МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ АГЕНТСТВО ПО ОБРАЗОВАНИЮ

ТИХООКЕАНСКИЙ ГОСУДАРСТВЕННЫЙ ЭКОНОМИЧЕСКИЙ УНИВЕРСИТЕТ

Е.А. Любченко, О.А. Чуднова

ПЛАНИРОВАНИЕ И ОРГАНИЗАЦИЯ ЭКСПЕРИМЕНТА

Часть 1

Учебное пособие по дисциплине «Планирование и организация эксперимента» для студентов специальностей 200503 «Стандартизация и сертификация», 220501 «Управление качеством»

Владивосток Издательство ТГЭУ 2010 УДК 519.242: 519.242.7

Л 93

Любченко Е.А., Чуднова О.А. Планирование и организация экспери-

мента: учебное пособие. Часть 1. – Владивосток: Изд-во ТГЭУ, 2010. – 156 с.

Пособие содержит начальные сведения из теории планирования и ор-

ганизации эксперимента, информацию по проверке статистических гипотез и

различных инструментах статистического анализа данных с последующим

построением математической модели эксперимента.

Рекомендовано для студентов 3-5 курсов специальностей 200503

«Стандартизация и сертификация и 220501 «Управление качеством», а также

всех, интересующихся вопросами прикладной статистики, обработки данных

и планирования эксперимента.

Печатается по решению УМС ТГЭУ

Рецензенты: А.В. Огнев, канд. физико-математических наук,

доцент (ДВГУ);

В.А. Петров, д-р медицинских наук, профессор (ВГМУ);

С.А. Щеголева, канд. физико-математических наук, доцент

(ДВГУ);

©Любченко Е.А.,

Чуднова О.А., 2010

© Изд-во ТГЭУ, 2010

СОДЕРЖАНИЕ

ВведениеОшибка! Закладка не опред	целена.
1 Введение в теорию «Планирования и организации	
эксперимента»	6
1.1 Планирование эксперимента и его задачи. Виды экспериментов.	
1.2 Параметры оптимизации и требования, предъявляемые к ним	
1.3 Факторы и требования к ним	
1.4 Выбор модели эксперимента	18
1.5 Принятие решений перед планированием	24
2 Статистическая проверка статистических гипотез	29
2.1 Статистические гипотезы. Виды ошибок при выдвижении	
статистических гипотез	29
2.2 Статистические критерии	33
2.3 Виды критериев согласия и области их применения	39
3 Статистические методы анализа данных и планировани	Я
экспериментов	
3.1 Дисперсионный анализ	
3.2 Корреляционный анализ	
3.3 Регрессионный анализ	
4 Ввеление в факторные планы	107
4.1 Полный факторный эксперимент типа 2 ^k	108
4.2 Полный факторный эксперимент и математическая модель	
эксперимента	114
4.3 Возвращение назад	118
4.4 Дробный факторный эксперимент типа 2 ^{k-p} : выбор полуреплик	
4.5 Выбор $1/4$ -реплик в ДФЭ- 2^k . Обобщающий определяющий контр	раст 130
Послесловие	133
Список литературы	134
Приложение А. О функциях случайных величин и их	
	136
Приложение В. Использование возможностей MS Excel в	100
статистических исследованиях	148
СТАТИСТИЧЕСКИХ ИССЛЕЛОВАНИЯХ	140

Введение 4

Введение

Многие специалисты, занятые исследованием тенденций развития российского рынка, отмечают появление плеяды руководителей нового поколения: людей, которых интересуют прибыли не здесь и сейчас, а с перспективой развития, роста их компаний. Никакое развитие предприятия невозможно без планирования процессов функционирования и без их грамотной организации. И тем более это невозможно без использования научных подходов планирования и организации деятельности. Как раз последнее и относится к сфере рассмотрения дисциплины «Планирование и организация эксперимента», о которой пойдет речь в данном пособии.

Проблемами «Планирования и организации эксперимента», насколько можно судить с высоты нашего опыта преподавания дисциплины, является:

- во-первых, ее «заброшенность» литературы очень мало и датируется она, в своей основе, 70-80-ми годами прошлого века;
- во-вторых, изначально данная дисциплина преподавалась на технических специальностях, и, соответственно, вся та немногая литература, которая имеется, ориентирована на соответствующую отрасль, рассчитанную именно на технически «подкованную» аудиторию.

Учитывая только две перечисленные проблемы, а при желании их можно найти и больше, вполне понятным становится, почему данное пособие актуально. В нем собрано все то, что в разрозненных вариантах есть в других литературных источниках по данной дисциплине, но систематизировано и переложено на простой, доступный студентам язык.

Отдельно хотелось бы остановиться на логике изложения представленного материала. На первый взгляд, первый раздел никоим образом не связан с материалом второго и третьего разделов, а четвертый раздел — вообще стоит особняком. Однако, это не так.

В первом разделе описываются основные понятия, которыми оперирует «Планирование и организация эксперимента». Эти понятия будут встречаться и в остальных разделах, по мере затрагивания проблем исследуемой дисциплины. Во втором и третьем разделах приводится описание статистических методов, с помощью которых проводится предварительный анализ априорной информации, отбор влияющих факторов и построение модели эксперимента.

При использовании большого числа влияющих факторов довольно остро встает вопрос сокращения числа опытов. Именно на это и направлена методика построения дробных реплик, описанная в четвертом разделе. В этом же разделе показано, как можно построить модель эксперимента без использования методики регрессионного анализа.

Таким образом, все четыре раздела пособия логически взаимосвязаны.

Дополнительный материал, изложенный в приложениях, содержит информацию по параметрам статистических функций, анализу и планированию экспериментов с использованием статистических функций Exel.

Представленное пособие является своеобразным помощником и проводником в изучении курса «Планирование и организация эксперимента» приобрести элементарные познания в области теории вероятностей и математической статистики.

1 Введение в теорию «Планирования и организации эксперимента»

Мысль о том, что эксперимент можно планировать восходит к глубокой древности. Пожалуй, как только человек взял в руки палку, он уже начал заниматься проблемами планирования с целью выработки наиболее оптимального способа добычи пропитания. Результатами подобных изысканий, проводившимся в течение столетий, стали современные блага цивилизации. Однако, первобытному человеку, да и средневековому рыцарю в том числе, абсолютно не были знакомы понятия статистики.

Подобная теория появилась (имеется в виду статистика) в начале – середине XX века. Вслед за развитием аппарата статистического анализа, его положения стали применяться и в планировании эксперимента. Автором идеи привлечения статистики в планирование являлся один из основоположников английской школы статистики – Рональд Фишер. Именно он доказал целесообразность использования статистических методов в проблеме поиска оптимальных условий проведения эксперимента. Так появилась совершенно новая наука, имеющая важное практическое значение – «Планирование и организация эксперимента».

1.1 Планирование эксперимента и его задачи.

Виды экспериментов

Так что же представляет собой планирование эксперимента? Для того чтобы представить себе этот процесс достаточно сказать, что мы с Вами ежедневно, ежечасно и даже ежеминутно занимаемся планированием эксперимента, и этот эксперимент называется жизнь.

Давайте для примера представим себе одно наше утро. Просыпаясь утром и собираясь выйти из дома, мы вспоминаем уже заранее намеченные на этот день дела или же намечаем их в эту самую минуту. При этом каждый из нас, рассматривая список предполагаемых дел, сразу проводит корректировку, что он точно способен сделать, что вероятнее всего сделает, на что сил может не хватить, но на всякий случай запишем это в реестр сегодняшних дел и т.д. Таким образом, каждый из нас прикидывает условия существования в дне сегодняшнем, чтобы данный эксперимент (мы все по-прежнему имеем ввиду – жизнь) у нас удался.

Точно таким же образом проводятся и промышленные эксперименты. С одной лишь оговоркой. При проведении различных лабораторных, промышленных или других экспериментов существуют какие-то нормативы точности полученных результатов. Ну, например, вес слона к концу проведения откорма должен составлять не менее (5000 ± 150) кг. И, откармливая слона, вполне естественно вы будете планировать свою животноводческую кампанию с учетом требуемого конечного веса *с точностью до 150 кг*.

Учитывая сказанное, можно сформулировать следующее определение.

Планирование эксперимента — это процедура выбора числа и условий проведения опытов, необходимых и достаточных для решения поставленной задачи с требуемой точностью.

При этом, как учит нас теория, необходимо придерживаться следующих ограничений:

- 1. общее число опытов должно быть по возможности минимальным;
- 2. необходимо одновременно изменять все переменные, определяющие (влияющие) процесс. Причем это изменение должно происходить по определенным правилам—алгоритмам;
- 3. при описании исследований необходимо использовать математический аппарат, формализующий действия экспериментатора;

4. в процессе проведения и планирования эксперимента необходимо придерживаться четкой стратегии, позволяющей принимать обоснованные решения после каждой серии экспериментов.

Задачей «Планирования эксперимента» является разработка рекомендаций или производственного процесса на основе исследования предварительных опытных данных для дальнейшей их реализации и построения математической модели исследуемого процесса с целью дальнейшего прогнозирования производства. Как правило, результатами таких исследований являются разработки наиболее оптимальных рекомендаций, технологического процесса, имеющих важные экономические, технические, технологические последствия и влекущих за собой как модернизацию отдельного технологического процесса, так и целого производства.

В зависимости от условий эксперименты делятся на несколько видов:

- промышленный –
 это эксперимент, поставленный в условиях предприятия с целью улучшения производства;
- научно-исследовательский –
 эксперимент, поставленный в научно-исследовательских лабораториях с
 целью исследования нового или улучшения существующего процесса,
 явления;
- 3) лабораторный эксперимент, поставленный в научно-исследовательских лабораториях с целью изучения хорошо известного, существующего процесса, явления;
- 4) *оптимальный* (экстремальный) эксперимент, поставленный с целью поиска наиболее оптимальных условий его реализации в заранее заданном смысле. С математической точки зрения, это эксперимент по поиску экстремумов некоторой функции, отсюда и второе название эксперимента;

дущих.

- 5) пошаговый эксперимент, состоящий из отдельных серий опытов. Причем условия проведения каждой следующей серии определяются результатами преды-
- 6) активный эксперимент, в ходе которого экспериментатор имеет возможность изменять и/или поддерживать на заданном уровне сколь угодно долго значение параметров, задающих условия проведения эксперимента;
- 7) пассивный эксперимент, в ходе которого экспериментатор НЕ имеет возможности изменять и/или поддерживать на заданном уровне сколь угодно долго значение параметров, задающих условия проведения эксперимента

На практике чаще всего приходится иметь дело со смешанным активно-пассивным экспериментом.

Как и в любой другой науке, «Планирование и организация эксперимента» имеет свой собственный язык, т.е. какие-то определенные термины, понятия. Ниже как раз и поговорим об этом.

1.2 Параметры оптимизации и требования, предъявляемые к ним

Прежде, чем проводить любой эксперимент, неважно научный он будет или нет, каждый из нас четко определяет для себя, а чего собственно он ждет в результате своей бурной деятельности? Причем желательно, особенно в случае промышленных или научных экспериментов, чтобы этот результат выражался количественно. В «Планировании и организации эксперимента» результат проведения опытов называется *параметром оптимизации* или *откликом системы* на воздействие.

Параметр оптимизации (отклик) — величина, описывающая результат проведенного эксперимента и зависящая от факторов, влияющих на эксперимент.

В зависимости от объекта и цели исследования параметры оптимизации могут быть самыми разнообразными. Введем классификацию параметров оптимизации:

<u>1 класс</u> Экономические параметры оптимизации.

К данному классу относятся *прибыль*, *себестоимость*, *рентабельность* (эти параметры используются при исследовании действующих промышленных объектов), *затраты на эксперимент* (оценивается в любых исследованиях, в т.ч. и научно-исследовательских).

2 класс Технико-экономические параметры оптимизации.

Среди этих параметров наиболее распространенными являются *производительность* и *коэффициент полезного действия*; такие параметры как *стабильность*, *надежность*, *долговечность* связаны с длительными наблюдениями и используются в основном при изучении дорогостоящих ответственных объектов.

<u>3 класс</u> Технико-технологические параметры оптимизации.

К этим параметрам оптимизации относятся физические характеристики продукта, механические характеристики продукта, физико-химические характеристики продукта, медико-биологические характеристики продукта, выход продукта. Как видно из перечня, данная категория параметров оптимизации оценивает качество выпускаемой продукции.

<u>**4 класс**</u> Прочие.

Эта категория содержит *психологические*, *эстемические*, *статистические* параметры оптимизации. Несмотря на кажущуюся простоту этой группы, данные параметры являются не менее важны-

ми, чем все предыдущие. С ростом сложности объекта растет и психологическая нагрузка на исполнителя, отчего очень сильно может измениться качество продукции. Эстетические же параметры прежде всего учитываются в вопросах повышения реализации.

В качестве примера выбора параметра оптимизации можно рассмотреть процесс обучения студента. Оценивать успешность проходящего процесса обучения можно различными вариантами, но наиболее оптимальным до сих пор остается балльная оценка знаний обучающегося. Исходя из приведенной выше классификации, данный параметр оптимизации относится, скорее всего, к четвертому виду – *прочие*.

Рассмотрим требования, предъявляемые к параметрам оптимизации.

Требование №1.

Прежде всего, параметр оптимизации должен быть количественным, задаваться числом. Исследователь должен иметь возможность его измерять при любом фиксированном наборе уровней факторов.

Вернемся, к оценке знаний. Не будь балльной оценки знаний, обучающемуся трудно было бы понять насколько его уровень знаний соответствует предъявляемым требованиям.

Множество значений, которые принимает параметр оптимизации, называется областью его определения.

Области определения могут быть дискретными и непрерывными. На практике, как правило, области определения дискретные.

Измерение параметра оптимизации предполагает наличие соответствующего прибора. В случае отсутствия такового по каким-либо причинам, приходится пользоваться приемом, называемым *ранжированием*: каждому параметру оптимизации присваиваются оценки по заранее выбранной шкале (двухбалльной, пятибалльной и т.д.), и в дальнейшем пользуются такой шкалой ранговой оценки при исследованиях. Фактически, мы качественным ве-

личинам присваиваем количественные значения. Яркий пример ранжированного подхода – балльная система оценки знаний.

Требование №2.

Параметр оптимизации должен выражаться одним числом. Не должно возникать таких ситуаций, когда один и тот же параметр описывается разными значениями. В противном случае возникают неясности и разночтения.

Примером таких разночтений может являться несоответствие в прочтении оценок, полученных при обучении. Приведу один яркий исторический пример. Однажды один мой знакомый рассказал, как он посещал Царскосельский лицей и там видел табель А.С. Пушкина. «Представляешь, – воскликнул мой знакомый, – а Пушкин-то был двоечником! У него в табеле одни двойки и колы стоят!» Конечно, можно и огорчиться, какого ужасного неуча записали в гении нации, если бы не одно «НО». В Царскосельском лицее была принята следующая система оценок:

- 1 отлично разбирается в предмете, имеет к нему склонность, желание, использует творческий подход;
- 2 неплохо разбирается в предмете, изучает без особого рвения, хотя и имеет склонность;
- 3 слабо разбирается в предмете, изучает без особого рвения, склонности к предмету слабые;
- 4 очень слабо разбирается в предмете, склонностей практически нет, изучает по принуждению;
- 0 не разбирается в предмете, склонностей не обнаружено, усвоение предмета практически отсутствует.

Вот тебе и двоечник! К слову сказать, во всем табеле у Пушкина была единственная плохая отметка – ноль по математике. Ну не его это был предмет.

Требование №3.

Однозначность параметра оптимизации в статистическом смысле: заданному набору *уровней факторов**) должно соответствовать, с точностью до ошибки эксперимента, одно значение параметра оптимизации. При этом обратное утверждение неверно, т.е. одно и то же значение параметра оптимизации может встречаться для разных наборов факторов.

Приведу пример. Хорошо известно, что для того, чтобы закипятить воду при нормальном давлении необходимо ее нагреть до 100 °C. И сколько бы раз вы не проводили этот опыт, результат будет один и тот же – при нормальном давлении и температуре 100 °C вода закипит. Однако при понижении давления температура кипения воды также снизится, т.е. получаем следующую ситуацию: другое сочетание значений температуры и давления даст тот же результат эксперимента – вода закипит.

Требование №4.

Параметр оптимизации должен быть эффективным с точки зрения достижения цели и в статистическом смысле. Фактически, это означает, что выбирать параметр оптимизации необходимо таким образом, чтобы он определялся с наибольшей возможной точностью.

Требование №5.

Параметр оптимизации должен удовлетворять требованию универсальности и полноты. Под универсальностью и полнотой параметра понимается его способность всесторонне охарактеризовать объект исследования.

Требование №6.

Параметр оптимизации должен иметь физическим смысл, быть простым и легко вычисляемым.

Требование физического смысла объясняется необходимостью дальнейшей интерпретации результатов эксперимента. Вообще говоря, можно

^{*)} О факторах и их уровнях будет подробно рассказано в следующем параграфе, но, для внесения ясности скажу, что уровни фактора – это те значения, которые фактор может принимать. Например, температура – фактор, значения температуры (0, 20, 100... °C) – его уровни.

параметр оптимизации описывать каким угодно выражением или способом, если только потом сможете объяснить, что это описание означает.

Легкость и простота вычислений позволяют проконтролировать правильность вычисления параметра оптимизации в процессе построения модели эксперимента.

Требование №7.

И, наконец, параметр оптимизации должен существовать для всех состояний системы.

Если жизнь на Марсе невозможна ни при каких состояниях, то выбирать в качестве результата эксперимента данное требование крайне неразумно.

Исходя из перечисленных требований, видно, что выбрать подходящий параметр оптимизации является делом довольно-таки трудоемким. Однако, именно правильный выбор параметра оптимизации является залогом успеха при дальнейшем планировании, поскольку выбор параметра оптимизации диктует вид математической модели эксперимента.

1.3 Факторы и требования к ним

После того, как выбран объект исследования и определен параметр оптимизации, необходимо определиться с величинами, которые могут влиять на процесс. В «Планировании и организации эксперимента» эти величины называются факторами. Упущенный существенный фактор ведет к абсолютно неправильным прогнозам и модели эксперимента, а лишний несущественный фактор только добавит хлопот при исследовании модели. Обычно рекомендуется использовать при планировании не более 15 факторов, если

же их больше – выбирать наиболее значимые, оставляя менее значительные факторы в стороне.

Фактор — измеряемая величина, описывающая влияние на объект исследования. Каждое значение, принимаемое фактором, называется уровнем фактора.

Так же как и параметр оптимизации, каждый фактор имеет область определения — совокупность всех значений, которые может принимать данный фактор.

Каждый фактор может принимать в опыте одно из нескольких значений. Фиксированный набор уровней нескольких факторов, т.е. их определенных фиксированных значений, будет определять какие-то конкретные условия проведения эксперимента. При изменении хотя бы одного из факторов в таком наборе приведет к изменению и условий и, как следствие, к изменению значения параметра оптимизации.

Для иллюстрации вернемся к примеру с кипящей водой, описанному в предыдущем параграфе. В рассмотренном примере используются два фактора — температура и давление, каждый из которых принимает определенные значения, т.е. принимает определенные уровни. Например, для давления — нормальное давление (760 мм рт. ст.), повышенное давление (скажем, 900 мм рт. ст.), пониженное давление (700 мм рт. ст.); для температуры — 50, 100, 1000 °C. Задавая те или иные значения температуры и давления, мы получим, что в одних случаях вода испариться почти мгновенно, в других — лишь слегка нагреется, в третьих — она закипит. Таким образом, меняя комбинации давления и температуры, говоря научным языком используя разные комбинации уровней двух факторов, мы определяем новые условия для проведения эксперимента и в то же время получаем другой результат.

Если перебрать все возможные наборы состояний, мы получим полное число возможных различных опытов. При этом число различных состояний

системы определяет ее сложность. Если обозначить число факторов, оказывающих влияние на эксперимент, как k, а число уровней, принимаемых каждым из факторов, буквой m, то число возможных состояний системы, т.е. число всех возможных опытов, определяется формулой:

$$N = m^k$$
.

Факторы бывают двух типов:

количественные – их можно оценивать количественно: измерять, взвешивать, титровать и т.п.;

качественные – количественно данный фактор задать не удается. Это разные вещества, технологические способы и т.п.

Требования, предъявляемые к факторам.

Требование №1.

Факторы должны быть управляемыми, т.е. экспериментатор должен иметь возможность, выбрав нужное значение фактора, поддерживать его постоянным на протяжении всего эксперимента.

Например, температура конфорки, на которую поставили подогревать воду — управляемая величина, мы можем ее величину менять самостоятельно и поддерживать постоянной сколько нам угодно; температура в комнате, где проходит эксперимент — неуправляемая величина, т.к. способов воздействовать на нее у нас практически нет и поддерживать ее на том или ином уровне для экспериментатора проблематично. В этом случае, при планировании эксперимента по нагреву воды мы в качестве фактора можем учитывать лишь первую температуру. Второй же показатель мы можем лишь принять во внимание.

Требование №2.

Фактор должен быть операциональным, т.е. можно указать последовательность действий (операций), необходимых для задания того или иного значения фактора.

Для того, чтобы переключить регулятор температуры на конфорке, каждый из нас предпринимает определенную последовательность действий, и мы можем ее точно описать (подойти к конфорке, повернуть регулятор и т.д.). А попробуйте маленькому ребенку лет трех-четырех просто сказать:

- Включи чайник!

Если он делает это впервые, он просто-напросто вас не поймет. Во втором случае мы имеем дело с нарушением принципа операциональности.

Требование №3.

Точность замера фактора должна быть как можно выше. Степень точности определяется диапазоном изменения факторов.

Требование №4.

Факторы должны быть однозначны, т.е. непосредственно влиять на объект исследования. Трудно изменять фактор, который является функцией других факторов.

Например, в качестве влияющего фактора мы бы очень не рекомендовали использовать женское настроение, поскольку трудно понять, что именно влияет на него в ту или иную минуту. А даже если и поймете, то в этом случае в качестве фактора лучше выбрать именно то, что влияет, дабы регулировать это настроение.

При планировании эксперимента редко рассматривается один фактор, обычно берется в рассмотрение сразу несколько факторов. Поэтому возникает необходимость формулировать *требования*, *предъявляемые к совокупности факторов*.

Требование №1.

Прежде всего, факторы должны быть совместимы. Совместимость факторов означает, что все их комбинации осуществимы и безопасны. Несовместимость факторов может наблюдаться на границах областей их определения. Избавиться от несовместимости можно, если в каждой области брать подобласть несколько меньшего размера. Положение усложняется, если несовместимость наблюдается внутри областей определения факторов. В этом случае приходится производить разбиение областей определения на несколько подобластей, «вырезая» кусок несовместимости, и ставить несколько планов экспериментов.

Требование №2.

При планировании также важна независимость факторов, т.е. возможность установления факторов на каком-либо уровне вне зависимости от значений уровней других факторов. Иначе это требование называют требованием отсутствия корреляции между факторами. Если между факторами наблюдается зависимость среднего или высокого уровня, один из двух факторов не принимают в рассмотрение.

1.4 Выбор модели эксперимента

Нередко при построении модели приходится принимать решение о выборе самого объекта, а именно, какие его характеристики и поведенческие функции следует учитывать, а какие – не вписываются в рамки поставленной задачи. В планировании эксперимента любого исследователя, прежде всего, интересует как поведет себя система, если на нее подействовать определенным образом. При этом ни одного из экспериментаторов абсолютно не интересует, что при этом «чувствует» сама система. Модели подобного рода, когда рассматривается только влияние на объект и его ответ на это влияние без учета внутренних процессов объекта, часто представляются так называемым

черным ящиком. При этом, воздействие на систему интерпретируется как входы черного ящика, а ответ системы на влияние – его выход, рисунок 1.1.

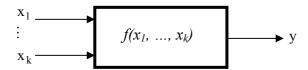


Рисунок 1.1. Модель объекта исследования в виде черного ящика, $(x_1...x_k - \phi$ акторы, действующие на объект, y – отклик системы)

В изучаемой нами теории под моделью также часто понимают модель черного ящика, в которой используется функция, устанавливающая зависимость между параметром оптимизации и факторами, рисунок 1.1:

$$y = f(x_1, x_2, ..., x_k).$$

Данная функция носит название *функции отклика*. С этих позиций, выбрать модель — значит выбрать вид этой функции, записать ее уравнение. Тогда только останется провести эксперимент по вычислению численных коэффициентов данной модели.

Иногда вместо алгебраической формы, т.е. уравнения, функцию отклика удается представить в геометрической форме. В этом случае речь заходит о *поверхности отклика*. Поиск решения в геометрической форме намного более нагляден, чем в виде уравнения. Однако, если число фактора больше двух, построение функции отклика невозможно, и приходится ограничиваться только алгебраической формой.

Остановимся на поверхности отклика подробнее. Для удобства рассмотрения представим систему, на которую влияют два фактора — x_1 и x_2 . Для того чтобы отобразить модель, достаточно располагать плоскостью с обычной Декартовой системой координат, по осям которых располагаются уровни каждого из факторов. Тогда каждому состоянию системы, т.е. «ящика» будет соответствовать точка на плоскости. Так как для каждого из факторов существуют области определения, у каждого фактора есть максимальное и минимальное возможные значения, между которыми и изменяется тот или иной фактор. Если факторы совместимы, границы их областей определения образуют на плоскости некоторый прямоугольник — область совместного существования факторов, рисунок 1.2.

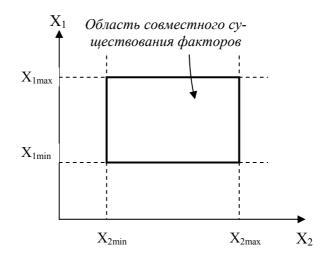


Рисунок 1.2. Пример факторного пространства

Пространство, образованное осями факторов (иногда осями факторов и осью параметра оптимизации), называется факторным пространством.

Чтобы указать значения параметра оптимизации требуется еще одна ось координат – ось отклика. Если ее добавить, графическая модель эксперимента примет вид, представленный на рисунке 1.3. Объект подобного вида носит название *поверхности отклика*.

Размерность факторного пространства зависит от числа факторов. Однако, если число факторов больше двух, построить поверхность отклика уже нельзя и приходиться ограничиваться только алгебраическим языком, т.е. уравнением функции отклика.

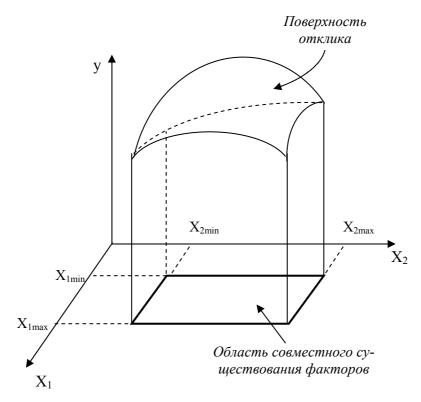


Рисунок 1.3. Поверхность отклика

Но для двух факторов можно даже не переходить к трехмерному пространству, а ограничиться плоскостью. Для этого достаточно произвести сечение поверхности отклика плоскостями, параллельными плоскости x_1Ox_2 , и полученные в сечениях линии спроектировать на эту плоскость, рисунок 1.4. Каждая линия, полученная в результате сечения, соответствует постоянному значению параметра оптимизации. Такая линия называется *линией равного отклика*.

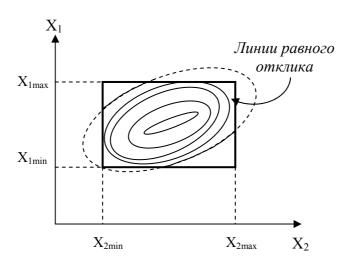


Рисунок 1.4. Проекция сечений поверхности отклика на плоскость

Как же найти те оптимальные условия эксперимента, которые нас интересуют? Причем было бы неплохо, чтобы этот поиск не требовал особых затрат. В этом случае мы прибегаем к математической модели эксперимента, с помощью которой можно предсказывать отклик системы в тех состояниях, которые экспериментально не изучались. В этом случае появляется возможность прогнозирования результатов эксперимента в точках, являющихся оптимальными в рамках поставленной задачи. И здесь мы переходим к пошаговому принципу.

Однако, прежде, чем приступать к моделированию, необходимо определиться с основными требованиями к поверхности отклика, на основе которой мы и собираемся делать прогнозы.

Требование №1.

Непрерывность поверхности — если к какой-либо точке факторного пространства функция отклика терпит разрыв, нет никакой гарантии, что при реальном осуществлении эксперимента данное состояние либо вообще невозможно, либо приведет к фатальным последствиям. При выборе большого шага перебора уровней факторов можно просто не заметить этот разрыв, «перешагнув» через него, однако вероятность попадания в эту критическую область на практике довольно-таки велика, и результат будет самым непредсказуемым.

Требование №2.

Гладкость поверхности отклика (соображения те же, что и в предыдущем пункте).

Требование №3.

Наличие единственного оптимума. Данное требование, пожалуй, одно из самых важных. При планировании эксперимента поиск оптимума может вестись в разных направлениях – и вправо, и влево. Если же оптимумов не-

сколько, да они еще и неравноценны, нет никакой гарантии, что наткнувшись на один из них, мы посчитаем данный оптимум именно тем решением, которое мы ищем, в то время, как это предположение неверно. Если же оптимум будет единственным, неважно с какой стороны мы будем к нему приближаться.

Суть шагового принципа сводится к следующему. Если нам известен вид поверхности отклика, кроме того, выполняются все требования для нее, можно заранее теоретически выбрать направление, в котором следует двигаться в поисках оптимального решения, будь то максимум или минимуму функции отклика (в зависимости от поставленной цели). Проведя эксперимент в выбранном направлении, по результатам определяемся, в каком направлении двигаться дальше. В конце концов, рано или поздно, реализовывая такие серии экспериментов и постоянно согласовываясь с видом поверхности отклика, мы найдем требуемый максимум.

Вообще говоря, моделей существует великое множество, а нам нужна одна единственная. Чтобы выбрать ее необходимо определиться, какие требования нужно предъявлять к модели.

Требование №1.

Главное требование к модели эксперимента — способность предсказывать дальнейшее направление опытов с требуемой точностью. При этом точность предсказания не должна зависеть от направления, в котором мы двигаемся при планировании, т.е. точность предсказания должна быть одинакова во всех направлениях.

Требование №2.

Адекватность модели. Данное требование означает, что модель <u>действештельно</u> должна предсказывать экспериментальные данные.

Требование №3.

Среди всех моделей необходимо выбирать ту, которая является наиболее простой. При этом понятие простоты довольно-таки относительно и зависит от решаемой проблемы. Прежде чем выбирать ту или иную функции

нужно дополнительно задаться вопросом, а что подразумевается в данном случае под простотой – вид уравнения или легкость описания?

Наиболее часто в планировании эксперимента останавливаются на полиномиальных моделях вида

 $y=b_0-\text{полином нулевой степени};$ $y=b_0+b_1x_1+b_2x_2-\text{полином 1-ой степени (линейный)};$ $y=b_0+b_1x_1+b_2x_2+b_{12}x_1x_2+b_{11}x_1^2+b_{22}x_2^2-\text{полином 2-ой степени}.$

Увеличивая степень полинома, можно задать приблизительное описание (аппроксимацию) функции любой сложности. Для экспериментатора же выбор полиномиальной модели позволяет значительно упростить поиск числовых коэффициентов. При выборе степени полинома нужно не забывать о простоте описания. Слишком высокие степени, несмотря на увеличение точности предсказания, редко приветствуется, поскольку с каждой новой степенью затрудняется поиск числовых коэффициентов. При увеличении коэффициентов растет и число опытов, необходимых для их вычисления. Чаще всего экспериментаторы стараются ограничиваться линейными полиномами, а если они недостаточно точны, полиномами второй степени (квадратичными). Дальнейшее увеличение степени полинома ведет, как правило, только к увеличению сложности прогнозирования и не больше.

1.5 Принятие решений перед планированием

Подытоживая все выше сказанное, отмечу, что прежде чем заниматься планированием эксперимента, необходимо определиться с некоторыми вопросами.

- **І.** Во-первых, следует точно определиться с понятием объекта исследования, дав ему точное формальное определение.
- **II.** Во-вторых, прежде чем приступать к эксперименту, необходимо однозначно и непротиворечиво сформулировать основную цель эксперимента, определиться с параметром оптимизации. Параметр оптимизации должен быть единственным, хотя он и может принимать различные значения.
- III. В-третьих, необходимо определиться с факторами, влияющими на ход эксперимента и с тем, какие значения принимают эти факторы. Влияющих факторов, вообще говоря, может быть сколько угодно, при этом каждый из них может принимать бесконечное число значений. Однако не следует забывать, что в зависимости от числа факторов и их уровней катастрофически растет и число экспериментов. Выбирая, скажем, порядка двадцати факторов, каждый из которых имеет, например, по два уровня, мы можем обречь себя на долгие годы «мучений».
- **IV.** В-четвертых, необходимо озадачиться поиском области проведения эксперимента. И здесь должны учитываться следующие соображения.
 - **1.** Прежде всего необходимо оценить границы областей определения факторов. При выборе границ учитываются ограничения нескольких типов:
 - *а)* принципиальные ограничения для значений факторов, которые ни при каких условиях не могут быть нарушены. Например, температура никак не может по значению оказаться ниже абсолютного нуля;
 - **b)** технико-экономические ограничения. Например, стоимость сырья, дефицитность отдельных компонентов, время протекания процесса;
 - *с)* конкретные условия проведения процесса наиболее часто встречающийся тип ограничений. Например, существование аппаратуры, стадия разработки технологии и т.п.

Таким образом, выбор экспериментальной области факторного пространства связан с тщательным анализом априорной 1 информации.

- 2. На втором этапе необходимо найти локальную область для планирования эксперимента. Данная процедура включает в себя два этапа:
 а) выбор основного уровня. Наилучшим условиям, определенным из анализа априорной информации, соответствует одна или несколько комбинаций уровней факторов. Каждая комбинация является многомерной точкой в факторном пространстве. Ее можно рассматривать как исходную точку для построения плана эксперимента. Такая точка называется основным или нулевым уровнем. Построение плана сводится к выбору точек, симметричных относительно основной. В разных условиях мы обладаем различной информацией об области наилучших условий. Выбор основной точки легко представить в виде схемы, рисунок 1.5.
 - *b)* Выбрав основной уровень, необходимо провести выбор интервалов варьирования. Необходимо выбрать два уровня, желательно симметричных относительно основного, которые называют верхним и нижним уровнями. Обычно за верхний уровень принимается тот, который соответствует наибольшему значению фактора, хотя данное требование и не является обязательным.

Интервалом варьирования факторов называется некоторое число (свое для каждого фактора), прибавление которого к основному уровню дает верхний уровень, а вычитание — нижний уровень.

¹ *Априорной* называется информация, извлеченная из результатов предшествующих опытов. Если информация берется из последующих опытов, она называется *апостериорной*.

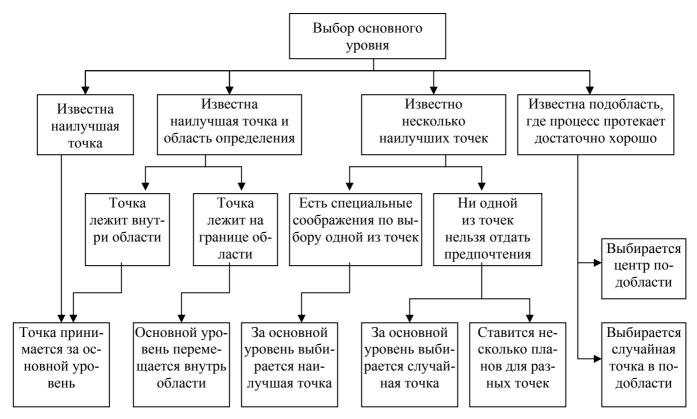


Рисунок 1.5. Схема выбора основного уровня

Для упрощения записи условий эксперимента и обработки экспериментальных данных масштабы по осям выбираются таким образом, чтобы верхний уровень соответствовал (+1), нижний (-1), а основной – нулю. Это всегда можно сделать с помощью преобразования

$$\widetilde{\mathbf{x}}_{i} = \frac{\mathbf{x}_{i} - \mathbf{x}_{0}}{\mathbf{m}_{i}},$$

где $\widetilde{\mathbf{x}}_{\mathrm{i}}$ – кодированное значение фактора,

хі – истинное значение фактора,

 x_0 – истинное значение нулевого уровня,

 m_i – интервал варьирования,

і – номер фактора.

Для качественных факторов, имеющих два уровня, один уровень обозначается (+1), другой (-1); порядок значения не имеет. Графически процедуру перекодировки можно представить как смещение осей факторного пространства в центр области проведения эксперимента, рисунок 1.6 (представлен случай двух факторов).

На выбор интервалов варьирования накладываются естественные *ограничения* сверху и снизу Интервал варьирования не может быть меньше ошибки фиксирования уровня фактора, иначе верхний и нижний уровни ока-

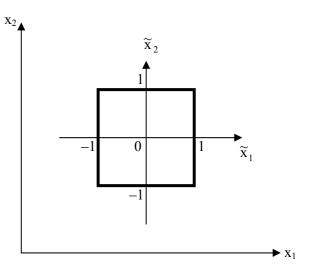


Рисунок 1.6. Графическое изображение процедуры кодировки факторов

жутся неразличимыми. С другой стороны, интервал варьирования не может быть настолько большим, чтобы выйти за пределы области определения. Кроме того, выбор интервалов варьирования напрямую зависит от информации относительно кривизны поверхности отклика и о диапазоне изменения параметра оптимизации. В зависимости от этих трех условий выбор интервалов варьирования будет различным.

V. И, наконец, необходимо помнить, что для грамотного исследователя является главной целью не поиск материальных благ, приобретаемых при оптимизации процесса, а построение математической модели объекта исследования, представляющей собой математической уравнение, связывающее параметр оптимизации и факторы, т.е. функции отклика. Наличие функции отклика «под рукой» поможет в дальнейшем решать новые задачи с наименьшими затратами по исследованию объекта.

2 Статистическая проверка статистических гипотез

2.1 Статистические гипотезы. Виды ошибок при выдвижении статистических гипотез

Часто необходимо знать закон распределения генеральной совокупности. Если имеются основания предположить, что он имеет определенный вид, то выдвигают гипотезу: генеральная совокупность распределена по данному закону. Таким образом, в этой гипотезе идет речь о виде предполагаемого распределения.

Другой случай, — когда закон распределения известен, но неизвестны его параметры (среднее, дисперсия). Если есть основания предполагать, что неизвестный параметр Θ равен определенному значению Θ_0 , , выдвигают гипотезу $\Theta = \Theta_0$. Таким образом, в этой гипотезе идет речь о предполагаемой величине параметра известного распределения.

Приведенные примеры представляют собой одни из многочисленных вариантов статистических гипотез. Таким образом,

статистической гипотезой называют гипотезу о виде неизвестного распределения или о параметрах известных распределений.

Наряду с первоначально выдвинутой гипотезой рассматривают и противоречащую ей. Если выдвинутая гипотеза будет отвергнута, ее место занимает противоречащая.

Нулевой (основной) гипотезой называют первоначально выдвинутую гипотезу. Гипотезу, противоречащую нулевой, называют конкурирующей (альтернативной) гипотезой.

Условно нулевую гипотезу обозначают H_0 , а альтернативную — H_1 . Приведу примеры обозначений статистических гипотез и варианты их прочтения:

 H_0 : $\bar{x} = 15$ - основная гипотеза состоит в том, что среднее значение случайной величины X *статистически неразличимо* с 15;

 H_1 : $\overline{x} > 15$ - альтернативная гипотеза состоит в том, что среднее значение случайной величины X *статистическим различимо* и больше 15.

ОБРАТИТЕ ВНИМАНИЕ: о равенстве показателей речи не идет. Корректно говорить «статистически неразличимо» или «статистически различимо» мо»

Когда выдвигается гипотеза, всегда существует вероятность, что она может быть правильной или неправильной, поэтому возникает необходимость ее проверки. Поскольку проверку производят статистическими методами, ее называют *статистической проверкой*.

При выдвижении гипотезы, независимо от того, статистическая она или нет, автор гипотезы берет на себя определенную ответственность. Ведь выдвинутая гипотеза, равно как и результаты ее проверки, могут быть ошибочными. Риск, который возникает при выдвижении статистической гипотезы, так и называют *ошибкой*, причем существуют ошибки I и II рода.

Ошибка I рода состоит в том, что будет отвергнута гипотеза, в то время как она верна. Ошибка I рода оценивается **уровнем значимости** α ^{*)}.

 $^{^{*)}}$ Об уровне значимости α и мощности критерия π более подробно будет сказано ниже.

Ошибка II рода состоит в том, что будет принята гипотеза, в то время как она неверна. Ошибка II рода оценивается **мощностью критерия** π .

При этом последствия таких ошибок могут оказаться весьма различными. Можно привести примеры, когда ошибка I рода влечет за собой более весомые последствия, чем ошибка II рода, и наоборот.

Пример 1. Идет прием у врача. Исследуя симптомы болезни, врач назначает лечение. Помимо лекарств, назначаемых при данных симптомах, врач выписывает некоторые анализы для подтверждения своего диагноза. При этом возможны следующие варианты:

Ошибка І рода. Назначение данных лекарств было правомерно, т.к. первоначальный диагноз оказался верным, что и подтвердили дополнительные анализа, но врач подверг первоначальный диагноз сомнению, т.е. фактически отверг его.

Ошибка II рода. Назначение данных лекарств недопустимо, т.к. первоначальный диагноз оказался неверным, что и показали дополнительные анализы, но врач назначил их в соответствии с первоначальным диагнозом, который он фактически принял.

Понятно, что в данном примере ошибка II рода приведет к более тяжким последствиям, чем ошибка I рода.

Пример 2. Стоит вопрос о замене строительных материалов, предусмотренных проектом, на другие, поскольку они более доступны и дешевы. Для этого проводится соответствующая экспертиза. При этом возможны следующие ошибки:

Ошибка I рода. Применение предлагаемых в качестве альтернативы строительных материалов невозможно, но эксперт разрешает их использование, т.к. считает их технические характеристики соответствующими нормам.

Ошибка II рода. Применение альтернативных материалов возможно, но эксперт запрещает замену.

В данном случае ошибка II рода менее тяжела, чем ошибка I рода.

Когда экспериментатор выдвигает ту или иную статистическую гипотезу, он предполагает, что может совершить ошибку. Решение, принимаемое экспериментатором должно иметь альтернативу, т.е. экспериментатор помимо выдвижения гипотезы должен держать наготове ответ на вопрос: «А что, если Вы ошиблись?» Про такую ситуацию говорят, что экспериментатор закладывает в гипотезу ошибку того или иного рода. Ошибку какого рода заложить в свою гипотезу экспериментатор решает в зависимости от тяжести последствий при совершении ошибки.

Чтобы было более-менее понятно, о чем идет речь, проиллюстрируем данные соображения на примере приема у врача, описанного выше. Фактически, назначая дополнительные анализы для подтверждения диагноза, врач закладывает в свою гипотезу ошибку первого рода, т.е. первоначальный диагноз может оказаться верным, но врач не верит без дополнительной проверки. Кстати, из этих же соображений врач первоначально, пока не выяснит окончательно диагноз, назначает лекарства лишь облегчающие симптоматику, но не решающие все кардинально. Согласитесь, в этом примере ошибка первого рода несет наименьшие последствия, и врач поступает правильно.

Хотя никто ему не запрещает заложить в свое решение ошибку второго рода: в этом случае врач начинает кардинальное лечение заболевания, рискуя назначить неправильное лечение при ошибочном первоначальном диагнозе. Думается, такого рода примеры известны из Вашей практики или практики Ваших знакомых.

2.2 Статистические критерии

Когда любой из нас проводит проверку чего-либо, принимает какоелибо решение, всем бывает необходим критерий соответствия полученного результата ожиданиям, тем или иным требованиям и т.д. Например, при покупке дивана человек оценивает его на соответствие многим критериям: габариты, цвет, форма... Точно также обстоит дело и в статистике. Только в данном случае необходимы критерии для проверки соответствия выдвинутой статистической гипотезы реальному положению дел. И критерии, соответственно, должны быть статистические.

Статистическим критерием (или просто критерием, критерием согласия) называют критерий проверки гипотезы о предполагаемом законе распределения случайной величины или значениях параметров распределений случайной величины.

При этом значение критерия, вычисленное по экспериментальным данным, называют *наблюдаемым значением критерия* $K_{\text{набл}}$.

Статистические критерии работают на всем множестве значений числовой прямой в пределах ($-\infty$; $+\infty$). При этом вся эта числовая прямая делится на два типа подобластей: критическую и область принятия гипотезы (решения).

Критической областью называют совокупность значений критерия, при которых нулевую гипотезу отвергают.

Областью принятия гипотезы (областью допустимых значений) называют совокупность значений критерия, при которых нулевую гипотезу принимают.

Критическими точками (границами) $k_{\kappa p}$ называют точки, отделяющие критическую область от области принятии решений.

Различают одностороннюю и двустороннюю критические области. Первая, в свою очередь, делится на правостороннюю и левостороннюю.

Правосторонней критической областью называют критическую область, определяемую неравенством $K_{\text{набл}} > k_{\kappa p}$, где $k_{\kappa p}$ – положительное число (см. рисунок 2.1,а).

Левосторонней критической областью называют критическую область, определяемую неравенством $K_{\text{набл}} < k_{\text{кр}}$, где $k_{\text{кр}}$ – отрицательное число (см. рисунок 2.1,б).

Двусторонней критической областью называют критическую область, определяемую неравенствами $K_{\text{набл}} < k_{\text{кр.1}}$ и $K_{\text{набл}} > k_{\text{кр.2}}$, где $k_{\text{кр.2}} > k_{\text{кр.1}}$ (см. рисунок 2.1,в).

Если критические точки двустороннего критерия выбирать симметрично, то определение двусторонней критической области перепишется как $|K_{\text{набл}}| > k_{\kappa p}$ (см. рисунок 2.1,г).

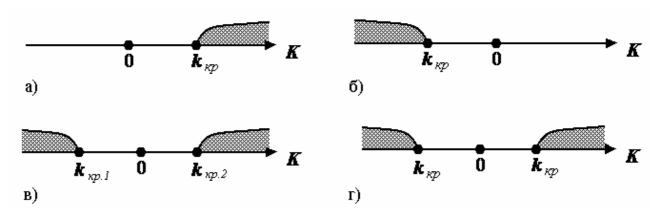


Рисунок 2.1. Графическое отображение критических областей. Штриховкой показаны критические области

Фактически, как видно из определений выше, нам необходимо какимлибо способом определить расположение двух выше названных подобластей, посмотреть в какую из них попало наблюдаемое значение критерия К_{набл}, и на основе таких простейших сравнений определиться принять основную гипотезу или нет. В общем случае алгоритм проверки основной гипотезы с помощью критериев согласия примет вид, рисунок 2.2.

С логической и понятийной точки зрения все достаточно просто. Но с практической позиции сразу же возникает вполне естественный вопрос, как отыскать критическую точку? Для ее отыскания задаются достаточно малой вероятностью – уровнем значимости α .

Уровнем значимости Q называют вероятность, при которой событие (в данной определенной задаче) практически невозможно, т.е. это вероятность того, что исследуемое событие при данных условиях не произойдет. С точки зрения проверки статистических гипотез, уровень значимости — вероятность того, что наблюдаемое значение критерия попадет в критическую область:

$$P(K_{\text{набл}} > k_{\kappa p}) = \alpha.$$

Вероятность того, что наблюдаемое значение критерия попадет в область допустимых значений называют доверительной вероятностью (надежностью)

$$P = 1 - \alpha$$

С общих позиций, надежностью называют вероятность того, что имеет место описываемое событие.



Рисунок 2.2. Общий алгоритм подтверждения основной гипотезы с помощью критериев согласия

Фактически, экспериментатор сам определяет ту степень вероятности, с которой данное событие, а в нашем случае — это выдвинутая гипотеза, не произойдет, т.е., попросту говоря, какова вероятность того, что экспериментатор ошибся, выдвинув свою гипотезу.

Задав уровень значимости, экспериментатор получает возможность найти критическую точку. Дело в том, что все статистические критерии (или критерии согласия) основываются на различных известных в статистике распределениях: распределении Пирсона, Фишера, Стьюдента и т.д. Для всех этих распределений уже давно рассчитаны так называемые критические значения, которые представляют собой *квантили**) упомянутых распределений.

Здесь необходимо сделать одно небольшое замечание.

В случае односторонних областей выбор критической точки определяется требованием

$$P(K_{\text{набл}} > k_{\text{кp}}) = \alpha - \text{при правостороннем критерии}$$
 или
$$P(K_{\text{набл}} < k_{\text{кp}}) = \alpha - \text{при левостороннем критерии}.$$

Однако, в случае двусторонней критической области данное условие примет вид

$$P(K_{\text{набл}} < k_{\text{кр.1}}) + P(K_{\text{набл}} > k_{\text{кр.2}}) = \alpha.$$

Ясно, что критические точки в этой ситуации могут быть выбраны бесчисленным множеством способов. Однако, как правило, критические точки стараются выбрать симметричными относительно нуля. Тогда

$$P(K_{\text{набл}} > k_{\text{кр}}) = P(K_{\text{набл}} < -k_{\text{кр}}),$$

 $^{^{*)}}$ Кратко о распределениях случайных величин, их параметрах и квантилях можно прочитать в **Приложении A**.

и критерий примет вид

$$P(K_{\text{набл}} > k_{\text{кр}}) = \alpha / 2.$$

– Хорошо, – скажете Вы, – с этим понятно. (Хотя на самом деле ничего не понятно). А как определиться с видом критической области: двусторонняя, левосторонняя или правосторонняя?

На самом деле здесь все еще проще. Вид критической области зависит от вида альтернативной гипотезы. Для простоты представим пример выбора критической области в виде стилизованной таблицы, таблица 2.1.

Таблица 2.1 – Алгоритм выбора вида критической области

Основная гипотеза:			
H_0 : $\overline{x}=a$ Альтернативная гипотеза:			
H_1 : $\overline{x} \neq a$	Двусторонняя (критические точки ищем при уровне значимости $\alpha/2$)		
H_1 : $\overline{x} < a$	Левосторонняя (критические точки ищем при уровне значимости α)		
H_1 : $\overline{x} > a$	Правосторонняя (критические точки ищем при уровне значимости α)		

Выше уже говорилось, что при статистической проверке статистических гипотез помимо основной принимается и альтернативная ей гипотеза. Вследствие этого целесообразно ввести в рассмотрение вероятность попадания критерия в критическую область при условии, что верна альтернативная гипотеза.

Мощностью критерия π называют вероятность nonadaния критерия в критическую область npu условии, что справедлива конкурирующая гипотеза. При этом, если вероятность совершения ошибки II рода равна β , то мощность критерия определяется как

$$\pi = 1$$
- β .

Если уровень значимости уже выбран, то критическую область следует строить так, чтобы мощность критерия была максимальной. Фактически, мощность критерия – вероятность того, что ошибка второго рода не будет допущена. При этом одновременно уменьшить α и β невозможно. При уменьшении одной величины, вторая неизбежно будет возрастать. Поскольку при проверке статистических гипотез выбирается уровень значимости α, относительно него и решается вопрос о выборе значения. Величина β автоматически будет уменьшаться или возрастать при увеличении или уменьшении α. Вопрос о выборе величины уровня значимости будет напрямую зависеть от тяжести последствий, вызываемых ошибками I и II рода. Если ошибка I рода влечет за собой более тяжелые последствия, то величину α выбирают как можно меньше.

2.3 Виды критериев согласия и области их применения

Итак, с методической точки зрения все более или менее понятно. Теперь осталось разобраться с видами критериев согласия и с тем, какой из них в каком случае использовать.

Около 90-95 % людей на Земле лучше всего воспринимают всю информацию визуально. Именно поэтому, вместо долгих рассуждений о видах

критериев согласия и случаях их применения, приведу схему, рисунок 2.3. Она проста, наглядна, да и Вы не запутаетесь в критериях.

Критерии согласия носят названия по имени тех ученых-статистиков, которые их и сформулировали. Исключение из общей картины на рисунке 2.3 составляет только один инструмент — однофакторный дисперсионный анализ (ОДА). Данный инструмент <u>НЕ</u> является критерием согласия. Однако чтобы классификатор инструментов сравнения был полон, ОДА был добавлен к критериям согласия. Дополнительно замечу, что сам ОДА будет рассмотрен в следующем разделе.

Все критерии согласия рассчитаны на то, что генеральные совокупности рассматриваемых в критериях случайных величин подчиняются нормальному закону. В противном случае результаты могут быть и неправильными. Кроме того, в критериях согласия рассматриваются так называемые исправленные оценки исследуемых параметров (среднего, дисперсии).

Принцип «работы» всех критериев согласия одинаков: по определенному правилу-алгоритму находим наблюдаемое значение критерия $K_{\text{набл}}$, сравниваем его с критическим значением $k_{\text{кр}}$ распределения, задействованного в данном критерии, и выносим суждение о подтверждении или отвержении основной гипотезы. Различие состоит лишь в алгоритмах поиска $K_{\text{набл}}$ и привлечении разных распределений для поиска $k_{\text{кр}}$.

Условие подтверждения / отвержения основной гипотезы будем демонстрировать на примере двусторонней критической области, за исключением первого случая. На примере первого критерия согласия покажем как выглядят условия подтверждения основной гипотезы для всех трех типов критических областей. Поскольку во всех случаях ситуация будет одна и та же, повторяться, думаем, не имеет смысла.

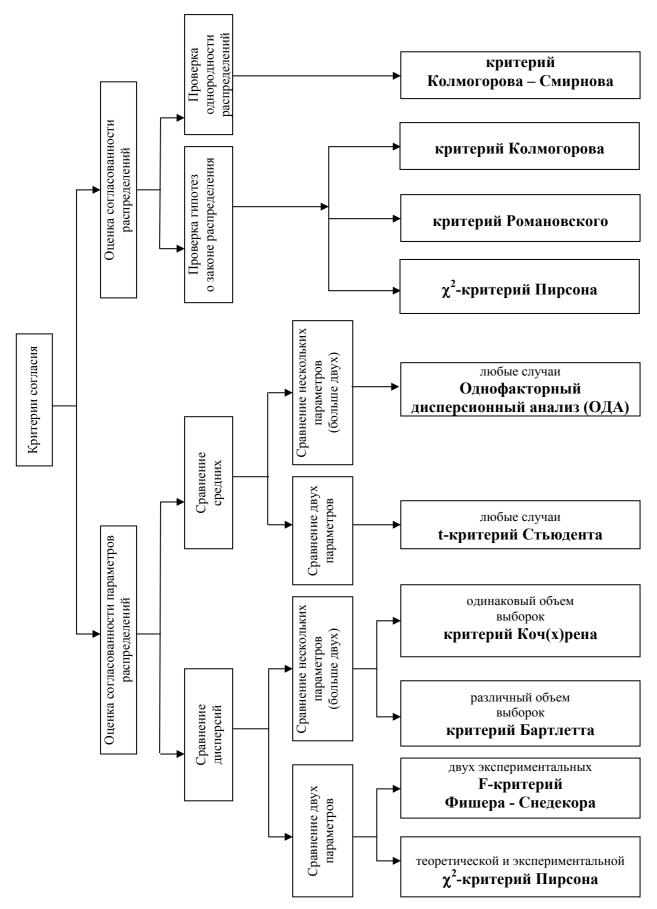


Рисунок 2.3. Классификация видов критериев согласия

$2.3.1 \chi^2$ -критерий согласия Пирсона *)

Критерий согласия Пирсона применяется для сравнения теоретического и экспериментального значений дисперсий.

В качестве теоретического значения дисперсии на практике используются значения, регламентированные какими-либо нормативными документами: ГОСТами, ТУ, техническим паспортом и т.п.

Обозначим s^2 — экспериментально полученное значение дисперсии по выборке объема n, σ^2 — теоретическое значение дисперсии. Основная гипотеза состоит в том, что данные значения дисперсий <u>статистически неразличимы;</u> в краткой записи наше предположение выглядит как

$$H_0: s^2 = \sigma^2.$$

При этом альтернативная гипотеза состоит в том, что

- 1) H_1 : $s^2 \neq \sigma^2$ экспериментальное и теоретическое значения дисперсий статистически различимы двусторонняя критическая область;
- 2) H_1 : $s^2 < \sigma^2$ теоретическое значения дисперсии превышает экспериментальное *левосторонняя критическая область*;
- 3) H_1 : $s^2 > \sigma^2$ теоретическое значения дисперсии меньше экспериментального *правосторонняя критическая область*.

Наблюдаемое значение χ^2 -критерия согласия Пирсона определяется по формуле:

$$K_{\chi^2} = \frac{(n-1) \cdot s^2}{\sigma^2}$$
.

^{*)} Следует читать «хи-квадрат критерий согласия...»

Критическая точка определяется как критическое значение χ^2 распределения Пирсона при заданном уровне значимости α (для двусторонней критической области — $\alpha/2$) с числом степеней свободы (n — 1). Все сказанное укладывается в следующее обозначение:

$$\chi_{\alpha}^{2}(n-1)$$
.

Основная гипотеза подтверждается, если:

- 1) двусторонняя критическая область $\left| K_{\chi^2} \right| < \chi^2_{\alpha/2} (n-1);$
- 2) левосторонняя критическая область $K_{\chi^2} > \chi_{\alpha}^2 (n-1);$
- 3) правосторонняя критическая область $K_{\chi^2} < \chi_{\alpha}^2 (n-1).$

2.3.2 Г-критерий согласия Фишера – Снедекора

Данный критерий согласия применяется для сравнения двух экспериментальных значений дисперсий.

Обозначим:

- s_1^2 экспериментальное значение дисперсии, полученное по выборке объема n_1 в первой серии опытов;
- s_2^2 экспериментальное значение дисперсии, полученное по выборке объема n_2 во второй серии опытов.

Причем, $s_1^2 > s_2^2$.

Основная и альтернативная гипотезы имеют вид:

$$H_0: s_1^2 = s_2^2$$
 $H_1: s_1^2 \neq s_2^2$.

Наблюдаемое значение F-критерия согласия Фишера определяется по ϕ ормуле * :

$$K_F = \frac{s_1^2}{s_2^2}$$
.

Критическая точка определяется как критическое значение Fраспределения Фишера при заданном уровне значимости (или $\alpha/2$) с числами степеней свободы $(n_1-1; n_2-1)$:

$$F_{\alpha}(n_1-1;n_2-1)$$
.

При определении критического значения следует помнить, что первым в скобках стоит значение числа степеней свободы для той дисперсии, которая находится в числителе формулы наблюдаемого значения критерия.

Основная гипотеза подтверждается, если

$$|K_F| < F_{\alpha/2}(n_1 - 1; n_2 - 1).$$

2.3.3 Критерий согласия Бартлетта

Заключается в сравнении нескольких дисперсий (больше двух) по выборкам различного объема. Главное условие применения критерия согласия Бартлетта – объем выборок должен быть не менее 4 испытаний.

Обозначим:

 $^{^{*)}}$ ОБРАТИТЕ ВНИМАНИЕ: в числителе всегда стоит то значение дисперсии, которое больше с математической точки зрения. Если бы было выдвинуто условие, что $S_1^2 < S_2^2$, то в формуле наблюдаемого значения критерия дробь перевернулась бы.

- s_1^2 экспериментальное значение дисперсии, полученное по выборке объема n_1 в первой серии опытов;
- s_2^2 экспериментальное значение дисперсии, полученное по выборке объема n_2 во второй серии опытов;

. .

 s_i^2 — экспериментальное значение дисперсии, полученное по выборке объема n_i в і-серии опытов.

При этом, некоторые объемы могут быть одинаковыми; если же все выборки имеют одинаковый объем, то предпочтительнее пользоваться критерием Коч(х)рена, описанном ниже.

Основная гипотеза имеет вид:

$$H_0$$
: $s_1^2 = s_2^2 = ... = s_i^2$.

Следует понимать, что формулировка альтернативной гипотезы в виде математического соотношения достаточно проблематична, т.к. отдельные значения дисперсий могут и совпадать между собой. Однако основная гипотеза состоит в статистической неразличимости **BCEX** значений дисперсий, и проверка будет состоять в оценке выполнимости именно этого требования. Соответственно, альтернативная гипотеза будет состоять в том, что основная гипотеза не выполняется. Если же вдруг встанет вопрос о попарном сравнении, то лучше воспользоваться критерием Фишера — Снедекора.

Обозначим через \bar{s}^2 среднюю арифметическую экспериментальных дисперсий, взвешенную по числам степеней свободы:

$$\bar{s}^{2} = \frac{\sum_{m=1}^{1} k_{m} s_{m}^{2}}{\sum_{m=1}^{i} k_{m}},$$

где $k_i = (n_i - 1)$ – число степеней свободы і-серии опытов.

В качестве критерия проверки основной гипотезы о статистической неразличимости (или однородности) дисперсий, т.е. наблюдаемого значения критерия, принимается случайная величина

$$B = \frac{V}{C}$$
,

где
$$V = 2,303 \cdot \left[k \lg \overline{s}^2 - \sum_{m=1}^{i} k_m \lg s_m^2 \right],$$

$$C = 1 + \frac{1}{3(i-1)} \left[\sum_{m=1}^{i} \frac{1}{k_m} - \frac{1}{\sum_{m=1}^{i} k_m} \right].$$

Бартлетт установил, что случайная величина **B** при условии справедливости нулевой гипотезы распределена приближенно как χ^2 с (i-1) степенями свободы, если все $k_i > 2$. Исходя из данного факта, критическую точку для проверки основной гипотезы определяют по таблицам критических значений χ^2 -распределения Пирсона при уровне значимости α с числом степеней свободы (i-1): $\chi^2_{\alpha}(i-1)$.

Основная гипотеза подтверждается, если

$$B < \chi_{\alpha}^2 (i-1)$$
.

Замечание. Не следует торопиться вычислять постоянную ${f C}$. Сначала надо найти ${f V}$ и сравнить с $\chi^2_{\alpha}(i-1)$. Если окажется, что $V<\chi^2_{\alpha}(i-1)$, то $B<\chi^2_{\alpha}(i-1)$ и подавно, т.к. C>1, и, следовательно, ${f C}$ вычислять не нужно.

Если же $V > \chi_{\alpha}^2 (i-1)$, то в этом случае вычисляем C и затем уже следуем стандартной процедуре, описанной выше.

2.3.4 Критерий Коч(х)рена*)

Применяется при сравнении нескольких дисперсий (больше двух) по выборкам одинакового объема.

Обозначим:

- s_1^2 экспериментальное значение дисперсии, полученное в первой серии опытов;
- s_2^2 экспериментальное значение дисперсии, полученное во второй серии опытов;

 s_{i}^{2} — экспериментальное значение дисперсии, полученное в і-серии опытов.

Причем, все серии опытов проводились с одинаковым числом испытаний n, т.е. для всех случаев имеем одинаковое число степеней свободы k=n-1.

Как и в предыдущем случае, основная гипотеза состоит в статистической неразличимости BCEX значений дисперсий

$$H_0: s_1^2 = s_2^2 = ... = s_i^2,$$

и проверка будет состоять в оценке выполнимости именно этого требования. Соответственно, альтернативная гипотеза будет состоять в том, что основная гипотеза не выполняется.

^{*)} По-английски фамилия данного ученого пишется Cochren. В соответствии с этим, при переводе различные литературные источники выдают либо Кохрен либо Кочрен. Поэтому в названии приведены оба варианта сразу.

В рассматриваемом случае выборок одинакового объема можно сравнить по критерию Фишера — Снедекора наибольшую и наименьшую дисперсии; если окажется, что различие между ними незначимо, то подавно незначимо и различие между остальными дисперсиями. Недостаток такого подхода состоит в том, что информация, которую содержат остальные дисперсии, кроме наименьшей и наибольшей, не учитывается.

В качестве наблюдаемого значения критерия, т.е. критерия проверки нулевой гипотезы, принимается случайная величина:

$$G_{\text{набл}} = \frac{s_{\text{max}}^2}{\left(s_1^2 + s_2^2 + ... + s_i^2\right)}.$$

Распределение этой случайной величины зависит только от числа степеней свободы k=n-1 и количества выборок i.

Критическую точку $k_{\kappa p}$ находят по таблицам критических значений распределения Коч(х)рена при заданном уровне значимости α с числом степеней свободы k и количестве выборок i: $G_{\alpha}(k,i)$.

Основная гипотеза подтверждается, если

$$G_{\text{Haff}} < G_{\alpha}(k;i)$$
.

Замечание. Если требуется оценить генеральную дисперсию, то при условии однородности дисперсий целесообразно принять в качестве ее оценки среднюю арифметическую выборочных дисперсий.

2.3.5 t-критерий Стьюдента

Критерий согласия Стьюдента применяется для сравнения двух средних значений. Причем, один и тот же критерий применяется как для сравне-

ния экспериментального и теоретического значений средних, так и для сравнение двух экспериментальных средних. Различие состоит в последовательности выполнения действий и определения наблюдаемых значений критерия.

При сравнении экспериментального и теоретического средних мы приступаем к вычислению наблюдаемого значения критерия сразу, а при сравнении двух экспериментальных средних - необходимо сначала определиться со статистической неразличимостью ДИСПЕРСИЙ рассматриваемых распределений, т.к. вид наблюдаемого значения критерия будет меняться. В дальнейшем рассмотрение критерия согласия Стьюдента будет проходить по следующей схеме, рисунок 2.4.

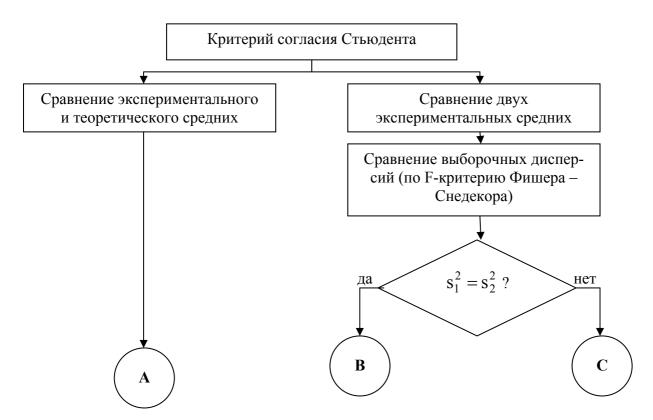


Рисунок 2.4. Схема рассмотрения критерия Стьюдента (A, B, C – порядковые пункты дальнейшего изложения)

Замечание. Поскольку распределение Стьюдента симметричное относительно нуля, обычно при построении критериев используют либо двустороннюю, либо правостороннюю критическую область.

А. Сравнение экспериментального и теоретического средних

Обозначим:

 \overline{x} — значение среднего, полученное в ходе эксперимента по выборке объема n,

 η — значение среднего, задаваемое нормативной документацией, т.е. теоретическое среднее.

Основная и альтернативная гипотезы имеют вид:

$$H_0: \overline{x} = \eta$$
 $H_1: \overline{x} \neq \eta$.

При построении наблюдаемого значения критерия возможно два варианта.

Вариант 1. Определена выборочная дисперсия s^2 . Тогда наблюдаемое значение критерия примет вид

$$K_{t} = \frac{\left|\overline{x} - \eta\right|}{\sqrt{s^{2}/(n-1)}}.$$

Критическую точку находим по таблицам критических значений распределения Стьюдента при заданном уровне значимости α (или $\alpha/2$) с числом степеней свободы (n-1): $t_{\alpha}(n-1)$.

Основная гипотеза подтверждается, если

$$K_{t} < t_{\alpha/2}(n-1).$$

Вариант 2. Известна дисперсия генеральной совокупности σ^2 , т.е. известно значение теоретической дисперсии. Тогда наблюдаемое значение критерия примет вид

$$K_{t} = \frac{\left| \overline{x} - \eta \right|}{\sqrt{\sigma^{2}/n}}.$$

В данном случае критическую точку находим из **функции Лапласа**. Нахождение критической точки $\mathbf{u}_{\kappa p}$ строим из предположения, что

$$\Phi(u_{\kappa p}) = \frac{1-\alpha}{2}$$
.

Основная гипотеза подтверждается, если

$$K_t < u_{\kappa p}$$
.

Понятно, что на практике чаще всего встречается первый вариант сравнения экспериментального и теоретического средних.

В. Сравнение двух экспериментальных средних при одинаковых выборочных дисперсиях

Обозначим:

 \overline{x}_1 — значение среднего, полученное в ходе эксперимента по выборке объема n_1 ,

 s_1^2 – значение дисперсии, полученное по первой выборке,

 \overline{x}_2 — значение среднего, полученное в ходе эксперимента по выборке объема n_2 ,

 s_{2}^{2} — значение дисперсии, полученное по второй выборке, причем $s_{1}^{2}>s_{2}^{2}$.

Проверка двух экспериментальных средних проводится в два этапа: на первом этапе проводится сравнение выборочных дисперсий по <u>F-критерию</u>

<u>согласия Фишера – Снедекора</u> (см. п. 2.3.2), а на втором этапе – уже непосредственно сравнение средних.

1 этап. Проводим сравнение выборочных дисперсий по критерию согласия Фишера – Снедекора. Строим основную и альтернативную гипотезы:

$$H_0: s_1^2 = s_2^2$$
 $H_1: s_1^2 \neq s_2^2$.

Наблюдаемое значение F-критерия согласия Фишера определяется по формуле:

$$K_F = \frac{s_1^2}{s_2^2}$$
.

Поскольку критическая область — двусторонняя, критическое значение распределения Фишера — Снедекора ищем при уровне значимости $\alpha/2$.

Пусть в ходе проверки гипотезы оказалось, что

$$|K_F| < F_{\alpha/2}(n_1 - 1; n_2 - 1).$$

Тогда, основная гипотеза о статистической неразличимости дисперсий обоих распределений – справедлива. В этом случае можно построить обобщенную оценку дисперсии

$$s_{\Sigma}^{2} = \frac{k_{1} \cdot s_{1}^{2} + k_{2} \cdot s_{2}^{2}}{k_{1} + k_{2}},$$

где $k_1 = (n_1 - 1)$ – число степеней свободы первого распределения, $k_2 = (n_2 - 1)$ – число степеней свободы второго распределения.

Теперь можно перейти к сравнению самих средних.

2 этап. Проводим сравнение экспериментальных средних.

Основная и альтернативная гипотезы имеют вид:

$$H_0: \overline{x}_1 = \overline{x}_2$$
 $H_1: \overline{x}_1 \neq \overline{x}_2$.

Наблюдаемое значение критерия будет определяться как

$$K_{t} = \frac{\left|\overline{x}_{1} - \overline{x}_{2}\right|}{\sqrt{s_{\Sigma}^{2} \cdot \left(\frac{1}{n_{1}} + \frac{1}{n_{2}}\right)}}.$$

Критическую точку находим по таблицам критических значений распределения Стьюдента при заданном уровне значимости α (или $\alpha/2$) с числом степеней свободы (n_1+n_2-2) : $t_{\alpha}(n_1+n_2-2)$.

Основная гипотеза подтверждается, если

$$K_t < t_{\alpha/2} (n_1 + n_2 - 2).$$

С. Сравнение двух экспериментальных средних при различных выборочных дисперсиях

Обозначим:

 \overline{x}_1 — значение среднего, полученное в ходе эксперимента по выборке объема $n_1,$

 s_1^2 – значение дисперсии, полученное по первой выборке,

 \overline{x}_2 — значение среднего, полученное в ходе эксперимента по выборке объема n_2 ,

 s_2^2 — значение дисперсии, полученное по второй выборке, причем $s_1^2 > s_2^2$.

Как и в предыдущем случае, проверка двух экспериментальных средних проводится в два этапа.

1 этап. Проводим сравнение выборочных дисперсий по критерию согласия Фишера – Снедекора. Строим основную и альтернативную гипотезы:

$$H_0: s_1^2 = s_2^2$$
 $H_1: s_1^2 \neq s_2^2$.

Наблюдаемое значение F-критерия согласия Фишера определяется по формуле:

$$K_F = \frac{s_1^2}{s_2^2}$$
.

Поскольку критическая область — двусторонняя, критическое значение распределения Фишера — Снедекора ищем при уровне значимости $\alpha/2$.

Пусть в ходе проверки гипотезы оказалось, что

$$|K_F| > F_{\alpha/2}(n_1 - 1; n_2 - 1).$$

Тогда, основная гипотеза о статистической неразличимости дисперсий обоих распределений отвергается в пользу альтернативной. В этом случае построить обобщенную оценку дисперсии нельзя и необходимо сразу перейти к сравнению самих средних.

2 э*тап.* Проводим сравнение экспериментальных средних.

Основная и альтернативная гипотезы имеют вид:

$$H_0: \overline{x}_1 = \overline{x}_2$$
 $H_1: \overline{x}_1 \neq \overline{x}_2$.

Наблюдаемое значение критерия будет определяться как

$$K_t = \frac{\left|\overline{x}_1 - \overline{x}_2\right|}{\sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}}.$$

Критическую точку находим по таблицам критических значений распределения Стьюдента при заданном уровне значимости α (или $\alpha/2$) с числом степеней свободы (n_1+n_2-2) : $t_{\alpha}(n_1+n_2-2)$.

Основная гипотеза подтверждается, если

$$K_t < t_{\alpha/2}(n_1 + n_2 - 2).$$

Ранее уже говорилось, что критерии согласия служат не только для оценки параметров распределений, но также и для оценки того, какому закону распределения подчиняется полученная экспериментальная выборка.

Строго говоря, большинство статистических инструментов корректно только в том случае, если закон распределения – нормальный или стремящийся к нормальному (это происходит с большинством известных законов распределения – Пирсона, Фишера, Стьюдента и т.д. – при определенном числе испытаний). Поэтому, прежде чем применять какие-либо методики оценки или планирования, грамотный экспериментатор сначала должен убедиться, а будет ли работать штатная методика, т.е. подчиняется ли экспериментальное распределение нормальному закону, или же предстоит придумывать другие пути.

Описанные ниже критерии согласия как раз и позволяют проводить проверку гипотез о виде распределения, т.е. сравнение эмпирической (или экспериментальной) и теоретической кривых.

Отдельно следует отметить, что при проверке гипотез о законе распределения контролируется лишь ошибка первого рода, оценить вероятность совершения ошибки второго рода — невозможно.

$2.3.6~\chi^2$ -критерий Пирсона для сравнения эмпирической и теоретической кривых

Известный английский статистик К. Пирсон в 1900 г. предложил для оценки расхождения между эмпирическими и теоретическими частотами критерий, который основан на распределении χ^2 . На сегодняшний день критерий согласия Пирсона является одним из наиболее часто используемых для сравнения экспериментальной и теоретической кривых.

В качестве меры расхождения между двумя кривыми берется величина К, равная:

$$K = \sum_{i=1}^{m} \frac{(n_i - np_i)^2}{np_i},$$

где n_i – эмпирические частоты значений измеряемой величины;

 p_i — вероятности появления значений измеряемой величины, рассчитанные по предполагаемому теоретическому закону распределения;

n – общее число проведенных экспериментов;

m — число интервалов, в которые сгруппированы экспериментальные данные.

Пирсон доказал, что статистика, определенная выше указанной формулой, имеет χ^2 -распределение с числом степеней свободы k=m-r-1, где r- число параметров теоретического распределения.

Замечание 1. Чаще всего, теоретические распределения описываются двумя параметрами — средним (т.е. математическим ожиданием) и дисперсией. Но бывают случаи, когда необходимо привлечь большее число параметров, таких как асимметрия, эксцесс. Подобные случаи оговариваются отдельно в справочниках по законам распределений случайных величин.

Основная гипотеза при проверке состоит в том, что расхождение между эмпирическим и теоретическим распределениями невелико, т.е.

$$H_0: K = 0$$
.

Основная гипотеза отвергается, если

$$K > \chi_{\alpha}^2 (m-r-1),$$

т.е. эмпирическая и теоретическая кривые не совпадают. Для экспериментатора подобный результат означает, что он неправильно подобрал вид теоретической кривой для описания экспериментальных данных. С точки зрения теории планирования эксперимента, это может означать, что построенная модель (теоретическая кривая) будет неправильно предсказывать поведение параметра оптимизации, т.е. не будет удовлетворять требованию адекватности модели эксперимента.

Замечание 2. При использовании данного критерия необходимо, чтобы при группировании экспериментальных данных в каждом интервале было не менее пяти значений. Если же в каком-либо интервале это требование не будет удовлетворяться, следует объединить соседние интервалы, чтобы число частот в интервале было не менее пяти. При этом в качестве величины m будет выбираться число объединенных интервалов.

2.3.7 Критерий Романовского

В.И. Романовский предложил использовать критерий χ^2 в другом виде. Сначала вычисляется величина

$$R = \frac{\left|\chi^2 - K\right|}{\sqrt{2K}},$$

где K = (m-3) – число степеней свободы, m – число групп.

В этом случае, если R < 3, то расхождение между эмпирическим и теоретическим распределениями считается несущественным, и принятый закон распределения можно принять в качестве модели эмпирического распределения.

Если же R > 3, то расхождение между распределениями существенно.

Отношение Романовского основывается на том, что математическое ожидание χ^2 равно числу степеней свободы K, а дисперсия — удвоенному числу степеней свободы (2K). В этом случае вероятность отклонения величины χ^2 на $3\sigma = 3\sqrt{2K}$ близка к единице.

2.3.8 Критерий Колмогорова

На практике для сравнения эмпирической и теоретической кривых кроме критерия χ^2 часто используется критерий Колмогорова, в котором в качестве меры расхождения между теоретическим и эмпирическим распределениями рассматривают максимальное значение абсолютной величины разности между эмпирической функцией распределения $F_n(x)$ и соответствующей теоретической функцией распределения F(x)

$$D = \max |F_n(x) - F(x)|,$$

называемое <u>статистикой критерия Колмогорова</u>. Мера расхождения функций определяется как

$$\lambda = D\sqrt{n}$$
.

Доказано, что какова бы ни была функция распределения F(x) непрерывной случайной величины X, при неограниченном увеличении числа наблюдений $(n \to \infty)$ вероятность неравенства $P(D\sqrt{n} \ge \lambda)$ стремится к пределу

$$P(\lambda) = 1 - \sum_{k=-\infty}^{+\infty} (-1)^k e^{-2k^2\lambda^2}$$
.

Задавая уровень значимости а, из соотношения

$$P(\lambda_{\alpha}) = \alpha$$

можно найти соответствующее критическое значение λ_{α} .

Схема применения критерия Колмогорова следующая:

- 1. Строятся эмпирическая функция распределения $F_n(x)$ и предполагаемая теоретическая функция распределения F(x).
- 2. Определяется мера расхождения между теоретическим и эмпирическим распределениями D по формуле и вычисляется величина $\lambda = D\sqrt{n}$.
- 3. Если $\lambda > \lambda_{\alpha}$, определенного на уровне значимости α , то нулевая гипотеза H_0 о том, что случайная величина X имеет заданный закон распределения, отвергается.

Если $\lambda \leq \lambda_{\alpha}$, то считают , что гипотеза H_0 не противоречит опытным данным.

Критерий Колмогорова достаточно часто применяется на практике благодаря своей простоте. Однако его применение возможно лишь тогда, когда теоретическая функция распределения F(x) задана полностью. Но такой случай на практике встречается весьма редко. Обычно из теоретических сообра-

жений известен лишь вид функции распределения, а ее параметры определяются по эмпирическим данным. При применении критерия χ^2 это обстоятельство учитывается соответствующим уменьшением числа степеней свободы. Такого рода поправок в критерии Колмогорова не предусмотрено. Поэтому, если при неизвестных значениях параметров применить критерий Колмогорова, взяв за значения параметров их оценки, то получим завышенное значение вероятности $P(\lambda)$, а значит, большее критическое значение λ_{α} . В результате есть риск в ряде случаев принять нулевую гипотезу H_0 о законе распределения случайной величины как правдоподобную, в то время как на самом деле она противоречит опытным данным.

2.3.9 Критерий Колмогорова – Смирнова

Гипотезы об *однородности выборок* – это гипотезы о том, что рассматриваемые выборки извлечены из одной и той же генеральной совокупности.

Пусть имеются две независимые выборки, произведенные из генеральных совокупностей с неизвестными теоретическими функциями распределения $F_1(x)$ и $F_2(x)$. Проверяемая нулевая гипотеза имеет вид

$$H_0$$
: $F_1(x) = F_2(x)$

против конкурирующей

$$H_1: F_1(x) \neq F_2(x).$$

Будем предполагать, что функции $F_1(x)$ и $F_2(x)$ непрерывны.

Статистика критерия Колмогорова – Смирнова имеет вид:

$$\lambda = \sqrt{\frac{n_1 n_2}{n_1 + n_2}} \cdot \max |F_{n1}(x) - F_{n1}(x)|,$$

где $F_{n1}(x)$ и $F_{n2}(x)$ — эмпирические функции распределения, построенные по двум выборкам объемов n_1 и n_2 .

Гипотеза H_0 отвергается, если фактически наблюдаемое значение статистики λ больше критического $\lambda_{\text{кp}}$, т.е. $\lambda > \lambda_{\text{кp}}$, и принимается в противном случае.

При малых объемах выборок (n_1 , $n_2 \le 20$) критические значения $\lambda_{\kappa p}$ для заданных уровней значимости критерия можно найти в специальных таблицах. При n_1 , $n_2 \to \infty$ распределение статистики λ сходится к распределению Колмогорова для статистики λ (см. n.2.3.9) . Поэтому гипотеза H_0 отвергается на уровне значимости α , если фактически наблюдаемое значение λ больше критического λ_{α} , т.е. $\lambda > \lambda_{\alpha}$, и принимается в противном случае.

Отличие между данным критерием согласия и критерием Колмогорова состоит в том, что в данном случае сравниваются две экспериментальных кривых, а в критерии Колмогорова — экспериментальная и теоретическая кривые.

Следует обратить особое внимание на тот факт, что данные, прежде всего, группируются в интервалы значений, а затем только проводится их статистический анализ. Группирование данных может проводиться как по стандартизированным статистическим процедурам, так и по каким-либо практическим соображениям самого экспериментатора.

3 Статистические методы анализа данных и планирования экспериментов

Рассмотренные в данном разделе статистические методы служат для анализа данных, полученных по предварительным экспериментам, и для планирования эксперимента, а точнее — для построения математической модели функции отклика.

Дисперсионный анализ служит для отсеивания факторов, не оказывающих существенного влияния на отклик эксперимента. В первом разделе этому вопросу уделено достаточно внимания. Поэтому дополнительно актуальность данного инструмента описывать не будем.

Корреляционный анализ по своим действиям сродни дисперсионному — он устанавливает связи между случайными величинами. Вполне естественно, что результаты будут перекликаться с результатами дисперсионного анализа. Но для нас этот инструмент интересен еще одним моментом: он позволяет установить наличие связей между влияющими факторами, что очень важно для выполнения требований к совокупности факторов.

Регрессионный анализ служит для построения математической модели зависимости между несколькими параметрами. В нашем случае – для построения функции отклика.

В основе всех этих инструментов лежат рассмотренные в предыдущем разделе критерии согласия Фишера, Стьюдента, Пирсона. В случае дисперсионного анализа критерии являются ключевым инструментом, т.е. без них не будет самого анализа. А в случаях корреляционного и регрессионного анализов критерии согласия служат для оценки адекватности полученных в ходе исследования результатов. Более подробно – в описании самих инструментов.

И последнее замечание. При описании приведенных ниже методик нами были упомянуты практически все критерии согласия, кроме тех, которые предназначены для оценки закона распределения. Однако, они присутствуют во всем рассмотренном ниже материале в неявном виде.

Дело в том, что все рассматриваемые статистические инструменты предполагают, что распределения подчиняются нормальному закону. Строго говоря, не убедившись в данном факте, нельзя использовать описанный ниже инструментарий. Поэтому, прежде всего, необходимо проводить проверку на нормальность имеющихся эмпирических распределений, а затем — пользоваться описанными ниже статистическими методами анализа и планирования. Кроме того, регрессионный анализ подразумевает построение теоретической кривой распределения, а как проверить ее адекватность без описанных выше критериев?

3.1 Дисперсионный анализ

В общем случае, задачей дисперсионного анализа является выявление тех факторов, которые оказывают существенное влияние на результат эксперимента. Помимо этого. Дисперсионный анализ может применяться для сравнения средних нескольких выборок, если число выборок больше двух. Для этой цели служит однофакторный дисперсионный анализ.

В целях решения поставленных задач принимается следующее. Если дисперсии полученных значений параметра оптимизации в случае влияния факторов отличаются от дисперсий результатов в случае отсутствия влияния факторов, то такой фактор признается значимым.

Как видно из формулировки задачи, здесь используются методы проверки статистических гипотез, а именно — задача проверки двух эмпирических дисперсий. Следовательно, дисперсионный анализ базируется на проверке дисперсий по критерию Фишера.

В зависимости от того, сколько факторов принимается в рассмотрение, различают однофакторный (случай простой группировки) и многофакторный дисперсионный анализ. Частным случаем второго является двухфакторный дисперсионный анализ (случай двойной группировки).

В рамках этих двух случаев различают следующие виды дисперсионного анализа:

- однофакторный дисперсионный анализ с одинаковым числом испытаний по уровням фактора (ОДА-ОЧИ);
- однофакторный дисперсионный анализ с неодинаковым числом испытаний по уровням фактора (ОДА-НЧИ);
- двухфакторный дисперсионный анализ без повторений (ДДА-БП);
- двухфакторный дисперсионный анализ с повторениями (ДДА-П).

3.1.1. Однофакторный дисперсионный анализ с одинаковым числом испытаний на уровнях фактора (ОДА-ОЧИ)

Прежде, чем проводить сам дисперсионный анализ, необходимо определить понятия группового среднего и общего среднего. Предположим, что в ходе проведения эксперимента «подключается» некоторый фактор A, который может принимать \mathbf{p} значений. Для каждого из этих \mathbf{p} значений фактора A проводится серия опытов, в ходе которых измеряется $\mathbf{n}_{\mathbf{p}}$ результатов. Результаты, принадлежащие одному и тому же уровню фактора. Будут составлять единую группу, и таких групп будет, как Вы догадываетесь, \mathbf{p} .

Результаты эксперимента обозначим как x_{ij} — значениями ј-членов в ігруппе. При этом, ј будет изменяться от 1 до n_p (индикатор члена группы), а і изменяется от 1 до p (индикатор группы).

Тогда среднее значение каждой группы равно:

$$\bar{x}_i = \frac{\sum_{j=1}^{n_p} x_{ij}}{n_p},$$

где $\sum_{j=1}^{n_p} x_{ij}^{}$ — сумма всех значений в і-группе,

 \mathbf{n}_{p} – число всех значений і-группы.

Общее среднее равно (n – общее число проведенных экспериментов, p – число i-групп)

$$\overline{x} = \frac{\sum_{i=1}^{p} \sum_{j=1}^{n_p} x_{ij}}{n}.$$

Построим схему однофакторного дисперсионного анализа (см. таблицу 3.1), попутно объясняя ее.

Таблица 3.1 – Схема ОДА-ОЧИ

Группировка (источник вариации)		Сумма квадратов	Число степеней свободы	Оценка дисперсии
1. Общая		$Q = \sum_{i=1}^{p} \sum_{j=1}^{n_p} (x_{ij} - \overline{x})^2$	n-1	$s^2 = \frac{Q}{n-1}$
2. По факторам А (между групп		$Q_A = \sum_{i=1}^{n_p} (\overline{x}_i - \overline{x})^2$	p – 1	$s_A^2 = \frac{Q_A}{p-1}$
3. Остаточная (внутри групп	1)	$Q_R = \sum_{i=1}^p \sum_{j=1}^{n_p} (x_{ij} - \overline{x}_i + \overline{x})^2$	n-p	$s_R^2 = \frac{Q_R}{n - p}$
4. Оценка влияни фактора	Я	$K_{F} = \frac{s_{A}^{2}}{s_{R}^{2}} \leq F_{\alpha}$		
5. Оценка вли- яния от- дельных значений фактора	K _t =	$= \frac{\overline{x}_{\text{база}} - \overline{x}_{\text{остальные значения}}}{\sqrt{\frac{\sigma_{\text{база}}^2}{n_{\text{база}}} + \frac{\sigma_{\text{остальные значения}}^2}{n_{\text{остальные значения}}}}} > 1$	$t_{\alpha} (n_{\text{база}} + n_{\text{остальны}})$	езначения — 2)

1) Определим дисперсию безотносительно к значению, который принимает фактор (общая группировка).

Для этого вычислим квадраты отклонений каждого из полученных значений от общего среднего, данные квадраты просуммируем

$$Q = \sum_{i=1}^{p} \sum_{j=1}^{n_p} (x_{ij} - \overline{x})^2.$$

Выбор квадратов отклонений связан с тем, что отклонения могут принимать как положительные, так и отрицательные значения. Если рассматривать просто сумму отклонений значений от среднего, может возникнуть эффект компенсации положительного и отрицательного значений. В этом случае получаемая сумма будет либо слишком мала, либо равна 0. Более того, оценка дисперсии будет неверной.

Затем сумму квадратов Q разделим на число степеней свободы (ЧСС) данного эксперимента, определяемое как *число опытов* – 1:

$$k = n - 1.$$

2) Определим дисперсию при условии влияния фактора A, для чего находятся <u>отклонения групповых средних от общей средней</u> (если фактор действительно оказывает влияние, такие отклонения должны быть значимыми, что можно оценить с помощью дисперсий). Схема рассуждений такая же, как и в предыдущем случае. Здесь число степеней свободы будет определяться как *число значений*, *принимаемых фактором*, — 1:

$$k_p = p - 1$$
.

3) И, наконец, определяем дисперсию значений, вызываемую случайными причинами (погрешность средств измерений, влияние окружающей среды и т.п.) Данную дисперсию вычисляют, учитывая следующее. Изменение значений эксперимента может вызываться либо случайными явлениями, либо изменением значений факторов. Если «убрать» изменение значений факторов, то вариация значений эксперимента будет проявляться только за счет случайной компоненты. Таким образом, необходимо отклонение значений эксперимента от среднего в каждой группе значений факторов. Для этого вычисляются квадраты отклонений внутри каждой группы, т.е. при значении фактора, равного A_1 , оценивается отклонение значений, полученных при таких условиях эксперимента, от своего среднего $\overline{x}_{i, для} A_1$; при A_2 — отклонение значений от своего среднего $\overline{x}_{i, для} A_2$ и т.д. Далее — по схеме, таблица 3.1.

Если посмотреть в таблицу 3.1, то сразу становится понятно, что вычислить сумму квадратов в третьем случае достаточно трудоемко. Однако, здесь можно прибегнуть к маленькой хитрости. Дело в том, что в статистике доказано, что

$$Q = Q_A + Q_R.$$

Два параметра из этой суммы нам известны, так что найти недостающее Q_R будет несложно.

4) Оценим, влияет ли исследуемый фактор А на результат эксперимента. Это можно сделать с помощью критерия согласия Фишера.

При проверке следуем простой логике: если разброс значений эксперимента при изменении фактора не отличается от разброса значений эксперимента при фиксированном значении фактора (т.е. вызываемого чисто случайными причинами), то фактор А не оказывает никакого влияния на результаты.

Строим основную и альтернативную гипотезы:

$$H_0: s_A^2 = s_R^2$$
 $H_1: s_A^2 \neq s_R^2$

Согласно критерию Фишера (см. таблицу 3.1), если отношение межгрупповой дисперсии к внутригрупповой меньше квантиля распределения Фишера при заданном уровне α

$$K_F = \frac{s_A^2}{s_R^2} < F_\alpha,$$

то дисперсии считаются статистически неразличимыми, т.е. фактор А не оказывает влияния на результат эксперимента. В противном случае – дисперсии статистически различимы, и фактор А оказывает влияние на результат.

Замечание. В отличие от классического критерия согласия Фишера — Снедекора, при проверке различия между межгрупповой и внутригрупповой дисперсиями в числителе ВСЕГДА стоит межгрупповая дисперсия, даже если она с математической точки зрения меньше внутригрупповой.

5) В случае, если дисперсионный анализ обнаруживает наличие существенного влияния факторов на результат эксперимента, необходимо оценить, какой из уровней (значений) факторов оказывает наиболее существенное влияние. С этой целью при помощи критерия согласия Стьюдента производится сравнение средних значений, полученных при различных значениях уровней факторов (см. схему анализа, последнюю строку). Для сравнения одно из средних значений принимается за основное (базовое), а остальные сравниваются именно с этим значением.

Например, при исследовании влияния вида удобрения на урожайность культуры за основное среднее значение можно выбрать средний показатель урожая, снимаемого без использования удобрений; при исследовании влия-

ния различной рекламы на уровень продаж за основное значение можно выбрать средний показатель продажи до использования какой-либо рекламы и т.д.

Если значение критерия меньше квантиля распределения Стьюдента при заданном уровне значимости α (t_{α})

$$K_{t} = \frac{\overline{x}_{\text{база}} - \overline{x}_{\text{остальные значения}}}{\sqrt{\frac{s_{\text{база}}^{2}}{n_{\text{база}}} + \frac{s_{\text{остальные значения}}^{2}}{n_{\text{остальные значения}}}}} < t_{\alpha} \Big(n_{\text{база}} + n_{\text{остальные значения}} - 2 \Big),$$

то средние считаются статистически неразличимыми, т.е. разницы в смене уровня фактора по сравнению с основным уровнем – нет. В противном же случае – данный уровень фактора признается как наиболее сильно влияющий на результаты.

3.1.2 Однофакторный дисперсионный анализ при неодинаковом числе испытаний по уровням фактора (ОДА-НЧИ)

Различие между случаями, когда производится неравное число наблюдений по факторам и когда производится равное число наблюдений по уровням фактора, состоит в промежуточных вычислениях, таблица 3.2.

Таблица 3.2 – Схема ОДА-НЧИ

Группировка	Сумма квадратов	Число степеней свободы	Оценка дисперсии
-------------	-----------------	------------------------------	---------------------

1. Общая	$Q = \sum_{i}^{p} \sum_{j=1}^{n_{p}} x_{ij}^{2} - \frac{\left(\sum_{i=1}^{p} \sum_{j=1}^{n_{p}} x_{ij}\right)^{2}}{n}$	n – 1	$s^2 = \frac{Q}{N-1}$
2. По факторам А (между группа-ми)	$Q_{A} = \sum_{i=1}^{p} \frac{\left(\sum_{j=1}^{n_{p}} X_{ij}\right)^{2}}{n_{p}} - \frac{\left(\sum_{i=1}^{q} \sum_{j=1}^{n_{p}} X_{ij}\right)^{2}}{n}$	p – 1	$s_A^2 = \frac{Q_A}{p-1}$
3. Остаточная (внутри групп)	$Q_{R} = \sum_{i}^{p} \sum_{j=1}^{n_{p}} x_{ij}^{2} - \sum_{i=1}^{p} \frac{\left(\sum_{j=1}^{n_{p}} x_{ij}\right)^{2}}{n_{p}}$	n – p	$s_{R}^{2} = \frac{Q_{R}}{N - p}$
4. Оценка влияния фактора	$K_{F} = \frac{s_{A}^{2}}{s_{R}^{2}} \leq F_{\alpha}$		
5. Оценка влияния отдельных значений фактора	Сдель- начений $K_t = \frac{1}{\sqrt{\frac{\sigma_{\text{база}}^2 + \frac{\sigma_{\text{остальные значения}}^2}{\sigma_{\text{остальные значения}}^2}}} $		

В остальном же схема рассуждений сохраняется прежней, в связи с чем не будем ее еще раз подробно расписывать.

3.1.3. Двухфакторный дисперсионный анализ без повторений (ДДА-БП)

Рассуждения при проведении дисперсионного анализа в случае двух факторов аналогичны предыдущей ситуации. Особенность состоит в том, что здесь рассматриваются две группировки: сначала по группам одного фактора, потом – по группам другого.

В данном случай на результат эксперимента влияют два фактора A и B. Причем, фактор A принимает значения $A_1,\ A_2...\ A_p;$ а фактор B принимает значения $B_1,\ B_2...\ B_q.$

В этом случае для построения схемы двухфакторного дисперсионного анализа без повторений (ДДА-БП) определяются групповые средние по фак-

тору А и по фактору В отдельно, а также общее среднее. Приведем формулы для вычисления:

1. групповое среднее по факторам А

$$\bar{x}_i = \frac{\sum_{j=1}^{q} x_{ij}}{q};$$

2. групповое среднее по факторам В

$$\overline{x}_{j} = \frac{\sum_{i=1}^{p} x_{ij}}{p};$$

3. общее среднее

$$\frac{-}{x} = \frac{\sum_{i=1}^{p} \sum_{j=1}^{q} x_{ij}}{pq}.$$

Думаю, вы обратили внимание, что групповое среднее определяется несколько странно: групповое среднее определяется по уровням фактора A, а индекс суммирования j — по фактору B, да и делится сумма на число уровней фактора B. И наоборот. Никакой ошибки здесь нет.

Дело в том, что очень часто, удобства ради записи результатов экспериментов по двухфакторному анализу записываются в виде таблицы, таблица 3.3.

Таблица 3.3 – Пример записи результатов ДДА-БП

Уровни	Уровни фактора В
--------	------------------

фактора А	B_1	B_2		$\mathrm{B}_{\mathfrak{q}}$
A_1	x ₁₁	X ₁₂	•••	x _{1q}
A_2	X ₂₁	X ₂₂	•••	x _{2q}
	•••	•••	•••	•••
A_{p}	X _{p1}	X _{p2}		X _{pq}

Из приведенного примера сразу видно, что среднее по уровням фактора А представляет собой среднее по строкам таблицы 3.3, а среднее по уровням фактора В – среднее по столбцам таблицы 3.3. С этих позиций сразу становится понятно и использование индекса суммирования, и почему в знаменателе при определении средних стоит число уровней другого фактора. Эти же соображения относятся и к умножению межгрупповых сумм квадратов на число уровней другого фактора (строки 2 и 3 таблицы 3.4)

Схема ДДА-БП имеет вид, представленный в таблице 3.4.

Следует обратить внимание на тот факт, что для **обоих** факторов проводится проверка на значимость (строка 5 таблицы 3.4); оценка влияния отдельных значений фактора проводится только для значимых факторов.

Несмотря на то, что вычислений прибавилось, общая схема рассуждений остается все той же (см. п. 3.1.1), поэтому вновь подробно описывать всю схему, по-моему, не имеет смысла. Единственное, что следует заметить, что шестой пункт схемы будет выполняться для тех факторов, которые будут признаны значимыми, т.е. если оба фактора окажутся значимыми, для каждого из них будет проводиться оценка влияния отдельных уровней на результат.

Группировка	Сумма квадратов	Число степеней свободы	Оценка дисперсии			
1. Общая	$Q = \sum_{i=1}^{p} \sum_{j=1}^{q} (x_{ij} - \overline{x})^{2}$	pq – 1	$s^2 = \frac{Q}{pq - 1}$			
2. Между группам фактора А	$Q_{A} = q \sum_{i=1}^{p} (\overline{x}_{i} - \overline{x})^{2}$	p – 1	$s_A^2 = \frac{Q_A}{p-1}$			
3. Между группам фактора В	$Q_{B} = p \sum_{j=1}^{q} (\overline{x}_{j} - \overline{x})^{2}$	q – 1	$s_{\rm B}^2 = \frac{Q_{\rm B}}{q-1}$			
4. Остаточная (внутри групп)	$Q_R = \sum_{i=1}^p \sum_{j=1}^q (x_{ij} - \overline{x}_i - \overline{x}_j + \overline{x})^2$	(p-1)· $(q-1)$	$s_R^2 = \frac{Q_R}{(p-1)(q-1)}$			
5. Оценка влияния фактора	$K_{F_A} = \frac{s_A^2}{s_R^2} \le F_{\alpha}(p-1; (p-1)(q-1))$ и $K_{F_B} = \frac{s_B^2}{s_P^2} \le F_{\alpha}(q-1; (p-1)(q-1))$					
6. Оценка влияния отдельных значений фактора	$K_{t} = \frac{\overline{x}_{\text{база}} - \overline{x}_{\text{остальные значения}}}{\sqrt{\frac{\sigma_{\text{база}}^{2}}{n_{\text{база}}} + \frac{\sigma_{\text{остальные значения}}^{2}}{n_{\text{остальные значения}}}}} \lesssim t_{\alpha} \Big(n_{\text{база}} + n_{\text{остальные значения}} - 2 \Big)$					

3.1.4. Двухфакторный дисперсионный анализ с повторениями (ДДА-П)

Двухфакторный дисперсионный анализ с повторениями проводится для случая влияния двух факторов на результат, но для каждой пары A_i и B_j проводится по несколько измерений случайной величины X. Главное условие для каждой пары факторов: число повторений должно быть одно и тоже. Для ясности приведем пример записи результатов в двухфакторном дисперсионном анализе с повторениями (ДДА-П), таблица 3.5.

Таблица 3.5 – Пример записи результатов ДДА-П

Уровни	Уровни фактора В						
фактора А	B_1	B_2		B_q			
A_1	X ₁₁₁	X ₁₂₁		X _{1q1}			
	X ₁₁₂	X ₁₂₂		x_{1q2}			
	X ₁₁₃	X ₁₂₃		X _{1q3}			
A_2	X ₂₁₁	X ₂₂₁		\mathbf{x}_{2q1}			
	X ₂₁₂	X ₂₂₂		X_{2q2}			
	X ₂₁₃	X ₂₂₃		X_{2q3}			
	•••						
A_p	X _{p11}	X _{p21}	• • •	X _{pq1}			
	X _{p12}	X _{p22}		X _{pq2}			
	X _{p13}	X _{p23}		X_{pq3}			

Определим индексы суммирования следующим образом:

- k = 1, 2, ..., n число измерений для одной пары факторов A_iB_j , т.е. число повторений (в нашем примере, таблица 3.5, k изменяется от 1 до 3);
- i = 1, 2, ..., p число уровней фактора А;
- j = 1, 2, ..., q число уровней фактора В.

Определим среднее для каждой пары факторов A_iB_j , групповое среднее для каждой пары факторов и общее среднее:

1. групповое среднее в ячейке, т.е. для каждой пары факторов $A_i B_j$

$$\bar{x}_{ij} = \frac{\sum_{k=1}^{n} x_{ijk}}{n};$$

2. групповое среднее по факторам А

$$\bar{x}_i = \frac{\sum_{j=1}^q \bar{x}_{ij}}{q};$$

3. групповое среднее по факторам В

$$\overline{x}_{j} = \frac{\sum_{i=1}^{p} \overline{x}_{ij}}{p};$$

4. общее среднее

$$\overline{x} = \frac{\sum_{i=1}^{p} \sum_{j=1}^{q} \overline{x}_{ij}}{pq}.$$

Учитывая введенные значения средних, схема ДДА-П примет вид, представленный в таблице 3.6.

Вообще говоря, после определения группового среднего ДДА с повторениями превращается в ДДА без повторений. Только вместо отдельного значения в ячейках стоят групповые средние.

Отдельно следует сказать о смысле такой группировки как взаимодействие. Здесь речь идет не об оценке степени взаимодействия между факторами, а об <u>оценке степени влияния взаимодействия факторов</u> на результат эксперимента.

Например, проводится оценка влияния смены рабочих и уровня их зарплаты на качество выполняемых работ. В результате было обнаружено, что влияющими оказались значения фактора А (смена) и взаимодействия. В этом случае результаты интерпретируются следующим образом: на качество выполняемых работ влияет тот факт, в какую смену они выполняются, а также уровень заработной платы за выполнение работ в ту или иную смену. В последнем предложении выделена фраза, соответствующая оценке влияния взаимодействия факторов.

Группировка	Сумма квадратов	Число степеней свободы	Оценка дисперсии			
1. Общая	$Q = \sum_{i=1}^{p} \sum_{j=1}^{q} \sum_{k=1}^{n} (x_{ijk} - \overline{x})^{2}$	pqn – 1	$s^2 = \frac{Q}{pqn - 1}$			
2. Между группами фактора А	$Q_{A} = qn \sum_{i=1}^{p} (\overline{x}_{i} - \overline{x})^{2}$	p – 1	$s_A^2 = \frac{Q_A}{p-1}$			
3. Между группами фактора В	$Q_{B} = pn \sum_{j=1}^{q} (\overline{x}_{j} - \overline{x})^{2}$	q – 1	$s_{\rm B}^2 = \frac{Q_{\rm B}}{q-1}$			
4. Взаимодействие	$Q_{AB} = n \sum_{i=1}^{p} \sum_{j=1}^{q} (\overline{x}_{ij} - \overline{x}_{i} - \overline{x}_{j} + \overline{x})^{2}$	(p-1) $(q-1)$	$s_{AB}^2 = \frac{Q_{AB}}{(p-1)(q-1)}$			
5. Остаточная (внутри групп)	$Q_{R} = \sum_{i=1}^{p} \sum_{j=1}^{q} \sum_{k=1}^{n} (x_{ijk} - \overline{x}_{ij})^{2}$	pqn – pq	$s_{R}^{2} = \frac{Q_{R}}{pqn - pq}$			
6. Оценка влияния фактора	2	$K_{F_{A}} = \frac{s_{A}^{2}}{s_{R}^{2}} \le F_{\alpha}(p-1; pqn-pq),$ $K_{F_{B}} = \frac{s_{B}^{2}}{s_{R}^{2}} \le F_{\alpha}(q-1; pqn-pq)$ и				
фиктори	$K_{F_{AB}} = \frac{s_{AB}^2}{s_R^2} \le F_{\alpha}((p-1)(q-1); pqn - pq)$					
7. Оценка влияния отдельных значений фактора	$K_{\rm t} = rac{\overline{x}_{ m база} - \overline{x}_{ m остальные значения}}{\sqrt{rac{\sigma_{ m база}^2}{n_{ m база}} + rac{\sigma_{ m остальные значения}^2}{n_{ m остальные значения}}}} \lesssim t_{ m a} \Big(n_{ m база} + n_{ m остальные значения} - 2 \Big)$					

3.2 Корреляционный анализ

Между любыми двумя, тремя... случайными величинами возможно существование следующих вариантов зависимостей:

- 1. отсутствие какой-либо зависимости;
- 2. статистическая зависимость это зависимость между случайными величинами, когда изменение одной величины вызывает изменение парамет-

ров распределения или вида самого распределения другой случайной величины;

3. функциональная зависимость — это зависимость между случайными величинами, которая может быть описана в виде функции X = f(Y).

Установить зависимость между случайными величинами можно либо графически, но это возможно только в случае двух-трех случайных величин, либо с помощью *корреляционного анализа*. Корреляционный анализ позволяет не только установить наличие зависимости между случайными величинами, но и дать качественную характеристику этой связи. В качестве такой меры служит *коэффициент корреляции*. Различают следующие виды коэффициентов корреляции:

- 1. парный линейный выборочный коэффициент корреляции r_{xy} ;
- 2. корреляционное отношение η_{xy} ;
- 3. множественный коэффициент корреляции $R_{i,jklm...}$ и частный выборочный коэффициент корреляции $r_{ij,klm...}$;
- 4. ранговые коэффициенты корреляции Спирмена и Кендалла.

Силу связи с помощью коэффициентов корреляции оценивают по так называемой шкале Чеддока, таблица 3.7, причем эта шкала универсальна для всех видов коэффициентов корреляции.

Думаю, Вы обратили внимание на тот факт, что коэффициент корреляции по абсолютному значению не превышает единицы. Знак количественной характеристики может быть положительным или отрицательным, об этом будет сказано в каждом отдельном случае, но качественно теснота связи будет определяться именно этой шкалой, без учета знака.

Таблица 3.7 – Шкала Чеддока для оценки силы связи между случайными величинами

Количественная мера					
тесноты связи, абсолютное значение					

Качественная характеристика силы связи

0,0-0,09	Весьма слабая	
0,1-0,29	Слабая	
0,3-0,49	Умеренная	
0,5-0,69	Заметная	
0,7-0,89	Высокая	
0,9-1,0	Весьма высокая	

С точки зрения теории «Планирования и организации эксперимента» особая ценность корреляционного анализа заключается в способности оценить наличие зависимости между влияющими на параметр оптимизации факторами, чтобы проверить выполнимость требования не коррелированности совокупности факторов. Существует следующее правило: если сила связи между двумя факторами не превышает умеренную, то можно оба фактора оставить в рассмотрении. В противном случае — один из двух факторов из рассмотрения выбрасывается. Конечно, это правило можно смягчить, разрешив использование факторов, между которыми наблюдается заметная связь, но, как правило, это делать не рекомендуется.

При оценивании связи между факторами существует одна опасность. Дело в том, что корреляционный анализ, как и любой другой математический аппарат работает, прежде всего, с цифрами, не обращая внимания на природу их возникновения и физический смысл. Поэтому в ходе проверки может возникнуть наличие высокой и более связи между теми случайными величинами, где ее не может вообще существовать по логике вещей. Конечно, для устранения подобных случаев существует проверка значимости коэффициента корреляции, но голову экспериментатора тоже не следует исключать из анализа. В моей практике был случай, когда студент утверждал, что качество сгущенного молока напрямую зависит от того, в банку какого цвета оно упаковано — синюю или зеленую. Свое утверждение он аргументировал тем, что между этими характеристиками существует весьма высокая корреляционная

зависимость. Чтобы Вам не попасться на подобную удочку, будьте внимательны!

А теперь рассмотрим все указанные выше виды коэффициентов корреляции. При этом будем приводить только используемые на практике формулы расчета коэффициентов корреляции и оценки их значимости. Теоретические выводы данных формул, при желании, можно найти в учебной и учебнометодической литературе по теории вероятностей и математической статистике.

3.2.1. Парный выборочный линейный коэффициент корреляции

Предназначен для выявления <u>линейной</u> связи между двумя случайными величинами. Определяется по экспериментальной выборке значений случайных величин X и Y. Отсюда и название данного коэффициента. Следует понимать, что равенство нулю данного коэффициента корреляции говорит от отсутствии ЛИНЕЙНОЙ зависимости между X и Y, т.е. зависимости типа

$$Y = kX + b$$
,

во всех ее возможных проявлениях. Отсутствие вообще какой-либо зависимости между случайными величинами (нелинейной или статистической) может подтвердить только *корреляционное отношение*, которое будет рассмотрено ниже.

Парный выборочный линейный коэффициент корреляции на практике удобнее всего определять по формуле:

$$r_{xy} = \frac{N\sum_{i=1}^{m}\sum_{j=1}^{k}n_{ij}x_{i}y_{j} - \left(\sum_{i=1}^{m}n_{i}x_{i}\right)\left(\sum_{j=1}^{k}n_{j}y_{j}\right)}{\sqrt{N\sum_{i=1}^{m}x_{i}^{2}n_{i} - \left(\sum_{i=1}^{m}n_{i}x_{i}\right)^{2}} \cdot \sqrt{N\sum_{j=1}^{k}y_{j}^{2}n_{j} - \left(\sum_{j=1}^{k}n_{j}y_{j}\right)^{2}}},$$
(3.1)

где і – индекс суммирования по значениям случайной величины Х;

ј – индекс суммирования по значениям случайной величины Y;

 n_{ij} – частота встречи пары $x_i y_i$;

n_i – число значений случайной величины X;

 n_i – число значений случайной величины Y;

N – общее число проведенных экспериментов.

Обозначение r_{xy} следует читать как «влияние случайной величины Y на изменчивость случайной величины X».

Чтобы произвести вычисления по формуле (3.1), необходимо данные заранее сгруппировать в корреляционную таблицу, общий вид которой представлен в таблице 3.8.

Таблица 3.8 – Общий вид корреляционной таблицы

	y ₁	y ₂	•••	y_k
X ₁	n ₁₁	n ₁₂	•••	n_{1k}
X_2	n ₂₁	n ₂₂		n _{2k}
•••	•••	•••	•••	•••
X _m	n _{m1}	n_{m2}		n _{mk}

Если же это сделать невозможно или каждая пара значений $x_i y_j$ встречается один раз, то формула (3.1) примет вид

$$r_{xy} = \frac{N\sum_{i=1}^{N} x_i y_i - \left(\sum_{i=1}^{N} x_i\right) \left(\sum_{i=1}^{N} y_i\right)}{\sqrt{N\sum_{i=1}^{N} x_i^2 - \left(\sum_{i=1}^{N} x_i\right)^2} \cdot \sqrt{N\sum_{i=1}^{N} y_i^2 - \left(\sum_{i=1}^{N} y_i\right)^2}},$$
(3.2)

где і – индекс суммирования по значениям случайных величин Х и Y;

N – общее число проведенных экспериментов.

Несмотря на громоздкость, формулы (3.1) и (3.2) действительно очень удобны для вычислений, поскольку не требуют дополнительных, промежуточных вычислений, а используют напрямую определяемые из опытов значения случайных величин.

Отметим *свойства* парного линейного выборочного коэффициента корреляции.

- 1. $-1 \le r_{xy} \le 1$. Причем:
 - а. $\left| \mathbf{r}_{xy} \right| = 1$ наличие линейной связи между случайными величинами X и Y, рисунок 3.1, а и 6;
 - b. $r_{xy} = 0$ отсутствие линейной связи между случайными величинами X и Y, рисунок 3.1, B;
 - с. $-1 < r_{xy} < 1$ наличие либо нелинейной, либо статистической связи между случайными величинами X и Y.
- 2. $0 < r_{xy}$ между случайными величинами X и Y наблюдается обратная зависимость, т.е. при возрастании значений одной случайной величины значения другой случайной величины уменьшаются, рисунок 3.1, а.
- 3. $r_{xy} > 0$ между случайными величинами X и Y наблюдается прямая зависимость, т.е. при возрастании значений одной случайной величины значения другой случайной величины также увеличиваются, рисунок 3.1, б.
- 4. $r_{xy} = r_{yx}$.

- 5. При увеличении (уменьшении) значений случайных величин на одно и то же число (или в одно и то же число раз) значение \mathbf{r}_{xy} остается неизменным
- 6. $r_{xx} = r_{yy} = 1$ поскольку это действие случайной величины самой на себя.

С Вашего позволения, еще раз обращу внимание на свойство 1b. Если $r_{xy}=0$, то мы говорим об отсутствии ТОЛЬКО линейной зависимости. Говорить об отсутствии зависимости между случайными величинами X и Y вообще можно после проверки корреляционного отношения (см. ниже).

Помимо коэффициента корреляции, можно также вычислить коэффи-*циент детерминации*, который показывает на сколько процентов изменчивость случайной величины Y зависит от изменчивости случайной величины X^{*} . Определяется коэффициент детерминации следующим образом

$$\Delta = (r_{xy})^2 \cdot 100 \% .$$

Значимость r_{xy} проверяется по критерию согласия Стьюдента. При этом в качестве основной гипотезы проверяется гипотеза об отсутствии линейной корреляции, т.е.

$$H_0: r_{xy} = 0;$$
 $H_1: r_{xy} \neq 0.$

Наблюдаемое значение критерия определяется по формуле:

$$K_r = \frac{r\sqrt{N-2}}{\sqrt{1-r^2}},$$

 $^{^{*)}}$ коэффициент детерминации можно определять по любому из представленных коэффициентов корреляции. Принцип определения везде один и тот же. Поэтому мы лишь один раз приведем его подробно. Далее — все по аналогии.

где N – общее число опытов.

Парный выборочный линейный коэффициент корреляции признается значимым (т.е. основная гипотеза <u>отвергается</u>), если

$$|K_r| \ge t_{\alpha}(N-2)$$

3.2.2 Корреляционное отношение

Корреляционное отношение позволяет выявить наличие или отсутствие связи между случайными величинами X и Y. Будем определять все формулы и выкладки по корреляционному отношению из предположения, что на изменчивость случайной величины Y влияют значения случайной величины X. Определяется корреляционное отношение на основе межгрупповой и общей дисперсий измеряемой величины. В нашем случае формула будет иметь вид:

$$\eta_{yx} = \sqrt{\frac{s_{\text{межгруп.по y}}^2}{s_{\text{общ.по y}}^2}}$$
 .

Определение общей дисперсии переменной Y $s_{\text{общ. по y}}^2$ производится по формуле:

$$s_{\text{общ.по y}}^2 = rac{\sum\limits_{j=1}^{m} (y_j - \overline{y})^2 n_j}{N},$$

где \overline{y} — среднее значение (математическое ожидание) случайной величины Y, оцененное по экспериментальным данным;

 n_{i} – частота встречи значения y_{i} ;

т – общее число значений у;

N – общее число проведенных экспериментов.

Для оценки межгрупповой дисперсии переменной Y $s_{\text{межгруп.по y}}^2$ необходимо произвести группировку значений переменной Y в зависимости от значений переменной X, т.е. отдельно «собрать» все y_j , которые были отмечены при значении x_1 , отдельно «собрать» все y_j , которые были отмечены при значении x_2 и т.д. По каждой полученной группе оценить средние значения величины y_i обозначим их y_i . Затем можно оценивать межгрупповую дисперсию:

$$s_{\text{межгруп.по y}}^2 = \frac{\sum\limits_{i=1}^k (\overline{y}_i - \overline{y})^2 n_i}{N},$$

где \overline{y} – среднее значение (математическое ожидание) случайной величины Y, оцененное по экспериментальным данным;

 \overline{y}_i – групповые средние значения (математические ожидания) случайной величины Y, оцененные по экспериментальным данным, сгруппированным по значением случайной величины X;

 n_{i} — частота встречи значения x_{i} ;

k – общее число значений x_i ;

N – общее число проведенных экспериментов.

Для дополнительного прояснения ситуации с определением общей и межгрупповой дисперсий можно обратиться к дисперсионному анализу (параграф 3.1) – принцип тот же самый.

Отметим свойства корреляционного отношения:

1. $0 \le \eta_{yx} \le 1$. Причем:

- а. $\eta_{yx} = 1$ наличие **функциональной** зависимости между случайными величинами X и Y;
- b. $\eta_{yx} = 0 \underline{\text{отсутствие какой-либо связи}}$ между случайными величинами X и Y;
- с. $0 < \eta_{yx} < 1$ наличие статистической связи между случайными величинами X и Y.
- 2. $\eta_{vx} \neq \eta_{xv}$.

Следует отметить, что направление связи между случайными величинами (прямая или обратная) в данном случае выявить не удастся.

Замечание. Фактически, после определения $\mathbf{r}_{xy} = 0$ необходимо оценить корреляционное отношение, и только по результатам последнего уже выносить «приговор» зависимости между двумя случайными величинами:

- а) если $r_{xy} = 0$, $\eta_{yx} = 1$ между случайными величинами X и Y наблюдается функциональная зависимость, но она носит нелинейный характер;
- b) если $r_{xy} = 0$, $\eta_{yx} = 0$ между случайными величинами X и Y не наблюдается какой-либо зависимости.

Значимость корреляционного отношения определяется по критерию согласия Фишера – Снедекора. По-прежнему,

$$H_0: \eta_{yx} = 0;$$
 $H_1: \eta_{yx} \neq 0.$

Наблюдаемое значение критерия определяется по формуле:

$$K_{\eta} = \frac{\eta_{yx}^2 \cdot (N - m)}{\left(1 - \eta_{yx}^2\right)(m - 1)},$$

где N – общее число опытов;

m – число полученных групп при определении межгрупповой дисперсии (фактически, это число значений случайной величины X).

Корреляционное отношение признается значимым (т.е. основная гипотеза <u>отвергается</u>), если

$$K_{\eta} \ge F_{\alpha}(m-1; N-m)$$

3.2.3 Множественный коэффициент корреляции и частный выборочный коэффициент корреляции

Множественный коэффициент корреляции, равно как и частный выборочный коэффициент корреляции, определяются в случае выявления зависимостей между случайными величинами, чье количество превышает два. Разница между этими двумя коэффициентами состоит в следующем:

- 1) множественный коэффициент корреляции оценивает влияние нескольких (больше двух) факторов на параметр оптимизации;
- 2) частный выборочный коэффициент корреляции оценивает зависимость между двумя параметрами (между двумя факторами, между фактором и параметром оптимизации и т.п.) при исключении влияния остальных параметров взаимодействия.

При взаимодействии нескольких случайных величин обычно строится корреляционная матрица, членами которой являются парные выборочные линейные коэффициенты корреляции между взаимодействующими случайными величинами. По главной диагонали данной матрицы располагаются единицы, а сама матрица — симметрична относительно главной диагонали. В общем случае корреляционная матрица имеет вид:

$$\mathbf{q} = \begin{bmatrix} 1 & r_{12} & r_{13} & \dots & r_{1k} \\ r_{21} & 1 & r_{23} & \dots & r_{2k} \\ r_{31} & r_{32} & 1 & \dots & r_{3k} \\ \dots & \dots & \dots & \dots \\ r_{k1} & r_{k2} & r_{k3} & \dots & 1 \end{bmatrix}.$$

Множественный коэффициент корреляции определяется по формуле*)

$$R_{j.1,2,...(j-1),(j+1),...k} = \sqrt{1 - \frac{|q|}{q_{jj}}},$$

где |q| - определитель корреляционной матрицы;

 ${
m q}_{\rm jj}$ — алгебраическое дополнение соответствующего элемента корреляционной матрицы.

Значимость множественного коэффициента корреляции определяется по критерию согласия Фишера – Снедекора.

$$H_0: R = 0;$$
 $H_1: R \neq 0.$

Наблюдаемое значение критерия определяется по формуле:

$$K_{R} = \frac{R^{2} \cdot (N-k)}{(1-R^{2})(k-1)},$$

где N – общее число опытов;

k – число переменных во взаимодействии.

 $^{^{*)}}$ Данное обозначение читается следующим образом: «коэффициент корреляции на случайную величину J случайных величин 1, 2, ..., K».

Множественный коэффициент корреляции признается значимым (т.е. основная гипотеза <u>отвергается</u>), если

$$K_R \ge F_{\alpha}(k-1; N-k).$$

Частный выборочный коэффициент корреляции определяется как $^{*)}$

$$r_{ij.1,...,k} = \frac{-q_{ij}}{\sqrt{q_{ii}q_{jj}}},$$

где $q_{ij}, q_{ii} q_{jj}$ – алгебраические дополнения соответствующих элементов корреляционной матрицы.

Значимость частного выборочного коэффициента корреляции определяется по критерию согласия Стьюдента.

$$H_0: r_{ij,1,...,k} = 0;$$
 $H_1: r_{ij,1,...,k} \neq 0.$

Наблюдаемое значение критерия определяется по формуле:

$$K_{r,1} = \frac{r_{ij,1,...,k} \sqrt{N-2}}{\sqrt{1-r_{ij,1,...,k}^2}},$$

где N – общее число опытов.

Частный выборочный коэффициент корреляции признается значимым (т.е. основная гипотеза отвергается), если

 $^{^{*)}}$ Данное обозначение следует читать как «взаимодействие между случайными величинами I и J при исключении влияния остальных случайных величин»

$$|K_{r,1}| \ge t_{\alpha}(N-k+2),$$

где k – число переменных во взаимодействии.

Замечание. Свойства множественного и частного коэффициентов корреляц3ии совпадают со свойствами корреляционного отношения и парного линейного выборочного коэффициента корреляции соответственно.

3.2.4 Ранговые коэффициенты корреляции

Все перечисленные выше коэффициенты корреляции, несмотря на всю свою необходимость, не позволяют, однако, оценивать зависимости качественных переменных. В лучшем случае качественные показатели можно подвергнуть процедуре ранжировки, но это не сделает их количественными, а значит – применять описанные выше показатели связи нельзя.

Для оценки ранжированных переменных существуют свои коэффициенты корреляции: коэффициенты Спирмена и Кендалла. Оба эти коэффициента оценивают совпадение (или не совпадение) рангов двух совокупностей по одному ранжируемому признаку. Ярким примером такого подхода является оценка участников в шоу «Ледниковый период». В этом случае ранжируемым признаком являются пары-участники, а рангами совокупностей — баллы, полученные участниками в ходе соревнований. В результате такого сравнения возможно выявление зависимости, например, между победами участников при «откатывании» различной программы.

Приведем методики оценки коэффициентов ранговой корреляции.

А. Коэффициент ранговой корреляции Спирмена

Для того, чтобы оценить коэффициент ранговой корреляции Спирмена, необходимо, прежде всего, определиться по какому признаку будет производиться ранжирование. Затем провести оценку рангов по этому признаку для двух совокупностей.

Коэффициент ранговой корреляции Спирмена определяется по формуле:

$$\rho = 1 - \frac{6\sum_{i=1}^{n} (r_i - s_i)^2}{n^3 - n}$$

где r_i , s_i – ранги i-го объекта по совокупностям X и Y;

n – число пар наблюдений.

Иногда при исследованиях сталкиваются со случаями, когда для разных значений признака ранжирования в одной совокупности существуют одинаковые ранговые значение. Такие случаи называются *случаями со связанными рангами*. Если невозможно решить, какие ранги приписать этим объектам, им всем приписывается одинаковый средний ранг.

В случае связанных рангов коэффициент Спирмена вычисляется по формуле:

$$\rho = 1 - \frac{\sum_{i=1}^{n} (r_i - s_i)^2}{\frac{1}{6} (n^3 - n) - (T_r + T_s)},$$

где
$$T_r = \frac{1}{12} \sum_{i=1}^{m_r} (t_r^3 - t_r), \ T_s = \frac{1}{12} \sum_{i=1}^{m_s} (t_s^3 - t_s)$$
 – поправочные коэффициенты;

 $m_{r}, \ m_{s}$ — число групп неразличимых рангов у первой и второй совокупности соответственно;

04

 $t_{\rm r},\,t_{\rm s}$ – число рангов, вошедших в соответствующую группу.

Для нашего примера с участниками шоу «Ледниковый период» случай связанных рангов будет выглядеть следующим образом. Пусть в ходе проведения соревнований участники шоу получили следующие баллы, таблица 3.9.

Номер	Набранн	ые баллы	Занято	е место
пары-участницы	1 этап	2 этап	1 этап	2 этап
01	15	19	1	2
14	14	20	2-3	1
10	14	18	2-3	3-5
02	12	17	4-6	4-8
03	12	18	4-6	3-5
07	12	18	4-6	3-5
11	11	17	7-8	4-8
17	11	17	7-8	4-8
05	10	17	9-10	4-8
15	10	17	9-10	4-8
12	9	14	11-14	9-10
16	9	14	11-14	9-10
06	9	13	11-14	11
	_			

Таблица 3.9 – Результаты соревнований в шоу «Ледниковый период»

При этом оказалось, что некоторые пары набрали одинаковое количество баллов и заняли, соответственно одинаковые места. Отдать кому-либо из них предпочтение перед другими не удалось. Тогда каждой из этих пар присваивается средний ранг, равный (таблица 3.10):

12

11-14

12

Номера пар	Расчет среднего ранга
14; 10	(2+3)/2=2,5
02; 03; 07	(4+5+6)/3=5
12; 16; 06; 04	(11 + 12 + 13 + 14) / 4 = 12,5
	ИТ.Д.

16

06

04

Сумма

11-14

11-14

11-14

9-10

11

12

	_			_			
Номер	Занято	е место	Приписыва	аемый ранг			
пары- участницы	1 этап	2 этап	1 этап, r _i	2 этап, s _i	$(\mathbf{r_i} - \mathbf{s_i})$	$(\mathbf{r}_{i}-\mathbf{s}_{i})^{2}$	
01	1	2	1	2	-1	1	
14	2-3	1	2,5	1	1,5	2,25	
10	2-3	3-5	2,5	4	-1,5	2,25	
02	4-6	4-8	5	6	-1	1	
03	4-6	3-5	5	4	1	1	
07	4-6	3-5	5	4	1	1	
11	7-8	4-8	7,5	6	1,5	2,25	
17	7-8	4-8	7,5	6	1,5	2,25	
05	9-10	4-8	9,5	6	3,5	12,25	
15	9-10	4-8	9,5	6	3,5	12,25	
12	11-14	9-10	12,5	9,5	3	9	
4.2						_	

Таблица 3.10 – Результаты распределения рангов пар-участниц шоу «Ледниковый период» и промежуточных расчетов коэффициента корреляции Спирмена

В нашем примере для первой совокупности (1 этап) число групп равно пяти, т.к. было выявлено пять групп совпавших значений; а для второй совокупности — число групп равно трем. Число рангов для каждой из групп первого этапа, соответственно, составило: два, три, два, два, четыре. Число рангов для каждой группы второго этапа соответственно равно: три, пять, два.

12.5

0.5

Рассчитаем коэффициент корреляции Спирмена для представленных данных и определим, зависят ли дальнейшие успехи пар от их предыдущих побед. Для этого рассчитаем поправочные коэффициенты для первой (T_r) и второй (T_s) совокупностей.

$$T_{r} = \frac{1}{12} \sum_{r=1}^{5} (t_{r}^{3} - t_{r}) = \frac{1}{12} [(2^{3} - 2) + (3^{3} - 3) + (2^{3} - 2) + (2^{3} - 2) + (4^{3} - 4)] =$$

$$= \frac{1}{12} [3 \cdot (8 - 2) + (27 - 3) + (64 - 4)] = \frac{18 + 24 + 60}{12} = 8,5;$$

$$T_{s} = \frac{1}{12} \sum_{r=1}^{3} (t_{s}^{3} - t_{s}) = \frac{1}{12} [(3^{3} - 3) + (5^{3} - 5) + (2^{3} - 2)] =$$

$$= \frac{1}{12} [(27 - 3) + (125 - 5) + (8 - 2)] = \frac{24 + 120 + 6}{12} = 12,5.$$

Тогда коэффициент корреляции Спирмена примет вид:

$$\rho = 1 - \frac{\sum_{i=1}^{n} (r_i - s_i)^2}{\frac{1}{6} (n^3 - n) - (T_r + T_s)} = 1 - \frac{58}{\frac{1}{6} (14^3 - 14) - (8,5 - 12,5)} = 1 - \frac{58}{455 + 4} \approx 0,87.$$

Учитывая оценку силы корреляционной связи по шкале Чеддока, таблица 3.7, видно, что связь между двумя совокупностями сильная. Таким образом, можно сделать вывод, что успехи пар-участниц шоу «Ледниковый период» напрямую зависят от их побед, одержанных ранее.

Оценка значимости коэффициента ранговой корреляции Спирмена, независимо от того, по какой из двух формул он вычислялся, производится по критерию согласия Стьюдента.

$$H_0: \rho = 0;$$
 $H_1: \rho \neq 0.$

Наблюдаемое значение критерия определяется по формуле:

$$K_{\rho} = \frac{\rho \sqrt{n-2}}{\sqrt{1-\rho^2}},$$

где п -число пар наблюдений.

Ранговый коэффициент корреляции Спирмена признается значимым (т.е. основная гипотеза <u>отвергается</u>), если

$$|K_{\rho}| \ge t_{\alpha}(n-2).$$

Для нашего примера, при уровне значимости $\alpha = 5$ %:

$$K_{\rho} = \frac{\rho\sqrt{n-2}}{\sqrt{1-\rho^2}} = \frac{0.87 \cdot \sqrt{14-2}}{\sqrt{1-0.87^2}} = \frac{3.01}{0.49} = 6.14;$$

$$t_{\alpha=0.05} (14-2) = 2.18.$$

Так как $\left|K_{\rho}\right| > t_{\alpha}(n-2)$, то рассчитанный нами коэффициент корреляции Спирмена признается значимым, т.е. выводы относительно связи успехов команд, сделанные ранее – справедливы.

В. Коэффициент ранговой корреляции Кендалла

Для того чтобы оценить коэффициент ранговой корреляции Кендалла, необходимо провести ранжировку исследуемого объекта (в нашем примере — пар-участниц) в порядке возрастания рангов по одной переменной (например, по первому этапу) и определить, сколько раз произошло нарушение порядка следования рангов по другой переменной. При этом определяется так называемое число инверсий. Инверсия — случай, когда большее число стоит слева от меньшего. Величина K, называемая статистикой Кендалла, равна общему числу инверсий в ранговой последовательности. Чтобы понять, как просчитывается число инверсий, приведем пример * .

^{*)} Данный пример заимствован нами из [13].

Пример. По результатам спортивных состязаний десять спортсменов в течение двух дней испытаний получили следующие баллы, таблица 3.11. Оценить, зависят ли результаты соревнований во второй день от результатов первого дня.

День	Условный код спортсмена									
депь	A	Б	В	Γ	Д	Е	Ж	3	И	К
I	1	2	3	4	5	6	7	8	9	10
II	1	4	2	6	3	9	10	8	7	5

Таблица 3.11 – Результаты соревнований.

В качестве ранжируемого признака будут выступать сами спортсмены. Расположим результаты в порядке возрастания значений в первый день. При этом значения во второй день несколько перемешаются, таблица 3.9.

Подсчитаем общее число инверсий, которое в результате получилось. Первое нарушение порядка в следовании рангов мы наблюдаем на второй позиции. С учетом последующих рангов, имеем последовательность

Рассмотрим образовавшиеся пары рангов:

Легко можно заметить, что образовалось всего две инверсии, они выделены полужирным шрифтом. Таким образом, для второй позиции записываем значение статистики Кендалла $K_2 = 2$.

Аналогично, можно подсчитать статистики Кендалла для инверсий на четвертой, шестой, седьмой, восьмой и девятой позиций. Они составят, соответственно,

$$K_4 = 2$$
; $K_6 = 3$; $K_7 = 3$; $K_8 = 2$; $K_9 = 1$.

(Советуем самостоятельно в этом убедиться, чтобы механизм расчета статистики Кендалла стал абсолютно ясен). Тогда статистика Кендалла для всей последовательности будет равна

$$\underline{\mathbf{K}} = \mathbf{K}_2 + \mathbf{K}_4 + \mathbf{K}_6 + \mathbf{K}_7 + \mathbf{K}_8 + \mathbf{K}_9 = \underline{\mathbf{13}}.$$

Вернемся к коэффициенту ранговой корреляции Кендалла. Коэффициент ранговой корреляции Кендалла определяется по формуле:

$$\tau = 1 - \frac{4K}{n(n-1)}.$$

Оценка значимости коэффициента ранговой корреляции Кендалла производится по критерию согласия Стьюдента.

$$H_0: \rho = 0;$$
 $H_1: \rho \neq 0.$

Наблюдаемое значение критерия определяется по формуле:

$$K_{\tau} = \tau \sqrt{\frac{9n(n-1)}{2(2n+5)}},$$

где п -число пар наблюдений.

Ранговый коэффициент корреляции Кендалла признается значимым (т.е. основная гипотеза <u>отвергается</u>), если

$$|K_{\tau}| \geq t_{1-\alpha}$$

где $t_{1-\alpha}$ определяется из выражения $\Phi(t_{1-\alpha}) = 1 - \alpha$;

 $\Phi(t_{1-\alpha})$ – функция Лапласа.

Для приведенного выше примера имеем:

$$\tau = 1 - \frac{4K}{n(n-1)} = 1 - \frac{4 \cdot 13}{10 \cdot (10-1)} = 0.42;$$

$$K_{\tau} = \tau \sqrt{\frac{9n(n-1)}{2(2n+5)}} = 0.42 \cdot \sqrt{\frac{9 \cdot 10 \cdot (10-1)}{2 \cdot (2 \cdot 10 + 5)}} = 0.42 \cdot 4.02 = 1.69$$

При уровне значимости $\alpha = 5 \%$:

$$\Phi(t_{1-\alpha}) = 1 - 0.05 = 0.95 \implies t_{1-\alpha} = 1.96.$$

Так как $|K_{\tau}| > t_{1-\alpha}$, то рассчитанный нами коэффициент корреляции Кендалла признается значимым, т.е. между результатами первого и второго дня соревнований действительно наблюдается зависимость. Однако, учитывая умеренный характер зависимости (см. таблицу 3.7), следует отметить, что данная связь не очень значительна.

Неоднократно мною отмечался тот факт, что чаще всего в статистике исследуют зависимость не между двумя, а между несколькими (больше двух) переменными. Тогда для оценки согласованности (читайте – корреляции) оценок используют так называемый

С. Коэффициент конкордации рангов Кендалла

Коэффициент конкордации рангов определяется по формуле:

$$W = \frac{12\sum_{i=1}^{n} \left(\sum_{j=1}^{m} r_{ij} - \frac{m(n+1)}{2}\right)}{m^{2}(n^{3} - n)},$$

где п – число объектов;

т – число анализируемых совокупностей.

Единственное условие для оценки коэффициента конкордации рангов Кендалла — число объектов $n \ge 7$.

Легко убедиться, что $0 \le W \le 1$, причем W = 1, если все совокупности совпадают между собой по рангам.

Значимость коэффициента конкордации рангов Кендалла оценивается по критерию согласия Пирсона. При этом

$$H_0: W = 0;$$
 $H_1: W \neq 0.$

Наблюдаемое значение критерия определяется по формуле:

$$K_{W} = m(n-1)W,$$

где п – число объектов;

т – число анализируемых совокупностей.

Коэффициент конкордации рангов Кендалла признается значимым (т.е. основная гипотеза <u>отвергается</u>), если

$$|K_{\rm W}| \ge \chi_{\alpha}^2 (n-1)$$

где $\chi^2_{\alpha}(n-1)$ – критическое значение χ^2 -распределения Пирсона при уровне значимости α с числом степеней свободы (n-1).

3.3 Регрессионный анализ

Регрессионный анализ представляет собой математический аппарат, который служит для построения математической модели эксперимента. Как уже упоминалось в параграфе 1.4, в планировании эксперимента чаще всего выбираются математические модели полиномиального характера.

Там же отмечалось, что экспериментатора после отбора полиномиальной модели заботит поиск ее коэффициентов. Фактически, этой фразой была определена задача регрессионного анализа с математической точки зрения. Поясним эту мысль на примере.

Простейшая полиномиальная модель имеет вид

$$y = b_0 + b_1 x.$$

Из предварительно проведенных опытов экспериментатору известны значения фактора x и результаты эксперимента y, которые при этих значениях фактора были зарегистрированы. Глядя на уравнение, сразу становится видно, что единственное, что неизвестно экспериментатору — коэффициенты x и y в y

Для определения коэффициентов полиномиальных моделей используются, чаще всего, метод моментов и метод наименьших квадратов. Причем, второй метод является самым популярным. Более того, в большинстве программных статистических пакетов для поиска коэффициентов уравнений используется именно метод наименьших квадратов.

Для того, чтобы имеет возможность контроля за программными статистическими пакетами, а также, чтобы понимать, откуда что берется, рассмотрим данный метод поиска коэффициентов регрессионной модели. Рассматривать данный метод будем на примере уравнения, приведенного выше.

Пусть была проведена серия из N опытов, при этом в каждом из проведенных опытов зависимость между установленным значением фактора x_i и полученным значением функции отклика y_i определялась выражением

$$y_i = b_1 x_i + b_0 + \varepsilon_i,$$

где ε_i - отклонение вследствие каких-либо случайных причин (погрешности).

После проведения всей серии опытов общая модель будет описываться совокупностью значений у_i на отдельных этапах, т.е.

$$\sum_{i=1}^{N} y_i = b_1 \sum_{i=1}^{N} x_i + b_0 + \sum_{i=1}^{N} \varepsilon_i.$$

При построении модели эксперимента исследователь, вполне естественно, старается свести к минимуму отклонения отдельных экспериментов, т.е. можно записать

$$\sum_{i=1}^{N} y_{i} - b_{1} \sum_{i=1}^{N} x_{i} - b_{0} = \sum_{i=1}^{N} \varepsilon_{i} \to \min.$$

Фактически, необходимо решить задачу по поиску минимума приведенной выше функции. Но прежде, чем заняться данной проблемой, нужно учесть еще один момент. Отклонения ε_i могут быть как положительные, так и отрицательные. В результате простого суммирования ε_i может возникнуть эффект компенсации: результат окажется либо ниже, чем есть на самом деле, либо вообще равным нулю. Чтобы избежать этого, обычно суммируют не сами отклонения, а их квадраты. Тогда получим:

$$\sum_{i=1}^{N} (y_i - b_1 x_i - b_0)^2 = \sum_{i=1}^{N} \epsilon_i^2 = S(b_0, b_1) \rightarrow \min.$$

Таким образом, наша задача в результате всех этих математических операций сводится к задаче поиска минимума функции S(a;b) при неизвестных коэффициентах a и b.

Для этого необходимо найти частные производные функции S(a;b) по неизвестным а и b, и решить систему уравнений относительно а и b, т.е.

$$\begin{cases} \frac{\partial S(b_0, b_1)}{\partial b_0} = 0 \\ \frac{\partial S(b_0, b_1)}{\partial b_1} = 0 \end{cases}$$

$$\begin{cases} 2\sum_{i=1}^{N} (y_i - b_1 x_i - b_0) \cdot (-x_i) = 0 \\ 2\sum_{i=1}^{N} (y_i - b_1 x_i - b_0) \cdot (-1) = 0 \end{cases} \Rightarrow \begin{cases} \sum_{i=1}^{N} x_i y_i = b_1 \sum_{i=1}^{N} x_i^2 + b_0 \sum_{i=1}^{N} x_i \\ \sum_{i=1}^{N} y_i = b_1 \sum_{i=1}^{N} x_i + b_0 N \end{cases}$$

Решив последнюю систему уравнений*), получаем:

$$b_{1} = \frac{N \sum_{i=1}^{N} x_{i} y_{i} - \sum_{i=1}^{N} x_{i} \sum_{i=1}^{N} y_{i}}{N \sum_{i=1}^{N} x_{i}^{2} - \left(\sum_{i=1}^{N} x_{i}\right)^{2}}.$$

Найти коэффициент b_0 можно, подставив полученный результат в одно из двух уравнений описанной системы. Однако есть путь попроще. В математической статистике доказано, что

^{*)} Советую Вам решить эту систему уравнений самостоятельно, получив описанные результаты.

$$b_0 = \overline{y} - a\overline{x}$$
.

Фактически, после этого можно спокойно записывать полученное уравнение регрессии. Однако на самом деле все не так просто.

Любой грамотный исследователь, прежде чем радоваться полученным результатам, проведет проверку значимости полученного уравнения регрессии и оценку значимости коэффициентов уравнения регрессии.

Провести оценку значимости — значит убедиться, что полученные результаты существуют для всей генеральной совокупности значений, выборка из которых была представлена результатами эксперимента. Проще говоря, необходимо убедиться, что построенная модель реально существует, а не является следствием случайного совпадения.

Сначала необходимо провести проверку значимости уравнения регрессии, поскольку, если уравнение не значимо, то оценивать значимость коэффициентов не имеет смысла. Вторым шагом проводиться проверка значимости коэффициентов уравнения регрессии.

1) Оценка значимости уравнения регрессии

Оценка значимости уравнения проводится по методике дисперсионного анализа. Проверить значимость уравнения регрессии — значит установить, является ли установленное из априорной информации уравнение регрессии адекватной моделью для исследуемого процесса (явления) и достаточно ли переменных для описания данного процесса было использовано.

В отличие от классического дисперсионного анализа (см. параграф 3.1), при оценке значимости уравнения регрессии рассматриваются следующие группировки данных: общая, регрессия и остаточная. Для оценки значимости коэффициента регрессии необходимо оценить:

1. Среднее значение параметра оптимизации \bar{y} во всей серии опытов:

$$\overline{y} = \frac{1}{N} \sum_{i=1}^{N} y_i ,$$

где N – общее число опытов.

2. Значения параметра оптимизации γ_i , рассчитанные по определенному ранее уравнению регрессии.

Схема дисперсионного анализа для оценки значимости уравнения регрессии примет вид, таблица 3.12.

Таблица 3.12 — Схема дисперсионного анализа для оценки значимости уравнения регрессии

Группировка	Сумма квадратов	Число степеней свободы	Оценка дисперсии			
Общая	$SS = \sum_{i=1}^{N} (y_i - \overline{y})^2$	N – 1	$\frac{SS}{N-1}$			
Регрессия	$SS_A = \sum_{i=1}^{N} (\gamma_i - \overline{y})^2$	m – 1	$s_{A}^{2} = \frac{SS_{A}}{m-1}$			
Остаточная	$SS_R = \sum_{i=1}^p (y_i - \gamma_i)^2$	N – m	$s_R^2 = \frac{SS_R}{N - m}$			
Оценка значимо- сти уравнения регрессии	$K_{F} = \frac{s_{A}^{2}}{s_{R}^{2}} > F_{\alpha}(m-1; N-m)$					

Здесь используются следующие обозначения: y_i — полученные экспериментально значения параметра оптимизации; \overline{y} — среднее значение экспериментально полученного параметра оптимизации; γ_i — предсказанные, т.е. рассчитанные по полученной модели, значения параметра оптимизации; N — число проведенных экспериментов; m — число коэффициентов в уравнении регрессии, включая свободный коэффициент b_0 .

Уравнение регрессии признается значимым, если

$$K_F = \frac{s_A^2}{s_R^2} \ge F_\alpha (m-1; N-m).$$

2) Оценка значимости коэффициентов уравнения регрессии

Оценка значимости коэффициентов проводится по критерию согласия Стьюдента.

Для определения наблюдаемых значений критерия для каждого из коэффициентов регрессионной модели, необходимо построить матрицу значений влияющих на эксперимент факторов $\|X\|$, причем: число строк матрицы равно числу проведенных экспериментов (одновременно — числу значений каждого из факторов); число столбцов — числу коэффициентов уравнения регрессии, включая свободный член b_0 . При этом первый столбец данной матрицы состоит из единиц, он предназначен для расчета коэффициента b_0 , второй и последующие — из значений факторов x_1 , x_2 и т.д.

Для нашего уравнения регрессии данная матрица будет состоять из двух столбцов: первый содержит единицы, а второй — значения x_i .

Далее необходимо рассчитать транспонированную матрицу \boldsymbol{X}^T и матрицу

$$\mathbf{C} = \left(\mathbf{X}^{\mathrm{T}} \cdot \mathbf{X}\right)^{-1}.$$

Наблюдаемое значение критерия для каждого из коэффициентов регрессии определяется по формуле:

$$t_{j-1} = \frac{\left|b_{j-1}\right|}{\sqrt{s_R^2 \cdot c_{jj}}},$$

где ј – номер строки и столбца матрицы С;

 s_R^2 — оценка дисперсии по остаточной группировке, определенной при расчете оценки значимости уравнения регрессии;

с_{јј} – соответствующий элемент матрицы С.

Критическое значение определяется выражением

$$t_{\alpha}(N-m)$$
.

В случае, если

$$t_{j-1} \ge t_{\alpha}(N-m),$$

коэффициент регрессии признается значимым. В противном случае, полученное значение коэффициента регрессии признается следствием ошибки, вызванной недостаточным объемом произведенной выборки данных. такое значение коэффициента регрессии в окончательную формулировку просто не записывается вместе со стоящим рядом обозначением фактора \mathbf{x}_i .

Фактически, на данном параграфе можно было бы и закончить учебное пособие, поскольку методика регрессионного анализа и есть то самое планирование эксперимента.

Чтобы пояснить данную мысль, необходимо вспомнить материал первого раздела. Модель эксперимента — уравнение функции отклика, желательно в явном виде. А функция отклика — это зависимость между получаемым значением параметра оптимизации и устанавливаемыми значениями уровней факторов.

Что мы получаем в результате регрессионного анализа? Зависимость результата опытов (читайте, параметра оптимизации) от некоторых влияющих на него характеристик (читайте, уровней факторов). Таким образом, мы достигли желаемого.

Однако, не стоит забывать, что построение модели эксперимента далеко не ограничивается линейной моделью, хотя именно на этом мы остановились в данном учебном пособии. Данный факт объясняется самой логикой повествования, заявленной в самом начале учебного пособия: данное пособие рассматривает лишь начальные, наиболее простые случаи использования того или иного инструментария. Более сложные варианты, в том числе и регрессионного анализа, можно рассмотреть самостоятельно по соответствующей литературе, либо прочитать вторую часть пособия.

4 Введение в факторные планы

Факторные планы, рассматриваемые ниже, позволяют упростить методику вычисления коэффициентов регрессионной модели эксперимента. Кроме того, факторные планы (см. дробный факторный эксперимент) позволяют сократить число опытов для построения модели эксперимента. При этом учитывается, где именно построенная модель потеряет чувствительность, т.е. где она не сможет оценить, благодаря каким компонентам происходит изменение значений отклика (в некоторых случаях подобный подход допустим).

Исходя из сказанного, изложение предложенного ниже материала ведется следующим образом. На примере полного факторного эксперимента показано, как можно рассчитать коэффициенты уравнения регрессии без использования метода наименьших квадратов или других методов.

Далее показывается лишь методика сокращения числа опытов и оценка потери чувствительности модели эксперимента. Расчет коэффициентов модели в случае дробных реплик будет производиться так же, как и для полного факторного эксперимента. Поэтому, читая параграфы, посвященные построению дробных реплик, следует помнить, что после построения реплик, определения систем смешивания производится оценка коэффициентов регрессионной модели по методике, описанной в параграфе 4.2.

Методика получения дробных реплик, равно как и методика оценки коэффициентов регрессионной модели приводятся на простейшем случае — для двухуровневых факторов.

4.1 Полный факторный эксперимент типа 2^k

Эксперимент, в котором реализуются все возможные сочетания уровней факторов, называется полным факторным экспериментом.

Наиболее простой вариант полного факторного эксперимента — эксперименты типа 2^k . Начнем изучение планов эксперимента и способов их построения именно с этого типа.

Эксперимент, в котором каждый из факторов имеет только два уровня, называется факторным экспериментом типа 2^k .

Зная число факторов, можно вычислить общее число экспериментов, которые необходимо провести в данном случае. Напомним, общее число опытов определяется по формуле

$$N = m^k$$
,

где т – число уровней фактора,

k – число факторов.

Тогда для полного факторного эксперимента данного типа общее число испытаний составит

$$N=2^k$$
.

Удобно представлять результаты априорных экспериментов в виде таблицы, каждый столбец которого соответствует значениям факторов, а каждая строка — различным опытам. Последний столбец такой таблицы отводится

под значения параметра оптимизации, которые он принимает при заданных значениях фактора. Такие таблицы называются матрицами планирования эксперимента или просто планами эксперимента, таблица 4.1. Каждый столбец матрицы называют вектор—столбцом, а каждую строку вектор—строкой.

№ опыта	Фак	торы	Буквенное обозначение	Параметр оптимизации
	X ₁	X ₂	ooosna ienne	у
I	-1	-1	(1)	У1
II	+1	-1	а	У2
III	-1	+1	b	У3
IV	+1	+1	ab	У4

Таблица 4.1 – Матрица планирования эксперимента 2²

Представленные в матрице результаты можно изобразить геометрически, рисунок 4.1. Для этого в области определения факторов находим основную

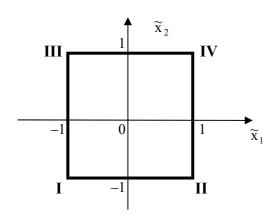


Рисунок 4.1. Геометрическое изображение матрицы 2^2 , представленной в таблице 4.1 (римскими цифрами обозначены номера опытов).

точку и проводим через нее новые оси координат, соответствующие перекодированным факторам. При этом область эксперимента пересекается осями в точках (+1) и (-1). Тогда условия проведения опытов будут соответствовать вершинам квадрата, центром которого является основной уровень, а стороны равны двум ин-

тервалам варьирования и параллельны осям факторов. Номера вершин квадрата соответствуют номерам опытов в матрице планирования.

Запись матрицы планирования, особенно для многих факторов, громоздка. Для ее сокращения вводят буквенные обозначения строк. Это делает-

ся следующим образом. Порядковый номер ставится в соответствие строчной букве латинского алфавита: $x_1 - a_1x_2 - b$ и т.д. Для каждой строки записывается своеобразный буквенный код. При этом соблюдается следующее правило: если фактор находится на верхнем уровне — буква ставится, в противном случае — нет; если все факторы находятся на нижнем уровне, вводится условное обозначение (1), таблица 4.1.

Если для эксперимента типа 2^k все возможные комбинации уровней легко найти простым перебором, то с ростом числа факторов появляется вероятность упустить из виду какое-либо состояние или продублировать его несколько раз. причем, чем больше факторов, тем выше эта вероятность. В результате возникает необходимость в разработке какого-либо алгоритма учета всех состояний системы. Таких алгоритмов несколько. Некоторые основаны на переходе от матриц меньших размерностей к матрицам более высоких размерностей. Рассмотрим их на примере перехода от планов 2^2 к планам 2^3 .

Способ №1. Метод перевода из низшей в более высокую размерность

При добавлении нового фактора каждая комбинация уровней исходного плана повторяется дважды: с верхним уровнем нового фактора и с его нижним уровнем. Отсюда возник следующий прием: записать матрицу меньшего размера, продублировать ее ниже, а затем для первого экземпляра исходной матрицы поставить верхний уровень нового фактора, а для второго экземпляра — нижний уровень нового фактора. Продемонстрируем данный прием на примере перехода $2^2 \rightarrow 2^3$, обозначив исходную матрицу во вновь сгенерированной, рисунок 4.2. Для простоты записи цифру «1» в обозначениях уровней факторов опустим.

№	X ₁	X ₂		N_2	X ₁	X ₂	X ₂	y
1	+	+		1	+	+	+	y ₁
2	_	+		2	_	+	+	y ₂
3	+	_		3	+	_	+	y ₃
4	_	_		4	_	_	+	y 4
	I	ı	1	5	+	+	_	y ₅
				6	_	+	_	y ₆
				7	+	_	_	y ₇
				8	_	_	_	y 8

Рисунок 4.2. Пример перевода матриц $2^2 \to 2^3$ по методу перевода из низшей размерности в более высокую

Способ №2. Метод перемножения

Как и в предыдущем случае, дважды вводим матрицу низшей размерности, одну под другой. Столбец нового фактора получаем по следующей процедуре. Для первой «маленькой» матрицы — перемножаем построчно значения факторов, а результат записываем в новый столбец. Для второй «маленькой» матрицы — перемножаем построчно значения факторов, а в новый столбец записываем результат, взятый с обратным знаком.

Продемонстрируем прием на примере перехода $2^2 \to 2^3$, обозначив исходную матрицу во вновь сгенерированной, рисунок 4.3.

\mathbf{x}_1	\mathbf{X}_{2}		№	$\mathbf{x_1}$	\mathbf{X}_{2}	\mathbf{X}_{2}	y
+	+		1	+	+	+	y ₁
_	+		2	_	+	_	y_2
+	_		3	+	_	_	y ₃
_	_		4	_	_	+	y ₄
		l	5	+	+	_	y ₅
			6	_	+	+	y 6
			7	+	_	+	y 7
			8	_	_	_	y_8
	+	+ + +	+ + + - +	+ + - + + - + - - - 4 5 6 7	+ + - + + - + - - - 5 + 6 - 7 +	$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$

Рисунок 4.3. Пример перевода матриц $2^2 \to 2^3$ по методу перемножения

Способ №3. Метод чередования знаков

Метод основан на следующей процедуре: в первом столбце знаки чередуются, во втором — вводятся попарно, в третьем — по четыре и т.д. Вообще число одинаковых знаков, идущих подряд определяется формулой 2^{k-1} , где k — номер фактора. В самом деле, для первого столбца $2^{1-1} = 2^0 = 1$ — подряд идет только по одному знаку, т.е. наблюдается чередование, для второго фактора $2^{2-1} = 2^1 = 2$ — идет по два одинаковых знака подряд (попарное расположение) ...

Продемонстрируем этот прием на примере матрицы 2^3 , рисунок 4.4.

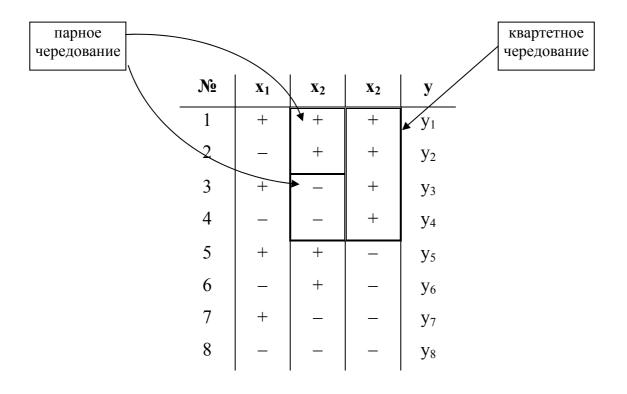


Рисунок 4.4. Построение матрицы 2^3 методом чередования

Какими же $\underline{ceoйcmeamu}$ обладают матрицы типа 2^k ?

1) Симметричность относительно центра эксперимента. Алгебраическая сумма элементов вектор-столбца каждого фактора равна нулю, т.е.

$$\sum_{i=1}^{N} x_{ji} = 0,$$

где ј – номер фактора,

N – число опытов, j = 1, 2, ..., k.

2) Условие нормировки. Сумма квадратов элементов каждого столбца равна числу опытов.

$$\sum_{i=1}^{N} x_{ji}^{2} = N.$$

3) Свойство ортогональности. Сумма по-членных произведений любых двух вектор-столбцов матрицы равна нулю

$$\sum_{i=1}^{N} x_{ji} x_{ui} = 0, \ j \neq u, j, u = 1, 2, ..., k.$$

4) Свойство ротатабельности. Точки в матрице планирования подбираются таким образом, что точность предсказания значений параметра оптимизации одинакова на равных расстояниях от центра эксперимента и не зависит от направления.

4.2 Полный факторный эксперимент и математическая модель эксперимента

Как уже говорилось ранее, для планирования эксперимента прежде всего необходима модель самого эксперимента и, как правило, математическая. В качестве таковой может рассматриваться то или иное уравнение, описывающее зависимость между значениями факторов и параметром оптимизации, т.е. функция отклика. Как правило, стараются выбрать линейную модель следующего вида

$$y = b_0 + b_1 x_1 + b_2 x_2 + \dots$$

В данном параграфе будем рассматривать эксперимент типа 2^k , т.е. математическая модель эксперимента имеет вид

$$y = b_0 + b_1 x_1 + b_2 x_2.$$

Цель работы с подобными моделями сводится к поиску неизвестных коэффициентов функции отклика. Ранее данную задачу мы решали методами регрессионного анализа, в частности, методом наименьших квадратов. Используя матрицу планирования, процедуру поиска коэффициентов можно упростить — они вычисляются по формуле

$$b_{j} = \frac{\sum_{i=1}^{N} x_{ji} y_{i}}{N}, \quad j = 1, 2, ..., k.$$
 (*)

Например, рассчитаем коэффициенты b_1 и b_2 для матрицы № 1, таблица 4.2

$$b_{1} = \frac{(-1)y_{1} + (+1)y_{2} + (-1)y_{3} + (+1)y_{4}}{4},$$

$$b_{2} = \frac{(-1)y_{1} + (-1)y_{2} + (+1)y_{3} + (+1)y_{4}}{4}.$$

Таблица 4.2 – Матрица планирования № 1

$N_{\underline{0}}$	\mathbf{x}_0	\mathbf{x}_1	\mathbf{x}_2	у
1	+1	-1	-1	y ₁
2	+1	+1	-1	y_2
3	+1	-1	+1	y ₃
4	+1	+1	+1	y 4

Таким образом, благодаря кодированию факторов, процедура вычисления коэффициентов значительно упростилась. Как же найти третий коэффициент, b_0 ? Если уравнение $y = b_0 + b_1x_1 + b_2x_2$ справедливо, то оно справедливо и для средних значений, т.е. $\overline{y} = b_0 + b_1\overline{x}_1 + b_2\overline{x}_2$. В силу свойства симметрии матрицы $\overline{x}_1 = \overline{x}_2 = 0$. Следовательно, $\overline{y} = b_0$. Чтобы привести формулу для вычисления b_0 в соответствие с формулой (*), в матрицу планирования удобно ввести вектор-столбец фиктивной переменной x_0 , которая во всех опытах приобретает значение +1 (см. выше). Тогда, формула (*) примет вид

$$b_{j} = \frac{\sum_{i=1}^{N} x_{ji} y_{i}}{N}, \quad j = 0, 1, ..., k,$$

а формула линейной модели — $y = b_0x_0 + b_1x_1 + b_2x_2$. Коэффициенты при независимых переменных указывают на силу влияния фактора. Чем больше численная величина коэффициента, тем большее влияние оказывает фактор. Если коэффициент имеет знак плюс, то между данным фактором и параметром оптимизации наблюдается прямая связь, т.е. при росте фактора возрастает и параметр оптимизации. Если же коэффициент имеет знак минус, то между данным фактором и параметром оптимизации обратная связь, т.е. при росте фактора параметр оптимизации уменьшается. Величина коэффициента соответствует вкладу данного фактора в величину параметра оптимизации при переходе фактора с нулевого на верхний или нижний уровень.

Иногда удобно оценивать вклад фактора при переходе от нижнего к верхнему уровню. Вклад, определенный таким образом, называется эффектом фактора (основным или главным эффектом). Численно он равен удвоенному коэффициенту.

Планируя эксперимент, на первом этапе мы стремимся получить линейную модель. Однако нет никакой гарантии, что в выбранных интервалах варьирования процесс описывается линейной моделью. Один из наиболее часто встречающихся видов нелинейности связан с тем, что эффект одного фактора зависит от уровня, на котором находится другой фактор. В этом случае говорят, что имеет место эффект взаимодействия двух факторов. Полный факторный эксперимент позволяет численно оценивать эффекты взаимодействия. Для этого, пользуясь правилом перемножения, получаем векторстолбец x_1x_2 (см. ниже), при вычислении коэффициентов взаимодействия пользуемся уже проверенной формулой (*), в которую в качестве значений факторов подставляем новый столбец. Например, имеем матрицу № 2, таблица 4.3:

Таблица 4.3 — Матрица планирования N_{2} 2

№	\mathbf{x}_0	\mathbf{x}_1	X ₂	x_1x_2	у
				+1	y ₁
2	+1	+1	-1	-1	y_2
3	+1	-1	+1	-1	y ₃
4	+1	+1	+1	+1	y ₄

Модель такой матрицы будет выглядеть следующим образом

$$y = b_0 x_0 + b_1 x_1 + b_2 x_2 + b_{12} x_1 x_2.$$

Вычислим коэффициенты по формуле (*) с учетом всего выше сказанного. Тогда, формулы для вычисления коэффициентов модели имеют вид:

$$b_0 = \frac{(+1)y_1 + (+1)y_2 + (+1)y_3 + (+1)y_4}{4},$$

$$b_1 = \frac{(-1)y_1 + (+1)y_2 + (-1)y_3 + (+1)y_4}{4},$$

$$b_2 = \frac{(-1)y_1 + (-1)y_2 + (+1)y_3 + (+1)y_4}{4},$$

$$b_{12} = \frac{(+1)y_1 + (-1)y_2 + (-1)y_3 + (+1)y_4}{4}.$$

Столбцы x_1 и x_2 задают планирование — по ним непосредственно определяются условия опытов, а столбцы x_0 и x_1x_2 служат только для расчета, это вспомогательные столбцы.

Эффект взаимодействия x_1x_2 носит название эффекта первого порядка или парного эффекта. Соответственно, эффект взаимодействия $x_1x_2x_3$ носит название эффекта взаимодействия второго порядка или тройного эффекта и т.д. Вообще эффект взаимодействия максимального порядка в полном факторном эксперименте имеет порядок, на единицу меньший числа факторов.

Полное число всех возможных эффектов, включая b_0 , линейные эффекты и взаимодействия всех порядков, равно числу опытов полного факторного

эксперимента. Чтобы найти число возможных взаимодействий некоторого порядка, можно воспользоваться формулой

$$C_k^n = \frac{k!}{n!(k-n)!}$$

где k – число факторов, n – число элементов во взаимодействии.

4.3 Возвращение назад

Приостановим наш «бег» по факторным планам и вернемся несколько назад – в регрессионный анализ, а именно – в оценку значимости коэффициентов регрессионной модели.

Еще раз напомню, бегло, саму процедуру. Мы строим матрицу X, состоящую из столбцов со значениями: «1» — для первого столбца матрицы; x_{i1} , x_{i2} ... — для последующих столбцов матрицы. Затем проводим различные математические операции с этой матрицей и обратной транспонированной матрицей (X^T)⁻¹, и находим наблюдаемые значения критерия значимости коэффициентов регрессионной модели.

А теперь давайте посмотрим на только что пройденный нами параграф, а точнее — на таблицы 4.2 и 4.3. Если внимательно присмотреться, сравнить с материалом параграфа 3.3, то можно легко заметить, что матрица X представляет собой всего лишь запись вектор-столбцов матрицы планирования, принадлежащих значениям факторов, в не кодированном (т.е. не в виде «+1» и «-1») виде. При этом учитываются и столбец x_0 , и, если это необходимо, столбец x_1x_2 . При этом транспонированная матрица X^T легко получается, если мы нашу матрицу планирования «положим на правый бок». А дальше — дело техники и Ваших математических навыков.

Если сравнение двух материалов Вам удалось, то в будущем, думается, будет легче понять принцип образования матриц X и X^T , а, следовательно, и осуществить все необходимые процедуры.

Ну а теперь, вернемся к факторным планам и усложним себе задачу: попробуем сократить число опытов, необходимых для построения математической модели эксперимента, без потери адекватности самой модели. И в этом нам поможет *дробный факторный эксперименти* ($\mathcal{Д}\Phi$ 3).

4.4 Дробный факторный эксперимент типа 2^{k-p} : выбор полуреплик

4.4.1 Основные определения дробного факторного эксперимента

Прежде, чем говорить о дробном факторном эксперименте, построении дробных реплик и т.п., договоримся о терминологии. Сами факторы, оказывающие влияние на параметр оптимизации, в дальнейшем называются эффектами или главными эффектами. Взаимодействия между двумя факторами называются двухфакторными (парными) или взаимодействиями первого порядка. Взаимодействия между тремя факторами, соответственно, называются взаимодействиями второго порядка или трехфакторными (тройными). Элементы типа х² (квадрат фактора), х³... называются квадратичными, кубическими и т.д. эффектами

Дробный факторный эксперимент (далее — ДФЭ) применяется для тех же целей, что и полный факторный эксперимент, т.е. для облегчения поиска коэффициентов математической модели, которые ранее решались методами регрессионного анализа. Однако, в довесок к этой проблеме ДФЭ решает еще

одну немаловажную задачу – уменьшение числа опытов, необходимых для планирования эксперимента.

Решение этой проблемы достигается путем переобозначения вектор— столбца матрицы планирования, содержащем незначительное (по предварительно проведенным экспериментам) взаимодействие факторов, как нового фактора. При этом новая матрица планирования не теряет своих свойств, описанных ранее. Полученная часть матрицы планирования называется репликой или, точнее, дробной репликой матрицы планирования.

Достоинства такого подхода очевидны: 1) уменьшается число экспериментов, проводимых для планирования; 2) больше внимания уделяется тем эффектам, которые оказывают наибольшее влияние на изменение параметра оптимизации. Однако, у подобного метода решения проблемы есть и недостатки, среди которых одним из основных является следующий: при переобозначении вектор—столбцов происходит смешивание эффектов различного порядка (об этом скажем ниже), в результате чего получаемая математическая модель становится нечувствительной к вкладам, вносимым смешиваемыми эффектами по-отдельности.

Необходимо отметить также, что любая дробная реплика строится на основе *полного факторного плана* матрицы более низкой размерности.

В зависимости от того, во сколько раз при образовании дробной реплики сокращается число опытов в полном факторном эксперименте (далее – ПФЭ), различают реплики различной дробности:

Число экспериментов сокращается:Уровень дробности репликив 2 раза1/2-реплика (полуреплика)в 4 раза1/4-реплика (четверть-реплика)в 8 раз1/8-реплика

и т.д.

Рассмотрим методику образования дробных реплик, их особенности на конкретном примере.

Запишем матрицу планирования эксперимента 2^2 с учетом взаимодействия между факторами, таблица 4.4.

№ опыта	\mathbf{x}_1	\mathbf{x}_2	X_3 X_1X_2	у
1	+	+	+	y ₁
2	_	+	_	y_2
3	+	_	_	y ₃
4	_	_	+	V ₄

Таблица 4.4 – Матрица планирования эксперимента с учетом взаимодействия

Результаты эксперимента, а, следовательно, и математическую модель эксперимента, можно представить в виде уравнения

$$y = b_0 + b_1 x_1 + b_2 x_2 + b_{12} x_1 x_2.$$

Предположим, что из априорной информации известно, что в выбранных интервалах варьирования процесс может быть описан линейной моделью, т.е. вклад в изменчивость параметра оптимизации от взаимодействия факторов очень незначителен. В таком случае вектор-столбец матрицы планирования x_1x_2 остается «свободным». Тогда его можно использовать для минимизации числа опытов, присвоив ему значение фактора x_3 .

В результате получаем четыре опыта для оценки влияния трех факторов, т.е. *половину* $\Pi\Phi\Theta-2^3$. Таким образом, полученная нами матрица представляет собой *полуреплику* $\Pi\Phi\Theta-2^3$.

Подобным образом можно получать реплики различной дробности для сокращения числа опытов полных факторных экспериментов типа 2^k за исключением ТОЛЬКО ПФЭ -2^2 . Отдельно отметим, что для ПФЭ -2^3 возможно получение ТОЛЬКО полуреплики, реплики большей дробности можно получить для экспериментов 2^k , где k > 3.

Для обозначения дробных реплик предусмотрено специальное обозначение:

$$2^{k-p}, (*)$$

где k – число факторов полного факторного эксперимента, для *которого* строится дробная реплика;

р – число взаимодействий, которыми пренебрегаем при построении дробной реплики.

Для нашего примера: число факторов в ПФЭ, для которого строится реплика равно 3, число взаимодействий, которыми пренебрегли равно 1. Тогда обозначение нашей полуреплики имеет вид 2^{3-1} .

Обозначение полуреплик имеет под собой несколько смыслов, один из которых мы только что привели. Есть и другие скрытые «резервы» у этого обозначения. Если уметь его правильно читать, можно выяснить и уровень дробности реплики, и на основе какого плана меньшей размерности будет строиться дробная реплика ПФЭ более высокой размерности. Покажем эти моменты.

Уровень дробности реплики можно в общем случае записать формулой $1/2^p$. Если в показателе степени у формулы (*) произвести вычитание k–p, то полученный показатель степени укажет, какой ПФЭ–план низшей размерности берется за основу при построении $1/2^p$ – реплики ПФЭ большей размерности.

Поясним данные моменты на нашем примере. Нами, как установлено выше была получена полуреплика $\Pi\Phi \ni -2^3$, которая обозначается как 2^{3-1} . В данном обозначении k=3, p=1. Дробность реплики, как сказано в предыдущем абзаце устанавливается формулой $1/2^p$. Применяя к нашему примеру, получим, что 2^{3-1} есть $1/2^1$ —реплика, т.е. *полуреплика*. Построение этой полуреплики производится на основе плана $\Pi\Phi \ni 2^{3-1}=2^2$, что и было проделано нами ранее.

Таким образом, можно сказать, что обозначение дробных реплик (*) предоставляет экспериментатору максимально возможную информацию о способе образования и разрядности дробной реплики.

4.4.2 Определяющий контраст, генерирующее соотношение. Планы с разрешающей способностью III

Можно заметить, что образование нашей полуреплики велось с использованием для нового фактора x_3 простого математического уравнения: $x_3 = x_1x_2$. Нетрудно видеть, что, если к этой матрице ниже дописать такую же полуреплику, но с использованием уравнения $x_3 = -x_1x_2$, то получим матрицу $\Pi\Phi \mathcal{F} - 2^3$. Таким образом, в зависимости от того, какое из двух упомянутых уравнений будем использовать для образования вектор—столбца нового фактора, получим различные части матрицы полного факторного эксперимента. Эти два уравнения можно переписать иначе, домножив обе части каждого из уравнений на «новый» фактор x_3 :

$$x_3 = x_1x_2$$
 \rightarrow $x_1x_2x_3 = +1$
 $x_3 = -x_1x_2$ \rightarrow $x_1x_2x_3 = -1$

Символическое обозначение произведения столбцов, равное +1 или -1, называется *определяющим контрастом*. Помимо того, чтобы указывать на ту часть матрицы ПФЭ, которая представлена в дробной реплике, определяющий контраст помогает определять какие эффекты будут смешаны между собой. Для этого необходимо умножить обе части определяющего контраста на столбец, соответствующий данному эффекту.

В нашем примере используется первый определяющий контраст. Используя его, определим, каким образом смешиваются эффекты в нашей математической модели. Для этого соответствующий определяющий контраст

будем последовательно умножать на эффекты x_1 , x_2 , x_3 . Напротив каждого полученного уравнения запишем, оценками каких эффектов являются коэффициенты нашей математической модели. При этом латинскими буквами b_i будем обозначать коэффициенты математической модели, а греческими буквами β_i — истинные значения вкладов каждого эффекта в изменчивость параметра оптимизации.

$$x_{1} = x_{2}x_{3}$$

$$x_{2} = x_{1}x_{3}$$

$$x_{3} = x_{1}x_{2}$$

$$b_{1} \rightarrow \beta_{1} + \beta_{23}$$

$$b_{2} \rightarrow \beta_{2} + \beta_{13}$$

$$b_{3} \rightarrow \beta_{3} + \beta_{12}$$

Соотношение, показывающее, с каким из эффектов смешан данный эффект, называется *генерирующим соотношением*. Т.е. выражения представляют собой генерирующие соотношения.

Что означают подобные соотношения? Фактически, эта запись показывает, что наша математическая модель проводит следующую замену:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2$$

$$b_1 x_1 + \beta_2 x_2 + \beta_{13} x_1 x_3 + \beta_{12} x_1 x_2$$

$$b_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2$$

в результате чего на свет появляется уравнение

$$y = \beta_0 + b_1 x_1 + b_2 x_2 + b_3 x_3.$$

«Чистым» в такой модели остается только свободный коэффициент. Однако, если помимо простого взаимодействия в математической модели наблюда-

ются и квадратичные коэффициенты, то свободный коэффициент будет смешиваться с ним.

С практической точки зрения подобные замены могут означать только одно: построенная таким образом математическая модель будет нечувствительна к оценке, какое именно взаимодействие дало тот или иной вклад в изменение параметра оптимизации – главный эффект (т.е. сам фактор) или же взаимодействие эффектов друг с другом. В том случае, когда взаимодействие эффектов носит незначительный характер, ничего фатального не произойдет. Однако, если хотя бы один из эффектов взаимодействия играет существенную роль, математическая модель, определенная таким образом, будет выдавать ошибку в прогнозировании результатов. Подобные ошибки могут быть как не заслуживающими внимание, так и фатальными. Поэтому очень важно знать, где именно модель будет наиболее уязвима в плане достоверности.

Вернемся к нашему примеру и построенной полуреплике. Из генерирующих соотношений видно, что при построении полуреплики 2^{3-1} происходит смешивание основных эффектов с эффектами взаимодействия первого порядка. Подобные планы называются *планами с разрешающей способностью* III. В общем случае разрешающая способность плана определяется по наибольшему числу факторов в определяющем контрасте. Обозначение нашей полуреплики в таком случае будет иметь вид 2^{3-1}_{III} – полуреплика $\Pi\Phi$ Э— 2^3 с разрешающей способностью III.

Еще раз отметим особенность планов с разрешающей способностью III. Не смешанным остается только свободный коэффициент, главные эффекты смешиваются с эффектами взаимодействия первого порядка.

4.43 Планы с разрешающей способностью IV, V

Попробуем произвести построение полуреплик для $\Pi\Phi \ni -2^4$. При этом в качестве основы для построения полуреплик выбирается $\Pi\Phi \ni -2^3$. Запишем математическую модель для $\Pi\Phi \ni -2^3$, учитывая парные и тройные взаимодействия:

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_{12}x_1x_2 + b_{13}x_1x_3 + b_{23}x_2x_3 + b_{123}x_1x_2x_3.$$

Для образования полуреплики необходимо пренебречь одним из эффектов взаимодействия. Очевидно, что, учитывая определяющий контраст, получим 8 вариантов:

контраст +1	контраст –1	Разрешающая способность
		_ плана
$\mathbf{x}_4 = \mathbf{x}_1 \mathbf{x}_2$	$\mathbf{x}_4 = -\mathbf{x}_1 \mathbf{x}_2$	
$\mathbf{x}_4 = \mathbf{x}_1 \mathbf{x}_3$	$\mathbf{x}_4 = -\mathbf{x}_1 \mathbf{x}_3$	> III
$\mathbf{x}_4 = \mathbf{x}_2 \mathbf{x}_3$	$\mathbf{x}_4 = -\mathbf{x}_2 \mathbf{x}_3$	
$\mathbf{x}_4 = \mathbf{x}_1 \mathbf{x}_2 \mathbf{x}_3$	$x_4 = -x_1 x_2 x_3$	IV

Нетрудно заметить, что две последних реплики имеют максимальную разрешающую способность и называются *главными*.

При отсутствии априорной информации об эффектах взаимодействия экспериментатор стремится выбрать реплику с наибольшей разрешающей способностью, т.к., как правило, чем выше степень взаимодействия, тем менее оно важно.

Определим систему смешивания для одной из главных полуреплик. Выберем определяющий контраст $x_1x_2x_3x_4 = +1$. Получим

$$x_1 = x_2x_3x_4$$
 $b_1 \rightarrow \beta_1 + \beta_{234}$
 $x_2 = x_1x_3x_4$ $b_2 \rightarrow \beta_2 + \beta_{134}$
 $x_3 = x_1x_2x_4$ $b_3 \rightarrow \beta_3 + \beta_{124}$
 $x_1x_2 = x_3x_4$ $b_{12} \rightarrow \beta_{12} + \beta_{34}$
 $x_1x_3 = x_2x_4$ $b_{13} \rightarrow \beta_{13} + \beta_{24}$
 $x_1x_4 = x_2x_3$ $b_{14} \rightarrow \beta_{14} + \beta_{23}$

Таким образом, можно заметить, что:

- 1) ни один из главных эффектов не смешан с другим главным эффектом или парным взаимодействием;
- 2) все парные взаимодействия смешаны друг с другом.

Два перечисленных момента представляют собой особенности и, одновременно, определение планов с разрешающей способностью IV.

Построим полуреплику, заданную определяющим контрастом $x_1x_2x_3x_4 = +1$. При этом вспомним правила буквенных обозначений строк матрицы и введем его. Тогда получим:

\mathcal{N}_{2}	\mathbf{x}_1	\mathbf{x}_2	X ₃	X ₄	Буквенное
	(a)	(b)	(c)	(d)	обозначение
1	+	+	+	+	abcd
2	_	+	+	_	bc
3	+	_	+	_	ac
4	_	_	+	+	cd
5	+	+	_	_	ab
6	_	+	_	+	bd
7	+	_	_	+	ad
8	_	_	_	_	(1)

т.е. все комбинации являются четными. Нетрудно убедиться, что при построении полуреплики, задаваемой определяющим контрастом $x_1x_2x_3x_4 = -1$, получим матрицу, задаваемую нечетными комбинациями: a, b, c, d, abc, abd, acd, bcd. Такие полуреплики называют *главными полурепликами*, отчасти потому, что они обладают максимальной разрешающей способностью.

Рассмотрим построение полуреплики 2^{5-1} . В нашем распоряжении будет множество вариантов, достаточно вспомнить уравнение $\Pi\Phi \mathcal{F} - 2^4$

$$y = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + b_4 x_4 + b_{12} x_1 x_2 + b_{13} x_1 x_3 + b_{14} x_1 x_4 + b_{23} x_2 x_3 + b_{24} x_2 x_4 + b_{34} x_3 x_4 + b_{123} x_1 x_2 x_3 + b_{134} x_1 x_3 x_4 + b_{234} x_2 x_3 x_4 + b_{124} x_1 x_2 x_4 + b_{1234} x_1 x_2 x_3 x_4.$$

Тогда можно получить следующие варианты:

контраст +1	контраст —1	Разрешающая способность
$\mathbf{x}_5 = \mathbf{x}_1 \mathbf{x}_2$	$\mathbf{x}_5 = -\mathbf{x}_1 \mathbf{x}_2$	плана
$\mathbf{x}_5 = \mathbf{x}_1 \mathbf{x}_3$	$\mathbf{x}_5 = -\mathbf{x}_1 \mathbf{x}_3$	
$\mathbf{x}_5 = \mathbf{x}_1 \mathbf{x}_4$	$\mathbf{x}_5 = -\mathbf{x}_1 \mathbf{x}_4$	
$\mathbf{x}_5 = \mathbf{x}_2 \mathbf{x}_3$	$\mathbf{x}_5 = -\mathbf{x}_2 \mathbf{x}_3$	> III
$\mathbf{x}_5 = \mathbf{x}_2 \mathbf{x}_4$	$\mathbf{x}_5 = -\mathbf{x}_2 \mathbf{x}_4$	
$\mathbf{x}_5 = \mathbf{x}_3 \mathbf{x}_4$	$\mathbf{x}_5 = -\mathbf{x}_3 \mathbf{x}_4$	
$\mathbf{x}_5 = \mathbf{x}_1 \mathbf{x}_2 \mathbf{x}_3$	$\mathbf{x}_5 = -\mathbf{x}_1 \mathbf{x}_2 \mathbf{x}_3$	j
$\mathbf{x}_5 = \mathbf{x}_1 \mathbf{x}_2 \mathbf{x}_4$	$x_5 = -x_1 x_2 x_4$	137
$\mathbf{x}_5 = \mathbf{x}_1 \mathbf{x}_3 \mathbf{x}_4$	$\mathbf{x}_5 = -\mathbf{x}_1 \mathbf{x}_3 \mathbf{x}_4$	> IV
$\mathbf{x}_5 = \mathbf{x}_2 \mathbf{x}_3 \mathbf{x}_4$	$\mathbf{x}_5 = -\mathbf{x}_2 \mathbf{x}_3 \mathbf{x}_4$	J
$x_5 = x_1 x_2 x_3 x_4$	$x_5 = -x_1 x_2 x_3 x_4$	V

Главными полурепликами здесь будут являться две последних, поскольку обладают наибольшей разрешающей способностью. Возьмем полуреплику, определяющуюся контрастом $+1 = x_1x_2x_3x_4x_5$ и определим для нее систему смешивания

$b_2 \rightarrow \beta_2 + \beta_{1345}$
$b_3 \rightarrow \beta_3 + \beta_{1245}$
$b_4 \rightarrow \beta_4 + \beta_{1235}$
$b_5 \rightarrow \beta_5 + \beta_{1234}$
$b_{12} \rightarrow \beta_{12} + \beta_{345}$
$b_{13} \rightarrow \beta_{13} + \beta_{245}$
$b_{14} \rightarrow \beta_{14} + \beta_{235}$
$b_{15} \rightarrow \beta_{15} + \beta_{245}$
$b_{23} \rightarrow \beta_{23} + \beta_{145}$
$b_{24} \rightarrow \beta_{24} + \beta_{135}$
$b_{25} \rightarrow \beta_{25} + \beta_{134}$
$b_{34} \rightarrow \beta_{34} + \beta_{125}$
$b_{35} \rightarrow \beta_{35} + \beta_{124}$
$b_{45} \rightarrow \beta_{45} + \beta_{235}$

Как хорошо видно из системы смешивания, главные эффекты здесь смешаны со взаимодействиями третьего порядка, а взаимодействия первого порядка — со взаимодействиями второго порядка. Такая система смешивания является одновременно особенностью и определением планов с разрешающей способностью V.

Как правило, на полурепликах 2^{5-1} работа с репликами данной дробности заканчивается, поскольку использование полуреплик 2^{6-1} и т.д. не выгодно: число опытов большое, а система смешивания слишком сложная.

4.5 Выбор 1/4-реплик в ДФЭ-2^k. Обобщающий определяющий контраст

Ранее уже отмечалось, что с увеличением дробности реплики уменьшается число опытов. А вместе с этим осложняется и система смешивания.

При исследовании влияния, например, пяти факторов можно поставить не 16 опытов, а только 8, т.е. воспользоваться репликой 2^{5-2} . Подобный подход возможен в том случае, если переименовывать не один из столбцов взаимодействий, а два. При этом взаимодействия должны быть разного порядка. Учитывая данные требования, возможно 12 решений для создания четверть реплики:

$$x_4 = x_1x_2,$$
 $x_5 = x_1x_2x_3$
 $x_4 = x_1x_2,$ $x_5 = -x_1x_2x_3$
 $x_4 = -x_1x_2,$ $x_5 = x_1x_2x_3$
 $x_4 = -x_1x_2,$ $x_5 = -x_1x_2x_3$
 $x_4 = x_1x_3,$ $x_5 = x_1x_2x_3$
 $x_4 = x_1x_3,$ $x_5 = -x_1x_2x_3$
 $x_4 = -x_1x_3,$ $x_5 = -x_1x_2x_3$
 $x_4 = -x_1x_3,$ $x_5 = -x_1x_2x_3$
 $x_4 = -x_1x_3,$ $x_5 = -x_1x_2x_3$
 $x_4 = x_2x_3,$ $x_5 = -x_1x_2x_3$
 $x_4 = x_2x_3,$ $x_5 = -x_1x_2x_3$
 $x_4 = -x_2x_3,$ $x_5 = -x_1x_2x_3$

Допустим, для создания дробной реплики был выбран вариант

$$x_4 = x_1x_3, x_5 = x_1x_2x_3.$$

Тогда определяющими контрастами являются

$$+1 = x_1x_3x_4, +1 = x_1x_2x_3x_5.$$

Если перемножить эти определяющие контрасты, получится третье соотношение, задающее элементы столбца $+1 = x_2x_4x_5$. Чтобы полностью охарактеризовать разрешающую способность реплики, необходимо записать *обобщающий определяющий контраст* $+1 = x_1x_3x_4 = x_2x_4x_5 = x_1x_2x_3x_5$. В этом случае система смешивания определяется умножением обобщающего определяющего контраста последовательно на главные эффекты. В результате система смешивания получается довольно сложной.

$$\begin{array}{lll} x_1 = x_3x_4 = x_2x_3x_5 = x_1x_2x_4x_5 & b_1 \rightarrow \beta_1 + \beta_{34} + \beta_{235} + \beta_{1245} \\ x_2 = x_1x_2x_3x_4 = x_4x_5 = x_1x_3x_5 & b_2 \rightarrow \beta_2 + \beta_{45} + \beta_{135} + \beta_{1234} \\ x_3 = x_1x_4 = x_2x_3x_4x_5 = x_1x_2x_5 & b_3 \rightarrow \beta_3 + \beta_{14} + \beta_{125} + \beta_{2345} \\ x_4 = x_1x_3 = x_2x_5 = x_1x_2x_3x_4x_5 & b_4 \rightarrow \beta_4 + \beta_{13} + \beta_{25} + \beta_{12345} \\ x_5 = x_1x_3x_4x_5 = x_2x_4 = x_1x_2x_3 & b_5 \rightarrow \beta_5 + \beta_{24} + \beta_{123} + \beta_{1345} \\ x_1x_2 = x_2x_3x_4 = x_1x_4x_5 = x_3x_5 & b_{12} \rightarrow \beta_{12} + \beta_{35} + \beta_{145} + \beta_{234} \\ x_1x_5 = x_3x_4x_5 = x_1x_2x_4 = x_2x_3 & b_{13} \rightarrow \beta_{15} + \beta_{23} + \beta_{124} + \beta_{345} \\ x_1x_3x_4 = x_2x_4x_5 = x_1x_2x_3x_5 & b_{134} \rightarrow \beta_{134} + \beta_{245} + \beta_{1235} \end{array}$$

Как можно заметить из системы смешивания, практически все эффекты перемешиваются между собой, за исключением свободного члена. Таким образом, модель теряет чувствительность к тому, какой именно эффект и какой степени дает вклад в изменчивость параметра оптимизации. Однако, если все

эффекты, кроме главных, особого значения не имеют, можно и воспользоваться данной моделью. В противном случае необходимо достроить четверть-реплику до полуреплики или полного факторного эксперимента, чтобы освободить необходимые эффекты взаимодействия и повысить чувствительность построенной модели.

Дополнительно отметим тот факт, что построение реплик большей дробности ведется подобно описанному здесь принципу четверть-реплик, т.е. строится обобщающий определяющий контраст, определяется система смешивания, затем — строится матрица планирования и т.д. Единственное, что будет отличаться — при увеличении степени дробности система смешивания будет все сложнее, а полученная модель — все менее чувствительна.

На этой оптимистичной ноте позвольте нам остановить Ваше знакомство с факторными планами. Конечно, это лишь самые азы. Однако, «перекормив» Вас информацией, мы получим обратный рефлекс: вместо усвоения материала и готовности продолжать, мы лишь внушим ужас.

Поэтому, дав лишь понятия о факторных планах и основных приемах работы с ними, мы адресуем Вас ко второй части учебного пособия и представленной ниже литературе, как только Вы будете готовы их осилить.