x_1 и x_2 — аргументы, то можно включить третью переменную x_1x_2, т. е. подобрать уравнение такого вида:

$$y = b_1x_1 + b_2x_2 + b_{12}x_1x_2.$$

Однако, если имеется три независимых переменных и нужно проверить возможные взаимодействия между ними, подбираемое уравнение будет типа:

$$y = b_1x_1 + b_2x_2 + b_3x_3 + b_{12}x_1x_2 + b_{23}x_2x_3 + b_{31}x_3x_1.$$

Подсчет соответствующих сумм квадратов и сумм произведений и последующее решение системы шести уравнений потребуют очень много работы; это может быть оправдано только в весьма важных случаях.

При факторном же испытании начальные допущения носят весьма общий характер. Никаких предположений о линейной зависимости не делается, и любое взаимодействие первого порядка может оказаться значимым. Факторное испытание может дать запутанные результаты лишь в очень редких случаях, когда каждый фактор сильно взаимодействует со всеми другими; но даже и при таких условиях результаты испытания могут, возможно, быть сделаны более ясными путем логарифмического преобразования. Кроме того, если некоторые из факторов имеют более двух уровней, возможная нелинейность в соотношениях эффектов может быть проверена помощью методов, указанных в главе XII (С). Требуемое количество расчетов гораздо меньше, чем при множественной корреляции.

Дальнейшим преимуществом факторных испытаний является то, что в общем случае точность результатов может быть заметно повышена применением метода смешивания.

В целом, не может быть сомнения, что факторные испытания должны быть гораздо предпочтительнее, по сравнению с исследованием данных с помощью метода множественной корреляции. Кроме изложенных преимуществ, достоинством факторных испытаний является еще и то обстоятельство, что они не подвержены погрешности из-за наличия неучтенных коррелированных переменных; в этом отличие эксперимента от наблюдения, рассмотренное в главе IX (К).

В промежуточных случаях, когда некоторые переменные могут контролироваться на уровнях, запроектированных планом проведения факторных испытаний, но другие переменные труднее поддаются контролю, наиболее подходящим будет метод смешанных вторых моментов. По применениям этого метода к производственным задачам* пока еще нет опубликованных работ.

(B) ПРЕИМУЩЕСТВА СТАТИСТИЧЕСКИ-ПЛАНИРОВАННЫХ ИСПЫТАНИЙ

Очевидно, что ни один из статистических приемов исследования нельзя применять, предварительно не продумав и не спланировав испытания. Статистические приемы нельзя применить к сумбурным случайным результатам.

Это ведет к далеко идущим заключениям. Например, традиционный стиль экспериментирования, когда пытаются сперва сделать что-то одно, затем, в случае неудачи, что-то другое и т. д., т. е. фактически придерживаются политики продвижения „малыми шагами", — совершенно неудовлетворителен. Наоборот, мы требуем сперва обдумать и оценить все обстоятельства, кажущиеся интересными или могущие представить интерес в будущем, и затем экспериментировать по ним всем одновременно. Таким путем мы будем двигаться прыжками, а не шагами. Большим преимуществом движения прыжками является то, что информация, полученная при прыжке, достигает, без дополнительных усилий с нашей стороны, гораздо большего объема, чем полученная при движении шагами.

В качестве примера обратимся к случаю, рассмотренному в главе XIII, (B), где испытывались 7 способов обработки в блоках, объемом 3. Дисперсия сравнения любой пары корректированных средних была $\frac{6}{7}\sigma_k^2 = 0{,}857\,\sigma_k^2$, где σ_k^2 была дисперсией ошибки одиночного наблюдения в пределах блоков.

Предположим теперь, что, вместо того чтобы начать экспериментирование с полным числом способов обработки, мы начали работать только с 4, а затем решили испытать прочие 3. Первые 4 способа обработки можно было бы испытать с помощью приема, данного в таблице 13.3; дисперсия сравнения любой пары этих средних была бы

$$2 \cdot \frac{k(t-1)}{N(k-1)}\sigma_k^2 = 2 \cdot \frac{3(4-1)}{3 \cdot 4(3-1)}\sigma_k^2 = 0{,}75\,\sigma_k^2.$$

* Описание метода смешанных вторых моментов см. R. A. Fisher, Statistical Methods for Research Workers. Section 49.1; K. Mather, Statistical Analysis in Biology, Section 34; G. W. Snedecor, Statistical Methods, ch. 12.

Эта цифра значительно лучше по сравнению с $0{,}857\,\sigma_k^2$ для всего эксперимента. Однако, когда мы подойдем к трем дополнительным способам обработки, окажется невозможным провести их исследование так, чтобы, по добавлении результатов к наличным данным, удалось получить ту же схему исследования, которая применялась для испытания 7 способов обработки. Лучшим вариантом того, что можно сделать, будет избрать один из способов обработки (например T_4) в качестве „стандарта", входящего в обе части эксперимента; из этого „стандарта" и трех оставшихся способов обработки образуется система сбалансированных неполных блоков, объема 3, что требует трех повторений.

Дисперсия ошибки сравнения любой пары из четырех способов обработки $T_4, \ldots, T_7$ будет $0{,}75\,\sigma_k^2$, т. е. той же величины, что и для любой пары из четырех способов обработки $T_1, \ldots, T_4$. Однако, когда потребуется сравнить любой из трех способов обработки группы $T_1, \ldots, T_3$ с любым из группы $T_5, \ldots, T_7$, сравнение можно выполнить только через промежуточную величину T_4; дисперсия ошибки такого сравнения будет $2 \cdot 0{,}75\,\sigma_k^2 = 1{,}5\,\sigma_k^2$, почти вдвое больше, чем при проведении всех испытаний сразу $(0{,}857\,\sigma_k^2)$. Общая средняя дисперсия сравнения в двухступенном эксперименте будет, в действительности, достигать $1{,}071\,\sigma_k^2$.

Принимая во внимание, что двухступенный эксперимент потребовал бы 24 наблюдения, а одноступенный (проведение всех испытаний, сразу) — только 21, становится очевидным явное преимущество последнего.

Наконец, можно было бы еще провести сравнение рассматривавшегося эксперимента с классическим методом проведения испытаний, с использованием повсюду единого „стандарта". Наиболее близкое приближение к сбалансированному плану проведения испытаний потребует сформировать 6 блоков, каждый из которых содержит стандарт и два способа обработки из остальных 6. Каждый из последних будет, таким образом, использован дважды.

Общее число требуемых наблюдений $6 \times 3 = 18$, т. е. несколько меньше, чем 21, требовавшиеся по методу сбалансированных неполных блоков. Точность же будет намного меньше.

При классическом приеме проведения испытаний будет три отдельных типа сравнений:

(a) Между каждым способом обработки и стандартом. Дисперсия их будет $0{,}667\,\sigma_k^2$.

(b) Между любой парой способов обработки, входящих в один и тот же блок. Их дисперсия будет $1{,}333\,\sigma_k^2$.

(c) Между любой парой способов обработки, не входящих в один и тот же блок. Их дисперсия будет равна приблизительно $4\,\sigma_k^2$. Общая средняя дисперсия будет равна приблизительно $2{,}285\,\sigma_k^2$.

Поскольку при применении метода сбалансированных неполных блоков все сравнения имели дисперсию 0,857 σ_k^2, сравнение в целом метода сбалансированных неполных блоков с классическим методом оказывается весьма неблагоприятным для последнего.

Таким же путем можно показать, что одно большое факторное испытание гораздо удовлетворительнее, чем два меньших по объему. Например, в случае 6 факторов, каждый на 2 уровнях, лучше провести одно полное испытание, требующее $2^6 = 64$ наблюдения, * чем разделить испытания на две части, с тремя факторами в каждой, что потребовало бы для обеих частей всего $2 \times 2^3 = 16$ наблюдений. Полное испытание потребовало бы гораздо больше работы, но принесло бы взамен следующие преимущества:

(a) Основные эффекты определятся с точностью в 8 раз выше.

(b) Все 15 возможных взаимодействий первого порядка оценятся со значительной точностью; испытания же, проведенные по частям, дадут оценку только 6 взаимодействий из 15, причем с точностью в 8 раз ниже.

(c) Получатся оценки всех 20 возможных взаимодействий второго порядка; при проведении же испытаний по частям только 2 таких взаимодействия получили бы оценку, причем они должны быть использованы как ошибки, так что полученные величины не имели бы никакой ценности как оценки самих взаимодействий.

(d) Оценятся все 15 возможных взаимодействий третьего порядка. Останутся еще 7 степеней свободы для величины ошибки; используя вместе с ними большую часть чисел свободы, относящихся к незначимым взаимодействиям, получим число степеней свободы для величины ошибки около 20. Эти характеристики гораздо удовлетворительнее, чем результаты проведения испытаний по частям, где мы начинаем работу с взаимодействием второго порядка в качестве ошибки, и только с 1 степенью свободы

(C) ЗАКЛЮЧЕНИЕ

Разделы о планировании испытаний написаны так, как будто бы эта тема была уже полностью разработана, подобно эвклидовой геометрии. Дело, однако, обстоит не так: эти статистические приемы имели до сих пор только очень ограниченное применение в химической промышленности; развивались же они в сельскохозяйственных науках, где теперь мы располагаем опытом, охватывающим более чем двадцатилетний период. Методы, очевидно, пригодны для переноса в другие отрасли по указанным путям; но мы должны быть осторожны в отношении чрезмерно буквального их применения. В химической промышленности требуется еще разработать целый ряд методов; другие, возможно, подлежат модификации с целью приспособления их к новым условиям. Так, например, повидимому для окончательных выводов, на основании которых могут быть проведены известные мероприятия, уровень значимости в 5% недостаточен; в действительности, требуется уровень в 1% или еще лучший. Поэтому в интерпретации выводов следует проявлять осторожность, пока мы не располагаем бо́льшим запасом данных опыта.

* В последнее время были разработаны методы факторных испытаний с большими количествами факторов, которые позволяют получить удовлетворительные оценки основных эффектов и взаимодействий низших порядков (первого и второго) с помощью только части полного числа наблюдений. См. D. J. Finney, *Annals of Eugenics*, **12**, (1945) 291.

ПРИЛОЖЕНИЯ

ТАБЛИЦА I

ТАБЛИЦА ЗНАЧЕНИИ t

Числа степеней свободы	t				
	0,10	0,05	0,02	0,01	0,001
1	6,31	12,71	31,82	63,66	636,62
2	2,92	4,30	6,97	9,93	31,60
3	2,35	3,18	4,54	5,84	12,94
4	2,13	2,78	3,75	4,60	8,61
5	2,02	2,57	3,37	4,03	6,86
6	1,94	2,45	3,14	3,71	5,96
7	1,90	2,37	3,00	3,50	5,41
8	1,86	2,31	2,90	3,36	5,04
9	1,83	2,26	2,82	3,25	4,78
10	1,81	2,23	2,76	3,17	4,59
11	1,80	2,20	2,72	3,11	4,44
12	1,78	2,18	2,68	3,06	4,32
13	1,77	2,16	2,65	3,01	4,22
14	1,76	2,15	2,62	2,98	4,14
15	1,75	2,13	2,60	2,95	4,07
16	1,75	2,12	2,58	2,92	4,02
17	1,74	2,11	2,57	2,90	3,97
18	1,73	2,10	2,55	2,88	3,92
19	1,73	2,09	2,54	2,86	3,88
20	1,73	2,09	2,53	2,85	3,85
21	1,72	2,08	2,52	2,83	3,82
22	1,72	2,07	2,51	2,82	3,79
23	1,71	2,07	2,50	2,81	3,77
24	1,71	2,06	2,49	2,80	3,75
25	1,71	2,06	2,48	2,79	3,73
26	1,71	2,06	2,48	2,78	3,71
27	1,70	2,05	2,47	2,77	3,69
28	1,70	2,05	2,47	2,76	3,67
29	1,70	2,04	2,46	2,76	3,66
30	1,70	2,04	2,46	2,75	3,65
40	1,68	2,02	2,42	2,70	3,55
60	1,67	2,00	2,39	2,66	3,46
120	1,66	1,98	2,36	2,62	3,37
∞	1,65	1,96	2,33	2,58	3,29

ТАБЛИЦА II

ТАБЛИЦА ЗНАЧЕНИЙ χ^2

Число степеней свободы	χ^2									
	0,99	0,98	0,95	0,90	0,50	0,10	0,05	0,02	0,01	0,001
1	0,000	0,001	0,004	0,015	0,455	2,71	3,84	5,41	6,64	10,83
2	0,020	0,040	0,103	0,211	1,386	4,61	5,99	7,82	9,21	13,82
3	0,115	0,185	0,352	0,584	2,366	6,25	7,82	9,84	11,34	16,27
4	0,297	0,429	0,711	1,064	3,357	7,78	9,49	11,67	13,28	18,47
5	0,554	0,752	1,145	1,610	4,351	9,24	11,07	13,39	15,09	20,52
6	0,872	1,134	1,635	2,204	5,35	10,65	12,59	15,03	16,81	22,46
7	1,239	1,564	2,167	2,833	6,35	12,02	14,07	16,62	18,48	24,32
8	1,646	2,032	2,733	3,490	7,34	13,36	15,51	18,17	20,09	26,13
9	2,088	2,532	3,325	4,168	8,34	14,68	16,92	19,68	21,67	27,88
10	2,558	3,059	3,940	4,865	9,34	15,99	18,31	21,16	23,21	29,59
11	3,05	3,61	4,57	5,58	10,34	17,28	19,68	22,62	24,73	31,26
12	3,57	4,18	5,23	6,30	11,34	18,55	21,03	24,05	26,22	32,91
13	4,11	4,76	5,89	7,04	12,34	19,81	22,36	25,47	27,69	34,53
14	4,66	5,37	6,57	7,79	13,34	21,06	23,69	26,87	29,14	36,12
15	5,23	5,99	7,26	8,55	14,34	22,31	25,00	28,26	30,58	37,70
16	5,81	6,61	7,96	9,31	15,34	23,54	26,30	29,63	32,00	39,25
17	6,41	7,26	8,67	10,09	16,34	24,77	27,59	31,00	33,41	40,79
18	7,02	7,91	9,39	10,87	17,34	25,99	28,87	32,35	34,81	42,31
19	7,63	8,57	10,12	11,65	18,34	27,20	30,14	33,69	36,19	43,82
20	8,26	9,24	10,85	12,44	19,34	28,41	31,41	35,02	37,57	45,32
21	8,90	9,91	11,59	13,34	20,34	29,61	32,67	36,34	38,93	46,80
22	9,54	10,60	12,34	14,04	21,34	30,81	33,92	37,66	40,29	48,27
23	10,20	11,29	13,09	14,85	22,34	32,01	35,17	38,97	41,64	49,73
24	10,86	11,99	13,85	15,66	23,34	33,20	36,42	40,27	42,98	51,18
25	11,52	12,70	14,61	16,47	24,34	34,38	37,65	41,57	44,31	52,62
26	12,20	13,41	15,38	17,29	25,34	35,56	38,89	42,86	45,64	54,05
27	12,88	14,12	16,15	18,11	26,34	36,74	40,11	44,14	46,96	55,48
28	13,56	14,85	16,93	18,94	27,34	37,92	41,34	45,42	48,28	56,89
29	14,26	15,57	17,71	19,77	28,34	39,09	42,56	46,69	49,59	58,30
30	14,95	16,31	18,49	20,60	29,34	40,26	43,77	47,96	50,89	59,70

ТАБЛИЦА III

ТАБЛИЦА ОТНОШЕНИЙ ДИСПЕРСИЙ (I)

N_2 \ N_1	Уровень значимости 0,20								
	1	2	3	4	5	6	12	24	∞
1	9,5	12,0	13,1	13,7	14,0	14,3	14,9	15,2	15,6
2	3,6	4,0	4,2	4,2	4,3	4,3	4,4	4,4	4,5
3	2,7	2,9	2,9	3,0	3,0	3,0	3,0	3,0	3,0
4	2,4	2,5	2,5	2,5	2,5	2,5	2,5	2,4	2,4
5	2,2	2,3	2,3	2,2	2,2	2,2	2,2	2,2	2,1
6	2,1	2,1	2,1	2,1	2,1	2,1	2,0	2,0	2,0
7	2,0	2,0	2,0	2,0	2,0	2,0	1,9	1,9	1,8
8	2,0	2,0	2,0	1,9	1,9	1,9	1,8	1,8	1,7
9	1,9	1,9	1,9	1,9	1,9	1,8	1,8	1,7	1,7
10	1,9	1,9	1,9	1,8	1,8	1,8	1,7	1,7	1,6
11	1,9	1,9	1,8	1,8	1,8	1,8	1,7	1,6	1,6
12	1,8	1,8	1,8	1,8	1,7	1,7	1,7	1,6	1,5
13	1,8	1,8	1,8	1,8	1,7	1,7	1,6	1,6	1,5
14	1,8	1,8	1,8	1,7	1,7	1,7	1,6	1,6	1,5
15	1,8	1,8	1,8	1,7	1,7	1,7	1,6	1,5	1,5
16	1,8	1,8	1,7	1,7	1,7	1,6	1,6	1,5	1,4
17	1,8	1,8	1,7	1,7	1,7	1,6	1,6	1,5	1,4
18	1,8	1,8	1,7	1,7	1,6	1,6	1,5	1,5	1,4
19	1,8	1,8	1,7	1,7	1,6	1,6	1,5	1,5	1,4
20	1,8	1,8	1,7	1,7	1,6	1,6	1,5	1,5	1,4
22	1,8	1,7	1,7	1,6	1,6	1,6	1,5	1,4	1,4
24	1,7	1,7	1,7	1,6	1,6	1,6	1,5	1,4	1,3
26	1,7	1,7	1,7	1,6	1,6	1,6	1,5	1,4	1,3
28	1,7	1,7	1,7	1,6	1,6	1,6	1,5	1,4	1,3
30	1,7	1,7	1,6	1,6	1,6	1,5	1,5	1,4	1,3
40	1,7	1,7	1,6	1,6	1,5	1,5	1,4	1,4	1,2
60	1,7	1,7	1,6	1,6	1,5	1,5	1,4	1,3	1,2
120	1,7	1,6	1,6	1,5	1,5	1,5	1,4	1,3	1,1
∞	1,6	1,6	1,6	1,5	1,5	1,4	1,3	1,2	1,0

ТАБЛИЦА ОТНОШЕНИЙ ДИСПЕРСИЙ (II)

N_2 \ N_1	Уровень значимости 0,05								
	1	2	3	4	5	6	12	24	∞
1	164,4	199,5	215,7	224,6	230,2	234,0	244,9	249,0	254,3
2	18,5	19,2	19,2	19,3	19,3	19,3	19,4	19,5	19,5
3	10,1	9,6	9,3	9,1	9,0	8,9	8,7	8,6	8,5
4	7,7	6,9	6,6	6,4	6,3	6,2	5,9	5,8	5,6
5	6,6	5,8	5,4	5,2	5,1	5,0	4,7	4,5	4,4
6	6,0	5,1	4,8	4,5	4,4	4,3	4,0	3,8	3,7
7	5,6	4,7	4,4	4,1	4,0	3,9	3,6	3,4	3,2
8	5,3	4,5	4,1	3,8	3,7	3,6	3,3	3,1	2,9
9	5,1	4,3	3,9	3,6	3,5	3,4	3,1	2,9	2,7
10	5,0	4,1	3,7	3,5	3,3	3,2	2,9	2,7	2,5
11	4,8	4,0	3,6	3,4	3,2	3,1	2,8	2,6	2,4
12	4,8	3,9	3,5	3,3	3,1	3,0	2,7	2,5	2,3
13	4,7	3,8	3,4	3,2	3,0	2,9	2,6	2,4	2,2
14	4,6	3,7	3,3	3,1	3,0	2,9	2,5	2,3	2,1
15	4,5	3,7	3,3	3,1	2,9	2,8	2,5	2,3	2,1
16	4,5	3,6	3,2	3,0	2,9	2,7	2,4	2,2	2,0
17	4,5	3,6	3,2	3,0	2,8	2,7	2,4	2,2	2,0
18	4,4	3,6	3,2	2,9	2,8	2,7	2,3	2,1	1,9
19	4,4	3,5	3,1	2,9	2,7	2,6	2,3	2,1	1,9
20	4,4	3,5	3,1	2,9	2,7	2,6	2,3	2,1	1,8
22	4,3	3,4	3,1	2,8	2,7	2,6	2,2	2,0	1,8
24	4,3	3,4	3,0	2,8	2,6	2,5	2,2	2,0	1,7
26	4,2	3,4	3,0	2,7	2,6	2,5	2,2	2,0	1,7
28	4,2	3,3	3,0	2,7	2,6	2,4	2,1	1,9	1,7
30	4,2	3,3	2,9	2,7	2,5	2,4	2,1	1,9	1,6
40	4,1	3,2	2,9	2,6	2,5	2,3	2,0	1,8	1,5
60	4,0	3,2	2,8	2,5	2,4	2,3	1,9	1,7	1,4
120	3,9	3,1	2,7	2,5	2,3	2,2	1,8	1,6	1,3
∞	3,8	3,0	2,6	2,4	2,2	2,1	1,8	1,5	1,0

14 Зак. 1691

ТАБЛИЦА ОТНОШЕНИЙ ДИСПЕРСИЙ (III)

N_2 \ N_1	Уровень значимости 0,01									
	1	2	3	4	5	6	8	12	24	∞
1	4052	4999	5403	5625	5764	5859	5981	6106	6234	6366
2	98,5	99,0	99,2	99,3	99,3	99,4	99,3	99,4	99,5	99,5
3	34,1	30,8	29,5	28,7	28,2	27,9	27,5	27,1	26,6	26,1
4	21,2	18,0	16,7	16,0	15,5	15,2	14,8	14,4	13,9	13,5
5	16,3	13,3	12,1	11,4	11,0	10 7	10,3	9,9	9,5	9,0
6	13,7	10,9	9,8	9,2	8,8	8,5	8,1	7,7	7,3	6,9
7	12,3	9,6	8,5	7,9	7,5	7,2	6,8	6,5	6,1	5,7
8	11,3	8,7	7,6	7,0	6,6	6,4	6,0	5,7	5,3	4,9
9	10,6	8,0	7,0	6,4	6,1	5,8	5,5	5,1	4,7	4,3
10	10,0	7,6	6,6	6,0	5,6	5,4	5,1	4,7	4,3	3,9
11	9,7	7,2	6,2	5,7	5,3	5,1	4,7	4,4	4,0	3,6
12	9,3	6,9	6,0	5,4	5,1	4,8	4,5	4,2	3,8	3,4
13	9,1	6,7	5,7	5,2	4,9	4,6	4,3	4,0	3,6	3,2
14	8,9	6,5	5,6	5,0	4,7	4,5	4,1	3,8	3,4	3,0
15	8,7	6,4	5,4	4,9	4,6	4,3	4,0	3,7	3,3	2,9
16	8,5	6,2	5,3	4,8	4,4	4,2	3,9	3,6	3,2	2,8
17	8,4	6,1	5,2	4,7	4,3	4,1	3,8	3,5	3,1	2,7
18	8,3	6,0	5,1	4,6	4,3	4,0	3,7	3,4	3,0	2,6
19	8,2	5,9	5,0	4,5	4,2	3,9	3,6	3,3	2,9	2,5
20	8,1	5,9	4,9	4,4	4,1	3,9	3,6	3,2	2,9	2,4
22	7,9	5,7	4,8	4,3	4,0	3,8	3,5	3,1	2,8	2,3
24	7,8	5,6	4,7	4,2	3,9	3,7	3,3	3,0	2,7	2,2
26	7,7	5,5	4,6	4,1	3,8	3,6	3,3	3,0	2,6	2,1
28	7,6	5,5	4,6	4,1	3,8	3,5	3,2	2,9	2,5	2,1
30	7,6	5,4	4,5	4,0	3,7	3,5	3,2	2,8	2,5	2,0
40	7,3	5,2	4,3	3,8	3,5	3,3	3,0	2,7	2,3	1,8
60	7,1	5,0	4,1	3,7	3,3	3,1	2,8	2,5	2,1	1,6
120	6,9	4,8	4,0	3,5	3,2	3,0	2,7	2,3	2,0	1,4
∞	6,6	4,6	3,8	3,3	3,0	2,8	2,5	2,2	1,8	1,0

ТАБЛИЦА ОТНОШЕНИЙ ДИСПЕРСИЙ (IV)

N_2 \ N_1	Уровень значимости 0,001									
	1	2	3	4	5	6	8	12	24	∞
1	Изменяется от 400 000 до 600 000									
2	998	999	999	999	999	999	999	999	999	999
3	167	148	141	137	135	133	131	128	126	123
4	74,1	61,3	56,2	53,4	51,7	50,5	49,0	47,4	45,8	44,1
5	47,0	36,6	33,2	31,1	29,8	28,8	27,6	26,4	25,1	23,8
6	35,5	27,0	23,7	21,9	20,8	20,0	19,0	18,0	16,9	15,8
7	29,2	21,7	18,8	17,2	16,2	15,5	14,6	13,7	12,7	11,7
8	25,4	18,5	15,8	14,4	13,5	12,9	12,0	11,2	10,3	9,3
9	22,9	16,4	13,9	12,6	11,7	11,1	10,4	9,6	8,7	7,8
10	21,0	14,9	12,6	11,3	10,5	9,9	9,2	8,5	7,6	6,8
11	19,7	13,8	11,6	10,4	9,6	9,1	8,3	7,6	6,9	6,0
12	18,6	13,0	10,8	9,6	8,9	8,4	7,7	7,0	6,3	5,4
13	17,8	12,3	10,2	9,1	8,4	7,9	7,2	6,5	5,8	5,0
14	17,1	11,8	9,7	8,6	7,9	7,4	6,8	6,1	5,4	4,6
15	16,6	11,3	9,3	8,3	7,6	7,1	6,5	5,8	5,1	4,3
16	16,1	11,0	9,0	7,9	7,3	6,8	6,2	5,6	4,9	4,1
17	15,7	10,7	8,7	7,7	7,0	6,6	6,0	5,3	4,6	3,9
18	15,4	10,4	8,5	7,5	6,8	6,4	5,8	5,1	4,5	3,7
19	15,1	10,2	8,3	7,3	6,6	6,2	5,6	5,0	4,3	3,5
20	14,8	10,0	8,1	7,1	6,5	6,0	5,4	4,8	4,2	3,4
22	14,4	9,6	7,8	6,8	6,2	5,8	5,2	4,6	3,9	3,2
24	14,0	9,3	7,6	6,6	6,0	5,6	5,0	4,4	3,7	3,0
26	13,7	9,1	7,4	6,4	5,8	5,4	4,8	4,2	3,6	2,8
28	13,5	8,9	7,2	6,3	5,7	5,2	4,7	4,1	3,5	2,7
30	13,3	8,8	7,1	6,1	5,5	5,1	4,6	4,0	3,4	2,6
40	12,6	8,2	6,6	5,7	5,1	4,7	4,2	3,6	3,0	2,2
60	12,0	7,8	6,2	5,3	4,8	4,4	3,9	3,3	2,7	1,9
120	11,4	7,3	5,8	5,0	4,4	4,0	3,5	3,0	2,4	1,6
∞	10,8	6,9	5,4	4,6	4,1	3,7	3,3	2,7	2,1	1,0

ТАБЛИЦА IV

ТАБЛИЦА КОЭФФИЦИЕНТОВ КОРРЕЛЯЦИИ

Число степеней свободы					
	0,10	0,05	0,02	0,01	0,001
1	0,988	0,997	0,999	1,000	1,000
2	0,900	0,950	0,980	0,990	0,999
3	0,805	0,878	0,934	0,959	0,992
4	0,729	0,811	0,882	0,917	0,974
5	0,669	0,754	0,833	0,874	0,951
6	0,621	0,707	0,789	0,834	0,925
7	0,582	0,666	0,750	0,798	0,898
8	0,549	0,632	0,716	0,765	0,872
9	0,521	0,602	0,685	0,735	0,847
10	0,497	0,576	0,658	0,708	0,823
11	0,476	0,553	0,634	0,684	0,801
12	0,457	0,532	0,612	0,661	0,780
13	0,441	0,514	0,592	0,641	0,760
14	0,426	0,497	0,574	0,623	0,742
15	0,412	0,482	0,558	0,606	0,725
16	0,400	0,468	0,543	0,590	0,708
17	0,389	0,456	0,528	0,575	0,693
18	0,378	0,444	0,516	0,561	0,679
19	0,369	0,433	0,503	0,549	0,665
20	0,360	0,423	0,492	0,537	0,652
25	0,323	0,381	0,445	0,487	0,597
30	0,296	0,349	0,409	0,449	0,554
35	0,275	0,325	0,381	0,418	0,519
40	0,257	0,304	0,358	0,393	0,490
45	0,243	0,287	0,338	0,372	0,465
50	0,231	0,273	0,322	0,354	0,443
60	0,211	0,250	0,295	0,325	0,408
70	0,195	0,232	0,274	0,302	0,380
80	0,183	0,217	0,256	0,283	0,357
90	0,173	0,205	0,242	0,267	0,337
100	0,164	0,195	0,230	0,254	0,321

ТАБЛИЦА V

КОЭФФИЦИЕНТЫ К ДИАГРАММАМ КОНТРОЛЯ КАЧЕСТВА

Объем выборки	$A'_{0,025}$	$A'_{0,001}$	$D'_{0,975}$	$D'_{0,999}$	d_n
2	1,229	1,937	2,81	4,12	1,13
3	0,668	1,054	2,17	2,98	1,69
4	0,476	0,750	1,93	2,57	2,06
5	0,377	0,594	1,81	2,34	2,33
6	0,316	0,498	1,72	2,21	2,53
7	0,274	0,432	1,66	2,11	2,70
8	0,244	0,384	1,62	2,04	2,85
9	0,220	0,347	1,58	1,99	2,97
10	0,202	0,317	1,56	1,93	3,08

КОММЕНТАРИИ РЕДАКТОРА ПЕРЕВОДА

1. СТАТИСТИЧЕСКИЙ МЕТОД И ТЕОРИЯ ВЕРОЯТНОСТЕЙ

Вероятности появляются в книге Браунли под двумя разными названиями: под их собственным именем и под именем относительных частот в „бесконечных совокупностях". Воображаемые „бесконечные совокупности", о которых, следуя Фишеру, говорит Браунли в разделе (С) главы II, в действительности не существуют, а в качестве математической фикции ничему не помогают: например, если представить себе бесконечную совокупность электрических лампочек, из которых бесконечное число исправно и бесконечное число дефектно, то для „относительной частоты" дефектных лампочек можно получить только неопределенное выражение $\frac{\infty}{\infty}$.

Однако является вполне естественным и соответствующим большому ряду наблюдений допущение, что каждому достаточно точно определенному состоянию процесса производства лампочек (в отношении качества материалов, изношенности оборудования, тщательности работы и т. п.) соответствует определенная вероятность для каждой отдельной произведенной лампочки оказаться дефектной. В соответствии с законами вероятностей в больших партиях лампочек, произведенных при стационарном состоянии производства, относительная частота появления дефектных лампочек будет близка к этой теоретической вероятности и, как правило, тем ближе, чем размеры партии больше.

Несмотря на то, что понятие вероятности довольно сложно по своему происхождению, оно лежит в основе всех теоретических оценок достоверности статистических выводов. В этом заключается верное зерно высказывания Фишера о том, что „основой для всей статистической работы является идея бесконечной совокупности".

Между понятиями, относящимися к конечным совокупностям объектов (в терминологии Браунли „индивидуумов"), и понятиями

теории вероятности имеется далеко идущая аналогия, некоторые стороны которой собраны в следующей таблице:

Для совокупности из n индивидуумов	В теории вероятностей
1. Относительная частота $\frac{m}{n}$ индивидуумов, обладающих каким-либо признаком (где m есть число индивидуумов с данным признаком)	Вероятность появления в данных условиях индивидуума с данным признаком
2. Переменное, принимающее для каждого индивидуума определенное значение $x_1, x_2, \ldots, x_n$ (где $1, 2, \ldots, n$ — номера индивидуумов)	Случайная величина ξ, способная принимать различные значения (число этих возможных значений может быть и бесконечным)
3. Среднее значение $\bar{x} = \frac{x_1 + x_2 + \ldots x_n}{n}$	Математическое ожидание $E(\xi)$ случайной величины ξ
4. Эмпирическая дисперсия $\sigma_x^2 = \frac{1}{n}\sum_{k=1}^{n}(x_k - \bar{x})^2$ или* $\sigma_x^2 = \frac{1}{n-1}\sum_{k=1}^{n}(x_k - \bar{x})^2$	Дисперсия $\sigma_\xi^2 = D(\xi) = E\{\xi - E(\xi)\}^2$ случайной величины ξ

Строгое различие между этими двумя рядами понятий является необходимой предпосылкой более отчетливого, с теоретической стороны, изложения вопросов, изучаемых в книге Браунли.

2. УРОВНИ ЗНАЧИМОСТИ ПРИ ПРОВЕРКЕ ГИПОТЕЗ

Главы III, IV и V книги Браунли посвящены различным простейшим задачам статистической проверки гипотез и оценки параметров. Рассмотрим несколько подробнее первый пример Браунли на проверку гипотез, изложенный в разделе (А) главы III.

* Браунли пользуется вторым выражением с $n-1$ в знаменателе. Объяснение этому обстоятельству см. в тексте (гл. II, (G)).

Оцениваемая гипотеза заключается в том, что между испытателями A и B нет „систематического" различия в отношении доставляемых ими результатов. Более точно, проверяемая гипотеза состоит в том, что математическое ожидание разности результатов, получаемых нашими испытателями, равно нулю. Так как далее используются таблицы критерия t Стьюдента, то по существу делается еще допущение, что разность результатов, получаемых (при исследовании одной и той же смеси) испытателями, подчинена закону распределения Гаусса, а результаты испытаний различных смесей независимы.

Таким образом, математическое содержание задачи сводится к следующему: рассматривается гипотеза, что случайные величины

$$\xi_1,\ \xi_2, \ldots,\ \xi_{10}$$

независимы между собой и подчинены одному и тому же закону распределения Гаусса со средним, равным нулю, и постоянной (хотя и неизвестной нам) дисперсией σ_ξ^2. В интересующем нас частном случае фактически наблюденные значения x_k случайных величин ξ_k оказались равными

$$-4;\ 2;\ -2;\ 0;\ -1;\ 1;\ -1;\ -1;\ -2;\ -3.$$

Так как эмпирическое среднее

$$\overline{x} = \frac{-4+2-2+0-1+1-1-1-2-3}{10} = -1{,}10$$

отрицательно, то возникает подозрение, не является ли и математическое ожидание $E(\xi)$ тоже отрицательным, т. е. не имеет ли испытатель B систематической склонности давать более высокие процентные содержания интересующего нас ингредиента смеси, чем испытатель A. Если бы такая систематическая склонность была точно установлена, то гипотеза

$$E(\xi) = 0$$

должна была бы считаться опровергнутой.

Для более точной количественной оценки положения, Браунли, в соответствии с процедурой, стандартной в современной английской статистике, рекомендует вычислить эмпирическую дисперсию

$$\sigma_x^2 = \frac{1}{10-1}\sum_{k=1}^{10}(x_k - \overline{x})^2 = 1{,}79$$

и образовать стьюдентово отношение *

$$t = \frac{\overline{x}}{\sigma_x/\sqrt{n}} = -1{,}95.$$

* Браунли опускает знак минус при 1,95 как несущественный для дальнейшего.

В таблице I в приложении даны для вероятностей

$$p = 0{,}10;\ 0{,}05;\ 0{,}02;\ 0{,}01;\ 0{,}001$$

значения таких t_p, для которых

$$\text{вероятность } \{|t| \geqslant t_p\} = p.$$

Вероятности эти зависят от „числа степеней свободы", которое в нашем случае равно

$$10 - 1 = 9.$$

В соответствии с этим для нашего случая

$$t_{0,10} = 1{,}83;\ t_{0,05} = 2{,}26;\ t_{0,02} = 2{,}82;\ t_{0,01} = 3{,}25;\ t_{0,001} = 4{,}78.$$

Это значит, что в предположении правильности гипотезы $E(\xi) = 0$ при неограниченном повторении серий, по десять определений значений

$$x_1,\ x_2, \ldots,\ x_{10}$$

в каждой, в описанной выше обстановке мы получили бы

$$|t| \geqslant 1{,}83 \quad \text{в } 10\% \text{ случаев,}$$
$$|t| \geqslant 2{,}26 \quad \text{в } 5\% \text{ случаев и т. д.}$$

Фактически наблюденное в рассматриваемом специальном случае значение $t = -1{,}95$ таково, что такие же или большие по абсолютной величине значения t должны встречаться при гипотезе $E(\xi) = 0$ в среднем более чем в 5% случаев, но менее чем в 10% случаев. По мнению Браунли, это еще не позволяет делать из наблюденных фактов каких-либо определенных выводов.

Наоборот, неравенство $|t| \geqslant t_{0,05}$ по Браунли уже достаточно, чтобы считать гипотезу $E(\xi) = 0$ опровергнутой. Впрочем, он сам не скрывает условности такой точки зрения. Несомненно также, что в каждом отдельном случае наше доверие к гипотезе после какой-либо ее проверки никогда нельзя отделить от ее оценки *a priori* до данной проверки. Это и учитывается в классической теории оценки вероятностей гипотез на основе теоремы Байеса.

Тем не менее, в производственной практике существенно *стандартизировать* правила проверки гипотез с тем, чтобы в неизменной общей обстановке эти правила могли применяться много раз совершенно автоматически. Этот расчет на многократное автоматическое применение, с одной стороны, делает такую стандартизацию неизбежной, но, с другой стороны, к счастью, делает ее *вполне обоснованной:* при установлении стандартной процедуры контроля нас совсем не интересует вопрос о том, будет ли в каждом отдельном случае ее результат соответствовать чьей-либо субъективной (или даже объективно обоснованной при учете индивидуальных обстоятельств случая) оценке положения, а интересует только процент ошибочных заключений при массовом применении данной процедуры.

В общей форме система проверки гипотез, разработанная английской статистической школой и принятая в книге Браунли, может быть изложена так. Для проверки гипотезы H измеряются величины

$$\xi_1, \xi_2, \ldots, \xi_n.$$

Из значений этих величин образуется такая комбинация:

$$k = f(\xi_1, \xi_2, \ldots, \xi_n),$$

которая при гипотезе H является случайной величиной с вполне определенным законом распределения. Эта величина называется „критерием“ для проверки гипотезы (t-, F-, χ^2-критерием). Для каждой вероятности p устанавливаются такие границы k'_p и k''_p, что при гипотезе H

$$\text{вероятность } \{k'_p < k < k''_p\} = 1 - p,$$
$$\text{вероятность } \{k \leqslant k'_p \text{ или } k \geqslant k''_p\} = p.$$

Если следовать правилу, в каждом отдельном случае отбрасывать гипотезу H, когда наблюденное значение k не укладывается в пределах $k'_p < k < k''_p$ и считать гипотезу согласующейся с результатами наблюдений, когда k укладывается в эти пределы, то при массовом применении правила мы будем отбрасывать гипотезу H ошибочно в доле случаев, равной p.

Вероятность p, которую кладут в основу избранного правила проверки, называется *уровнем значимости* данного правила. Выбор уровня значимости является очень важным и ответственным делом: если с увеличением вероятности p увеличивается вероятность ошибочного отбрасывания гипотезы, то с уменьшением этой вероятности для эффективного применения правила требуется все большее число испытаний. Часто целесообразно введение нескольких уровней значимости, например, по такой схеме: гипотеза ставится под сомнение при переходе критерием границ, соответствующих уровню значимости $p = 0{,}05$, и в этом случае производятся дополнительные испытания для ее проверки, безоговорочно же отбрасывается лишь при переходе границ, соответствующих уровню значимости $p = 0{,}01$.

Кроме того, для проверки одной и той же гипотезы может быть предложено много различных критериев, и целесообразный выбор критерия для проверки данной гипотезы или данного типа гипотез является одной из центральных проблем математической статистики. Но практик, ограничивающийся простейшими применениями уже разработанных методов математической статистики,

этим вопросом может не заниматься; если уровень значимости в каждом случае ему придется выбирать самому, исходя из конкретных условий производства, то в отношении выбора критерия в простейших случаях он может следовать стандартным указаниям руководств: при оценке гипотез о значении математических ожиданий, или равенстве математических ожиданий, лежащих в основе двух выборок, он будет пользоваться критерием t, при оценке гипотезы о равенстве двух дисперсий — критерием F и т. д.

В заключение сделаем еще два предостережения:

1) Статистическая проверка гипотез в изложенной выше форме в определенных случаях, как мы видели, приводит к *отбрасыванию* рассматриваемой гипотезы H. Наоборот, для окончательного утверждения гипотезы H в качестве безусловно истинной сформулированные приемы не дают ничего определенного. Естественно, что попадание критерия, вычисленного по данным выборочного наблюдения, в разумные границы служит „подтверждением“ гипотезы, но никакой количественной оценки возникающей отсюда „уверенности“ в справедливости гипотезы H правила рассмотренного типа не дают.

2) Вероятность p, названная выше *уровнем значимости*, должна правильно интерпретироваться так, как это было выше объяснено.

В тех случаях, когда закон распределения критерия k непрерывен (как это имеет место во всех случаях, рассмотренных у Браунли), для каждого наблюденного значения k можно найти такой уровень значимости p, при котором наблюденное значение k совпадает с одной из границ k'_p или k''_p (в рассмотренном выше примере на применение критерия t имеем $t'_p = -t_p$, $t''_p = t_p$ и дело сводится к определению такого p, для которого $|t| = t_p$). Часто смешивают это значение уровня значимости p с апостериорной вероятностью гипотезы H после данного результата наблюдений (получения данного значения критерия k). Как уже указывалось, такое смешение *грубо ошибочно*. Переходя же к практическим выводам, мы рекомендуем принять, что по существу нормативное значение имеют только уровни значимости, установленные заранее для целой серии однотипных испытаний.

3. ДОВЕРИТЕЛЬНЫЕ ГРАНИЦЫ ПРИ ОЦЕНКЕ ПАРАМЕТРОВ

Некоторым развитием изложенной выше методики оценки отдельной гипотезы является принятая в книге Браунли методика оценки параметров при помощи доверительных границ с данным уровнем значимости.

Для иллюстрации рассмотрим пример, изложенный в разделе (С) главы IV. Предполагается, что десять случайных величин

$\xi_1, \xi_2, \ldots, \xi_{10}$ независимы и подчинены одному и тому же закону распределения Гаусса с неизвестным нам средним и неизвестной дисперсией σ_ξ^2. Желательно иметь общее правило оценки неизвестной теоретической дисперсии σ_ξ^2 по эмпирической дисперсии

$$\sigma_x^2 = \frac{1}{9} \sum_{k=1}^{10} (x_k - \bar{x})^2,$$

составленной из фактически наблюденных значений x_k случайных величин ξ_k, имеющих уровень значимости 0,10, т. е. правило, которое при систематическом применении в указанной выше обстановке приведет к ошибочному результату лишь в 10% случаев. Браунли в качестве такого правила предлагает оценку

$$\frac{9\sigma_x^2}{16{,}92} < \sigma_\xi^2 < \frac{9\sigma_x^2}{3{,}32}.$$

Оценка эта основана на том, что отношение

$$\chi^2 = \frac{9\sigma_x^2}{\sigma_\xi^2}$$

имеет закон распределения χ^2 с девятью степенями свободы, для которого, в соответствии с таблицей II,

$$\text{вероятность } \{3{,}32 < \chi^2 < 16{,}92\} = 0{,}05 + 0{,}05 = 0{,}10.$$

Вполне аналогична и общая теория доверительных границ в том ее простейшем варианте, который достаточен для случаев, рассмотренных у Браунли. Значение критерия k составляется из оцениваемого параметра θ и наблюдаемых величин так, чтобы критерий k имел вполне определенный закон распределения. Тогда из неравенств

$$k_p' < k < k_p'',$$

имеющих вероятность $1-p$, где p — избранный уровень значимости, вытекают определенные оценки $\theta' < \theta < \theta''$ параметра θ через наблюдаемые величины. Этими оценками и рекомендуется пользоваться, если решено, что уровень значимости p отвечает требующейся в данной обстановке достоверности оценок. Значения θ' и θ'' называются *доверительными границами* для параметра θ, соответствующими уровню значимости p.

4. ДИСПЕРСИОННЫЙ АНАЛИЗ

Изложение дисперсионного анализа у Браунли чрезмерно подчинено пропаганде той идеи, что возможно получать достаточно интересные результаты, производя только по одному испытанию при каждой комбинации факторов. Мы советовали бы начинать ознакомление с идеями дисперсионного анализа с примера в конце раздела (Н) главы XI. Этот пример не слишком сложен (два фактора), но позволяет ознакомиться со способом оценки взаимодействия факторов и делает понятным значение повторных испытаний при одинаковых комбинациях факторов (для каждой из четырех комбинаций сделано по два испытания).

В этом примере рассматриваются измерения величины, о которой предположено, что она зависит от двух факторов D и N и, быть может, еще от многих не учтенных факторов (к которым можно отнести и ошибки измерений). Для каждого из факторов выбраны по два уровня (D_1 и D_2 для первого и N_1 и N_2 для второго) и при каждой комбинации уровней произведено по два измерения. Результаты таковы:

при уровнях D_1 и N_1 получилось -12 и -5,
„ „ D_1 „ N_2 „ $+27$ „ $+27$,
„ „ D_2 „ N_1 „ -29 „ -22,
„ „ D_2 „ N_2 „ $+34$ „ $+24$.

Теоретические допущения дисперсионного анализа для нашего случая таковы:

1) Для каждой комбинации уровней факторов имеется свое среднее значение измеряемой величины:

a_{11} для комбинации уровней D_1 и N_1,
a_{12} „ „ „ D_1 „ N_2,
a_{21} „ „ „ D_2 „ N_1,
a_{22} „ „ „ D_2 „ N_2.

2) Отклонение результата измерения от соответствующего среднего a_{ij} подчинено закону распределения Гаусса с некоторой, постоянной во всех случаях, дисперсией σ^2 (математическое ожидание этого отклонения по самому определению равно нулю). Отклонения эти при различных измерениях независимы.

Обозначим теперь

$$a_{1*} = \frac{a_{11}+a_{12}}{2}, \quad a_{2*} = \frac{a_{21}+a_{22}}{2},$$

$$a_{*1} = \frac{a_{11}+a_{21}}{2}, \quad a_{*2} = \frac{a_{12}+a_{22}}{2},$$

$$a_{**} = \frac{a_{11}+a_{12}+a_{21}+a_{22}}{2}.$$

При этих обозначениях

$$a_{ij} = a_{**} + (a_{i*} - a_{**}) + (a_{*j} - a_{**}) + \left(a_{ij} - \frac{a_{i*}+a_{*j}}{2}\right).$$

В дисперсионном анализе второй член полученного выражения для a_{ij} приписывается действию фактора D, третий — действию фактора N, а четвертый — их взаимодействию.

Так как результат k-го измерения при комбинации уровней D_i, N_j может зависеть еще от неучитывающихся факторов (в том числе от ошибок измерения), то он выражается в виде

$$\xi_{ij}^k = a_{**} + (a_{i*} - a_{**}) + (a_{*j} - a_{**}) + \left(a_{ij} - \frac{a_{i*} + a_{*j}}{2}\right) + \Delta_{ij}^k.$$

Легко теперь подсчитать, что среднее из восьми квадратов отклонений от общего среднего ξ_{ij}^k имеет вид

$$\frac{1}{8}\sum_{i,\,j,\,k}(\xi_{ij}^* - a_{**})^2 = \sigma_d^2 + \sigma_n^2 + \sigma_{dn}^2 + \varphi^2,$$

где

$$\sigma_d^2 = \frac{1}{2}\left[(a_{1*} - a_{**})^2 + (a_{2*} - a_{**})^2\right],$$

$$\sigma_n^2 = \frac{1}{2}\left[(a_{*1} - a_{**})^2 + (a_{*2} - a_{**})^2\right],$$

$$\sigma_{dn}^2 = \frac{1}{4}\sum_{i,j=1,2}\left(a_{ij} - \frac{a_{i*} + a_{*j}}{2}\right)^2,$$

$$\varphi^2 = \frac{1}{8}\sum_{i,\,j,\,k=1,2}(\Delta_{ij}^{(k)})^2.$$

В дисперсионном анализе считают σ_d^2 за меру влияния на измеряемую величину фактора D, σ_n^2 — фактора N, σ_{dn}^2 — их взаимодействия. Их называют *компонентами дисперсии**. Оценка этих σ_d^2, σ_n^2 и σ_{dn}^2 на основе наблюденных значений величин $\xi_{ij}^{(k)}$ и является задачей дисперсионного анализа. С этой целью составляются по определенным правилам, которые достаточно ясно изложены у Браунли, „средние квадратов“:

$$4\sigma_d^2 + 2\sigma_{dn}^2 + \sigma_{0(d)}^2 = 112{,}5.$$
$$4\sigma_n^2 + 2\sigma_{dn}^2 + \sigma_{0(n)}^2 = 4050{,}0,$$
$$2\sigma_{dn}^2 + \sigma_{0(dn)}^2 = 180{,}5,$$
$$\sigma_0^2 = 24{,}75.$$

---

* Обозначив через Σ^2 математическое ожидание

$$\frac{1}{8}\sum_{i,\,j,\,k}(\xi_{ij}^{(k)} - a_{**})^2,$$

мы получили бы

$$\Sigma^2 = \sigma_d^2 + \sigma_n^2 + \sigma_{dn}^2 + \sigma^2.$$

Σ^2 называют несколько условно „полной дисперсией“. Это объясняет как название „компонента дисперсии“, так и то, что σ^2 называется *остаточной дисперсией* (истинной).

Вместо наших $\sigma_{0(d)}^2$, $\sigma_{0(dn)}^2$, $\sigma_{0(n)}^2$ и σ_0^2 Браунли всюду пишет просто σ_0^2. В действительности это *различные величины*, имеющие лишь то общее, что все они при сделанных предположениях подчинены закону распределения Гаусса с математическим ожиданием, равным нулю, и дисперсией σ^2. Между собой же они совсем не связаны, а наоборот, *независимы*.

Единственный вопрос, на который в книге Браунли дается вполне обоснованный, при подчеркнутых выше допущениях, ответ, это вопрос о проверке гипотезы отсутствия влияния изучаемых факторов или их взаимодействий. Ответ на этот вопрос дается при помощи таблицы III по общей схеме теории проверки гипотез. Когда факторов много, то достигаемое таким образом выделение заведомо существенных факторов и их взаимодействий может оказаться очень полезным.

Следует, однако, помнить, что оговоренные выше допущения 1) и 2) редко выполняются в действительности, что заставляет относиться к результатам дисперсионного анализа с большей осторожностью, чем это видно из изложения Браунли. Кроме того, следует иметь в виду, что иногда существенность какого-либо фактора или взаимодействия факторов остается неуловленной из-за большого значения „истинной остаточной дисперсии“ σ^2, которую можно оценить по „эмпирической остаточной дисперсии“ σ_0^2, как указано в разделе (D) главы IX.

Браунли показывает на примерах, как для факторов и взаимодействий, существенность которых доказана, оценивать размеры их влияния. Но эти указания нельзя считать достаточно полными.

ОГЛАВЛЕНИЕ

$^1/_4$15 Зак. 1691

ГЛАВА X.

МНОЖЕСТВЕННАЯ КОРРЕЛЯЦИЯ

ГЛАВА XI

ОБЩИЙ ДИСПЕРСИОННЫЙ АНАЛИЗ

ГЛАВА XII

РАЗНЫЕ ВОПРОСЫ, СВЯЗАННЫЕ С ДИСПЕРСИОННЫМ АНАЛИЗОМ

ГЛАВА XIII

СБАЛАНСИРОВАННЫЕ НЕПОЛНЫЕ БЛОКИ

ГЛАВА XIV

СМЕШИВАНИЕ. ПРОБЛЕМА ОГРАНИЧЕННОСТИ ОБЪЕМА БЛОКА В ФАКТОРНЫХ ИСПЫТАНИЯХ

ГЛАВА XV

ОБЩИЕ ВЫВОДЫ

Редактор *Д. Васильков.*
Технический редактор *Б. Корнилов.*
Корректоры: *А. Алябьев* и *В. Соколов.*

Подписано к печати 5/II 1949 г. А-01488. Печ. л. $14^1/_4$. Уч.-издат. л. 13,8. Формат $60 \times 92^1/_{16}$. Издат. № 1/305. Цена 15 руб.
Заказ 1691.

2-я тип. „Печатный Двор" им. А. М. Горького треста „Полиграфкнига" ОГИЗа при Совете Министров СССР. Ленинград, Гатчинская, 26.