

Федеральное агентство по образованию

САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ
ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ

В. А. Лунев

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ И ПЛАНИРОВАНИЕ ЭКСПЕРИМЕНТА

Допущено учебно-методическим объединением
по образованию в области металлургии в качестве
учебного пособия для студентов высших
учебных заведений, обучающихся по
направлению «Металлургия»

Санкт-Петербург
Издательство Политехнического университета
2012

УДК 519.242; 621.43 (075.8)
ББК 34.3, в 6
Л847

Р е ц е н з е н т ы :

Доктор технических наук, профессор СПбГПУ *Н. Г. Колбасников*

Доктор технических наук, профессор СПб ГИЭУ *В. Д. Дурнев*

Лунев В. А. Математическое моделирование и планирование эксперимента: Учеб. пособие. СПб., 153 с.

Рассмотрены вопросы подготовки эксперимента с использованием теории подобия и фракционного анализа, основные определения и теоремы, методы обработки экспериментальных кривых и приемы построения интерполяционных полиномов; основы планирования эксперимента с целью получения модели процесса с минимальными затратами, математические аспекты оценок ее адекватности. Главное внимание уделено современным методам планирования эксперимента и технике обработки результатов. Предназначено для студентов технологических специальностей университета, готовящих стать специалистами-технологами на металлургических, машиностроительных и других предприятий. Может быть использовано также инженерным персоналом предприятий в системе повышения квалификации. Табл. 32. Ил. 41. Библиогр.: 13 назв. Печатается по решению редакционно-издательского совета Санкт-Петербургского государственного политехнического университета.

ISBN 5-7422-1023-X

© Лунев В. А.

© Санкт-Петербургский государственный политехнический университет, 2012

ОГЛАВЛЕНИЕ

Введение	5
1. Общие сведения о технологическом эксперименте	8
1.1. Формы моделирования, типы моделей	8
1.2. Кибернетическое представление модели	11
1.3. Геометрическое представление модели	13
1.4. Цели эксперимента	13
1.5. Типы эксперимента	15
<i>Контрольные вопросы</i>	18
2. Основные понятия теории подобия и размерностей	19
1.2. Характеристика входных и выходного параметров	19
2.2. Теорема о существенных параметрах	21
2.3. Размерные и безразмерные величины. Зависимые и независимые размерности	23
2.4. Структура связей между параметрами	24
2.5. П-теорема подобия	26
2.6. Критерии процесса	30
<i>Контрольные вопросы</i>	33
3. Основы физического моделирования ...	33
3.1. Теоремы “классической” теории подобия	33
3.2. Физическое моделирование	35
3.3. Фракционный анализ	37
3.4. Процедуры фракционного анализа	38
3.5. Возможности фракционного анализа	41
3.6. Естественные координаты системы	46
3.7. Общие выводы по теории подобия и фракционному анализу	49
<i>Контрольные вопросы</i>	50
4. Статистическая обработка результатов измерений	51
4.1. Нормальный закон распределения	51
4.2. Полигон распределения и гистограмма	54
4.3. Проверка нормальности закона распределения	56
4.4. Оценка статистических характеристик	58
4.5. Статистические операции	59
4.6. Построение экспериментальных кривых	64
<i>Упражнения</i>	68
<i>Контрольные вопросы</i>	69
5. Интерполяция	69
5.1. Постановка задачи	69
5.2. Теорема о количестве членов в уравнении	70
5.3. Построение интерполяционного уравнения методом Ньютона	72
<i>Упражнения</i>	74
5.4. Интерполяционные полиномы Лагранжа	74
5.5. Кубические сплайны	76
<i>Контрольные вопросы</i>	79

6. Дисперсионный анализ	80
6.1. Задачи дисперсионного анализа	80
6.2. Основные уравнения дисперсионного анализа	82
6.3. Общий анализ дисперсионной модели	87
6.4. Дисперсионный анализ количественных и качественных факторов	89
6.5. Задача	93
6.7. Латинские квадраты	97
6.8. Латинские кубы, греко-латинские квадраты, кубы и гиперкубы	100
<i>Контрольные вопросы</i>	102
7. Регрессионный анализ	103
7.1. Задачи регрессионного анализа	103
7.2. Построение регрессии	103
7.3. Обращенная матрица. Планирование эксперимента	106
7.4. Ортогональные планы первого порядка	109
7.5. Планы 2^k	114
7.6. Дробные реплики	116
7.7. Симплекс-планы. Планы Плакетта-Бермана	119
7.8. Дисперсионный анализ регрессионной модели	121
7.9. Центральный композиционный план	124
7.10. Планирование второго порядка	126
7.11. Пример построения рототабельного плана	128
7.12 Симплекс-решетчатый план второго порядка	133
7.13. Отсеивающие эксперименты. Метод случайного баланса	133
7.14. Экстремальные эксперименты. Крутое восхождение	135
7.15. Адаптивные модели в системах управления	141
7.16. Корреляционный анализ	143
<i>Контрольные вопросы</i>	149
Библиографический список	150
Приложения	151

ВВЕДЕНИЕ

Деятельность человека неразрывно связана с экспериментом. Экспериментом следует называть любую деятельность, которая направлена на получение некоторого ожидаемого результата. Необходимо различать эксперименты, которые проводят натуралисты и технологи. Натуралисты только наблюдают природу или объект. Например, орнитологи изучают поведение птиц, фиксируют, окольцовывают, ведут дневники и, но чем меньше они вмешиваются в жизнь птиц, тем интереснее и полнее будет исследование. Они не ставят своей целью изменить поведение птиц. Технологические эксперименты, наоборот, предполагают активное вмешательство исследователя в структуру исследуемого объекта. Цель его эксперимента – изучить, описать изучаемый процесс с тем, чтобы изменить, улучшить структуру объекта и его технологические, экономические и другие характеристики. Таким образом, мы выделяем технологический эксперимент в отдельный класс творческой деятельности человека.

Замечательный американский ученый-экономист У. Э. Деминг показал, что такая деятельность проходит в четыре стадии. На первой осуществляется планирование эксперимента, определяются цели и методы достижения результата. Вторая стадия эксперимента заключается в его реализации, изучении и анализе полученных результатов. Во многих случаях научным результатом эксперимента является некоторое описание изучаемого объекта, или, как говорят, модель процесса. В философской теории познания показано, что человек познает не сам объект, а некоторую его модель, некоторое представление этого объекта, объект же в целом он охватить не в состоянии.

Затем следует третья стадия – контроль результата, проверка его воспроизводимости, устойчивости, достоверности. Происходит оценка качества научной продукции как «товара». Проверяется, насколько полученная модель соответствует реальному объекту. Наконец, наступает четвертая стадия – корректировка цели и плана эксперимента, подготовка очередного эксперимента, который также будет проходить в четыре стадии.

Деятельность инженера-исследователя и инженера-технолога – это непрерывный эксперимент, протекающий по этому кругу. Директор предприятия ежедневно осуществляет сначала сбор информации о работе предприятия, затем принимает решения и контролирует ход технологического процесса, непрерывно вводя корректировки и, наконец, продумывает мероприятия по совершенствованию производства. Он не просто наблюдает за деятельностью предприятия, а непрерывно вмешивается в нее с целью обеспечить нормальную работу сейчас и в перспективе. Турист, совершающий восхождение на вершину горы также обязан сначала спланировать маршрут и технологию восхождения, затем приступить к реализации плана, постоянно сверяя реальность с планом и корректируя свои действия.

Циклический характер экспериментальной, производственной и других видов деятельности человека, видимо, является неким общим законом природы такого рода деятельности – законом Деминга. В каждом из ее видов

различие состоит только в конкретных особенностях поставленной задачи и специфике конечного продукта: в научной деятельности конечной продукцией является научный результат или модель объекта, в производстве – материальные вещи, у туриста – покоренная вершина и т. д.

В настоящем пособии мы будем заниматься только технологическими экспериментами, которые объединены общей методологией их исследования. Это широкий класс экспериментов, направленных на исследование технологических систем с целью описания происходящих в них процессов, выявления закономерностей и зависимостей, определения границ изменения параметров и т. п., приводящих к заданному результату. В отличие от исследований натуралистов, в технологических экспериментах заранее известно, что ответ на поставленные вопросы всегда существует, и требуется только обеспечить его получение. В настоящее время планирование, анализ и обработка результатов технологического эксперимента выделяется в самостоятельное направление математической статистики и осуществляется по вполне научно обоснованным правилам на достаточно мощной математической базе. Научная методика проведения эксперимента позволяет значительно экономить время, материалы и средства на достижение желаемого результата. На многих предприятиях инженер не допускается к проведению сложных и дорогостоящих экспериментов, если он не владеет математическими основами его проведения.

Возможны два подхода к изучению любого объекта – теоретический и экспериментальный. Они неразрывно связаны и взаимно дополняют друг друга. Теоретический подход обычно связан с математическим описанием объекта в виде некоторой системы уравнений и поиском методов их решения. Математическое моделирование объекта осуществляется обычно на ЭВМ путем анализа или решения этих уравнений. Однако чисто математическое моделирование в конечном счете решает лишь математические проблемы, оно формально оперирует некоторыми количественными характеристиками объекта, не вдаваясь в физическую сущность процессов. Математическое моделирование с ее мощным математическим аппаратом – математической физикой и численными методами решения задач – в основе своей занимается не изучением физического объекта, а тех математических связей, которые в какой-то мере соответствуют упорядоченным связям между параметрами физического процесса. Эти связи формулируются самим же исследователем на основе некоторых следствий, вытекающих из основных физических положений. Они не всегда полны, не вовлекают во внимание некоторые сопутствующие факторы. В связи с этим при математическом решении задачи не все проблемы физического процесса могут быть охвачены. Кроме того, часто выводы из математического решения получаются настолько сложными, что практическое использование их затрудняется. Надежность этих выводов полностью зависит от правильности сформулированных основных предпосылок, и чем сложнее процесс, тем меньше надежность. Достоверность сведений, добытых методами математического моделирования, зависит полностью от правильности математических моделей, соответствия их

физическим процессам. Чем сложнее объекты, тем выше вероятность того, что математическая модель не соответствует объекту, тогда математическое моделирование превращается в некоторую игру ума. Поэтому многие социальные, экологические, психологические и подобные сложные математические модели нередко приводят к выводам, которые затем не реализуются на практике.

При экспериментальных методах исследования объектов эти недостатки в значительной мере отсутствуют, физическая сущность объекта реализуется в эксперименте. Исследователь неформально анализирует объект, замечая всякие тонкости, нюансы процессов. Конечно, чем сложнее объект, тем дороже эксперимент, и не всегда эксперимент возможен в том диапазоне, который необходим. Существуют временные и материальные ограничения. Поэтому почти всегда требуется проведение математического моделирования объекта с последующей экспериментальной проверкой результата путем правильно поставленного эксперимента.

Основой экспериментального изучения объекта являются теория подобия и теория планирования и обработки эксперимента. Они излагаются в данном учебном пособии как элементы некой системы, которая имеет единую и стройную теоретическую основу. Обычно в литературе теорию подобия применяют только для того, чтобы помочь выделить критерии процесса и избежать грубых ошибок при экспериментальном моделировании. Ее мощный раздел - фракционный анализ – практически слабо используется. В данном пособии эти разделы представлены как методология мышления при подготовке и проведении эксперимента.

Теория планирования эксперимента во многих книгах излагается в мало понятных математических терминах. Надеемся, что этот недостаток удалось устранить. Надо отметить также, что издается мало книг, раскрывающих возможности применения этой теории к технологическим задачам.

Книга задумана как учебное пособие для студентов всех технических специальностей. Независимо от специальности, инженер-технолог и исследователь, а также руководящий работник, обязаны непосредственно на производстве использовать идеи планируемого эксперимента. Это, с одной стороны, позволяет значительно экономить время и средства на постановку и проведение производственного эксперимента, минимальными средствами добывать необходимую информацию, с другой – дает возможность развивать и внедрять статистические методы исследования и оценки качества продукции, использовать научный аппарат теории планирования эксперимента как инструмент повышения качества выпускаемой продукции.

Во многих учебных заведениях вводят в ГОС науку о планировании и методах обработки технологического эксперимента. Более широкое распространение ее в учебных заведениях и на производстве сдерживается только отсутствием соответствующей литературы разного уровня сложности и недостаточной квалификацией научных и руководящих работников, определяющих стратегию развития производства.

1. Общие сведения о технологическом эксперименте

Формы моделирования, типы моделей.

Познание любого процесса, явления или объекта начинается с его моделирования, то есть построения некоторого представления о нем, в главных чертах приближенного к реальности. Важно подчеркнуть, что модель и реальный объект всегда не совпадают, человек не может иначе познавать мир, как через постоянное совершенствование своего представления о нем, никогда не достигнув полной идентификации.

Термин “модель” имеет два значения. В первом значении под моделью понимают уменьшенную копию натурального объекта, на которой можно путем экспериментального исследования получить необходимую информацию о натурном объекте. Такие модели называются физическими. На физических моделях можно «проиграть» процесс. Видимо, наилучшей физической моделью является сам натуральный объект, так как на уменьшенных моделях не исключена потеря некоторой важной информации. Основным теоретическим аппаратом физического моделирования служит теория подобия с ее основными разделами - теорией размерностей и фракционным анализом.

Во втором значении моделью называется некоторое мысленное представление изучаемого объекта. Это может быть его словесное описание, чертеж, набор символов, система математических выражений и т. п. Однако в технических науках и технологии объект считается математически описанным тогда, когда имеется уравнение или система уравнений, соответствующие физическим процессам, протекающим в объекте. Важно отметить, что такие уравнения основаны на некоторой математической концепции объекта, выдвинутой самим же исследователем, добытой им на основе обобщения экспериментального материала и литературных данных. А исследователь может ошибаться.

Математическая модель и математическое моделирование играют ведущую роль в современной методологии исследования подавляющего большинства технических процессов и объектов. При этом роль физического моделирования несколько не умалется. Чаще всего необходимо применение одновременно обоих методов, чтобы добиться наиболее достоверных результатов.

Математическая модель в общем виде представляет собой уравнение или систему уравнений типа

$$y = f(x_1, x_2, x_3, \dots, x_k) , \quad (1)$$

которая связывает параметры $x_1, x_2, x_3, \dots, x_k$ с параметром y .

Модель в записи (1) называется *параметрической*, так как содержит только перечень всех параметров, участвующих в исследовании, и никакой другой информационной нагрузки не несет. Исследователь может задавать значения параметров $x_1, x_2, x_3, \dots, x_k$ в определенных пределах, поэтому они называются входными. Параметр y является откликом на эти изменения, поэтому он называется откликом или выходным параметром. Естественно, час-

то желательна более информативная форма связи всех параметров процесса, чем (1).

Для описания процессов применяются математические уравнения или системы уравнений разного типа. Наиболее простыми являются алгебраические уравнения вида

$$\begin{aligned} y = & a_0 + a_1x_1 + a_2x_2 + \dots + a_kx_k + \\ & + a_{12}x_1x_2 + a_{13}x_1x_3 + \dots + a_{k,k-1}x_kx_{k-1} + \\ & + a_{11}x_1^2 + a_{22}x_2^2 + \dots + a_{kk}x_k^2 + \dots \end{aligned} \quad (2)$$

Уравнение любого вида можно разложить в бесконечный ряд Тейлора, совпадающий с формой (2), поэтому такое математическое представление объекта является общим.

Алгебраические уравнения применяются для установления формальной связи между входными и выходными параметрами, в них не зашифрованы какие-то внутренние физические связи между параметрами. Так можно связать даже два независимо протекающих процесса, если они оба определяются каким-то третьим процессом. Алгебраические уравнения просты, поэтому удобны в системах автоматического управления и при статистической обработке массивов информации. Даже в том случае, когда для процесса имеются более сложные уравнения, более наполненные физическим содержанием, полезно упростить их и представить в алгебраической форме.

Более содержательными являются дифференциальные уравнения, описывающие, как правило, движение материальных точек, их траекторию. Они уже связывают параметры на основании анализа некоторых внутренних механизмов процесса и могут иметь, например, вид

$$\frac{d^2y}{dx^2} = f_1(x) \cdot \frac{dy}{dx} + f_2(x) \quad (3)$$

Если движение осуществляется по трем координатам, то модель процесса будет состоять из системы трех уравнений подобного типа (при этом функции, входящие в уравнения, будут зависеть от трех координат, и производные станут частными). Для описания движения системы материальных точек, потоков, полей в движущейся среде применяют системы дифференциальных уравнений в частных производных. Основные задачи механики сплошной среды, аэрогидродинамики, тепловых и других процессов связаны с анализом уравнений этого типа. Методы решения их разрабатываются математической физикой. Результатом их решения являются обычно функции предыдущего типа, описывающие траектории движения материальных точек.

Наконец, в качестве математической модели могут выступать функционалы, описывающие общее состояние движущейся среды. Функционалами, то есть функциями от функций движения материальных точек, могут служить уравнения сохранения энергии движения, потенциальная энергия системы, количество движения и т. п. Функционал содержит в себе в концентрированном виде ту же самую информацию, что и система дифференциальных уравнений в частных производных. Такая система может быть получена

из функционала после соответствующих преобразований. Решения функционалов сводятся к нахождению их максимума или минимума, результатом являются те же уравнения траекторий движения материальных точек.

Представленные математические модели необходимы для аналитического описания процесса, поэтому их следует отнести к классу аналитических моделей. Помимо них применяются математические модели-аналоги, которые основаны на том, что некоторые процессы разной физической природы описываются одинаковыми математическими уравнениями. Например, распространение тепла в сплошной среде, растекание воды в водоеме, течение металла при волочении, распространение электрического тока на электропроводной бумаге происходят по аналогичным закономерностям и описываются одним и тем же уравнением. В связи с этим электрическая модель может быть применена для исследования тепловых и гидротехнических задач, для анализа деформации металла при волочении и т. п. На такой аналогии основано применение прибора ЭГДА (электро-гидродинамической аналогии).

Широкий класс математических моделей составляют имитационные модели. Они применяются как в исследовательских экспериментах, так и в системах управления объектами. Это динамические аналитические модели, обычно состоящие из системы уравнений, имеющих более сложную форму записи, чем (2), которые имитируют поведения объекта при изменении различных входных параметров.

Видимо, на этом не ограничивается все многообразие математических моделей, однако для технологических задач и построения технологических экспериментов, более сложные модели применять нецелесообразно.

В технике часто разграничивают понятия «модель» и «закон». И то, и другое – это уравнения, описывающие объект. Но под законом понимают детерминированную, а под моделью – статистическую связь между входными и выходными параметрами. При детерминированной связи изменение x_i всегда одинаково влияет на изменение величины y . Например, по закону Ома для электрической цепи постоянного тока

$$I = U / R..$$

Изменение напряжения в 2 раза непременно приведет к изменению силы тока ровно в 2 раза (при постоянном сопротивлении R). Если провести экспериментальное исследование, то приведенная связь между напряжением и силой тока будет подтверждена во всех измерениях (на 100 %).

В истории науки был длительный период, когда целью научного исследования было открытие закона. В механике, гидравлике, теплотехнике, аэродинамике и других важных прикладных науках были открыты законы Гука, Бернулли, Клайперона-Менделеева и др. Но этот период, видимо, исчерпал себя, все законы уже открыты. Современной науке приходится довольствоваться исследованием моделей, которые имеют статистическую, или вероятностную, то есть не 100 %-ную связь между входными и выходными параметрами. При одинаковом изменении параметра x_i в разных измерениях параметр y может измениться по-разному. Например, в одном измерении y мо-

жет принимать значение 2,0, в другом – 1,7, в третьем – 2,2, и т. п. В вероятностных моделях показания выходного параметра колеблются вокруг некоторого среднего значения, которое лишь с некоторой достоверностью (или вероятностью) может считаться истинным значением.

Строго говоря, многие законы также в основе своей вероятностны. Например, в законе Клайперона-Менделеева, устанавливающем связь между давлением p , объемом V и температурой T газа в замкнутом сосуде ($pV = nRT$), детерминированные свойства проявляются только потому, что на стенки сосуда воздействует всегда большое количество молекул, и давление, определяемое количеством таких соударений, становится постоянным. С уменьшением количества молекул в сосуде закон начинает проявлять вероятностные свойства.

Есть еще одно отличие между законом и моделью. Закон нетерпим к гипотезам. Если появляется новая гипотеза, утверждающая, что закон не верен, то между ними будет происходить борьба до тех пор, пока не восторжествует истина: либо гипотеза станет законом, либо выживет прежний закон. Двух законов Ома не может быть. Модель в этом смысле более покладиста. Она допускает сосуществование нескольких моделей. Если исследователю не понравится существующая формула (модель), он имеет право, не отвергая предыдущую, вывести свою формулу. Например, для расчета уширения полосы при прокатке сосуществуют более ста формул, каждая для каких-то определенных условий прокатки. Таким образом, после открытия всех законов, поле деятельности исследователей не сузилось, а, наоборот, расширилось, переместившись в область исследования статистических моделей.

Кибернетическое представление модели.

Сначала полезно уточнить используемые термины. В теории подобия, математическом моделировании и планировании эксперимента нет единой установившейся терминологии, отчего возможна путаница и непонимание.

Величиной в технике называют такую характеристику объекта (процесса, явления, системы), которую можно измерить. Среди величин всегда имеются такие, которые в процессе исследования изменяют свои значения, поэтому они называются переменными в отличие от постоянных величин или констант. Часто переменные самостоятельно или в виде комбинаций с другими величинами выступают в роли координат, используемых при построении графиков.

Параметр – это величина, определяющая геометрические, физические и другие характеристики системы, например, высота H , длина L , температура T , напряжение U и т. д. К числу параметров относятся и различные их комбинации, имеющие какой-либо смысл, например, отношение H/L .

Нередко в экспериментальных исследованиях применяют термин “фактор”, который мы будем считать синонимом термина “входной параметр” и будем употреблять оба термина равнозначно.

Вернемся к рассмотрению параметрического уравнения (1):

$$y = f(x_1, x_2, x_3, \dots, x_k) .$$

Наглядной формой представления этой модели служит кибернетический «ящик», приведенный на рис 1.

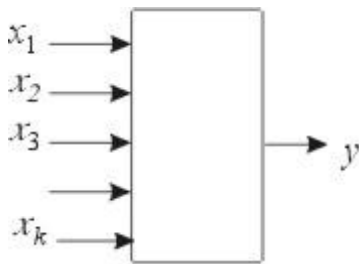


Рис 1. Модель «черного ящика».

Управляющие параметры $x_1, x_2, x_3, \dots, x_k$ входят в него по заданию и усмотрению исследователя, отсюда и названы входными, параметр y является выходным. Модель такого типа наглядно показывает, что значения входных параметров (факторов), “засыпаемых” в “ящик”, как на мельнице перемалываются в некоторое значение выходного параметра y . Математические выражения, осуществляющие это “перемалывание”, являются устройством мельницы. При наличии дифференциальных или иных уравнений связи между входными и выходными параметрами устанавливаются на основе физических закономерностей в изучаемых процессах. Если уравнения отсутствуют, то модель называют “черным ящиком”, в таких моделях связи можно установить только статистическим путем на основе эксперимента. При этом не раскрывается физическая природа таких связей. Чем сложнее объект, тем чаще мы имеем дело с “черным ящиком”, и вынуждены строить статистические модели. В основе такого построения лежат измерения.

Измерением называется один акт определения численного значения любого из входных x_i или выходного y параметров. Измерение осуществляется, как правило, в размерных единицах – м, кг, с, вольт и т. п.

Один акт измерения всех входных и выходного параметров (в уравнении (1)) называется опытом. В каждом опыте содержится только один определенный набор значений входных и выходного параметров. Иногда в опытах полезно указывать не только значения входных параметров, но и их порядковый номер по осевой шкале, то есть уровень изменения каждого фактора. В совокупности все опыты образуют эксперимент. Точкой эксперимента или узловой точкой (узлом) будем называть фиксированный набор входных параметров, соответствующий одному опыту. Таким образом, опыты проводятся в узловых точках эксперимента. В любой узловой точке можно провести несколько опытов. Такое дублирование необходимо для того, чтобы оценить воспроизводимость эксперимента и оценить степень разброса данных при воспроизводимости. Таким образом, для того, чтобы получить количественную оценку воспроизводимости эксперимента, необходимо хотя бы в одной точке (а, может быть, не только в одной и даже во всех) провести несколько опытов.

Геометрическое представление модели.

Параметрическую модель(1) можно представить в виде геометрического образа. Допустим, мы имеем дело только с одним входным параметром x_1 :

$$y = f(x_1).$$

Это однофакторный эксперимент, количество входных параметров $k = 1$. График представленной зависимости показан на рис. 2 а, который мы будем трактовать как поверхность в одномерном пространстве.

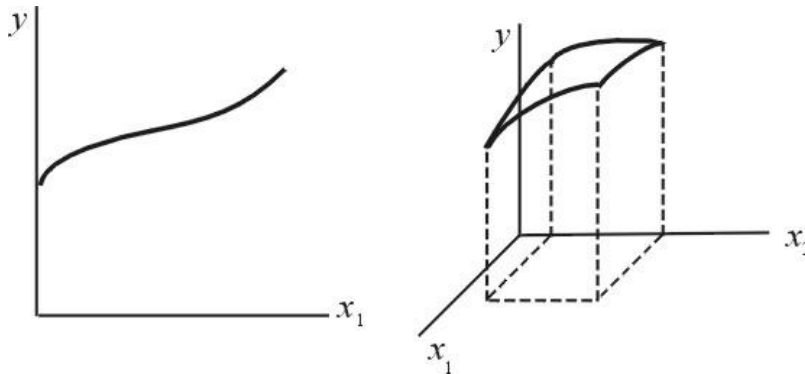


Рис. 2. Однофакторный (а) и двухфакторный (б) эксперименты.

На рис. 2, б количество факторов увеличено до двух: $k = 2$. Зависимость

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

представляет собой уравнение поверхности, нависшей над плоскостью координат $x_1 - x_2$, построенной по значениям y . В одномерном пространстве мы имели две взаимно перпендикулярные координатные оси, в двумерном пространстве таких осей три. В любом k мерном пространстве необходимо иметь k взаимно перпендикулярных осей для входных параметров и одну ось, перпендикулярную всем предыдущим, по которой будет откладываться значение отклика y . Полученная таким способом поверхность будет представлять собой гиперповерхность в k –мерном пространстве. Физически ее построить нельзя, но по аналогии с одномерным и двумерным случаями, можно представить в виде некоторой логической конструкции. Таким образом, параметрическое уравнение (1) представляет собой гиперповерхность в k –мерном пространстве, и в геометрическом плане задача эксперимента сводится к построению такой гиперповерхности.

1.4. Цели эксперимента.

Не претендуя на полноту, определим цели, ради которых проводятся экспериментальные исследования. Ограничимся только технологическими экспериментами, которые направлены на решение технико-экономических задач, поиск новых технологических возможностей и новых материалов. Отметим, что это очень широкий спектр деятельности, поэтому технологиче-

ский эксперимент занимает ведущее место в ряду многочисленных экспериментов.

1.4.1. Описание изучаемого процесса (объекта)

В технических науках процесс считается описанным, определенным, если записано уравнение или система уравнений, правильно отражающие наиболее существенные его характеристики, то есть имеется математическая модель процесса. В большинстве экспериментальных исследований построение математической модели является начальной и обязательной стадией изучения процесса. В некоторых задачах - это конечная цель, поэтому методам построения математической модели мы будем уделять особое внимание.

В общем случае математическая модель процесса записывается в виде параметрического уравнения (1) или системы уравнений типа (2).

1.4.2. Нахождение экстремальных условий процесса.

В таких исследованиях требуется найти такие условия протекания процесса, то есть такой набор факторов x_1, x_2, \dots, x_k , при которых выходной параметр y принимает минимальное или максимальное (экстремальное) значение. В качестве y в технологических задачах обычно выступают технико-экономические показатели процесса (стоимость продукции, производительность, расход материала и электроэнергии, процент брака, количество трещин и т. п.). Такая же цель ставится в распространенных в металлургии задачах “состав-свойство”, в которых, варьируя химическим составом сплава, добиваются его наилучших свойств. Основой таких экспериментов является математическая модель (1).

1.4.3. Ранжирование переменных.

Среди входных переменных (факторов) $x_1, x_2, x_3, \dots, x_k$ есть наиболее сильно влияющие на параметр y и с ними, видимо, следует считаться в первую очередь. Это значимые параметры. Мало значимые входные параметры, изменение которых мало отражается на величине y , наоборот, следует исключить из уравнения (1), чтобы упростить исследование. Применительно к технико-экономическим и производственно-технологическим задачам существует теорема, известная в экономических и технологических науках как закон Парето, утверждающая, что значимых параметров всегда немного, точнее, в исследовании могут участвовать только некоторые из них при стабильном значении других значимых параметров. К примеру, когда изучают какой-нибудь параметр качества выпускаемых заводом деталей, то в каждом конкретном случае выделяют не более 3 - 4 наиболее важных причин брака в продукции. Остальные параметры либо слабо влияют на процесс, в нашем случае на процент бракованных деталей, либо на первых порах не заслуживают внимания и в исследовании устанавливаются на каком-то постоянном уровне.

Все факторы в любом процессе можно ранжировать по степени их влияния. Когда проводится эксперимент, то его сложность и стоимость сильно за-

висит от количества входных параметров, поэтому очень важно выделить наиболее значимые факторы и исключить из рассмотрения остальные.

Но, с другой стороны, в теории информатики существует теорема о существенных параметрах, утверждающая, что в случае пропуска какого-нибудь из существенных входных параметров процесса, которые существенно влияют на величину y , результат эксперимента будет неверным и, возможно, прямо противоположным. Очень опасно не учитывать существенные параметры. В сложных процессах часто трудно сразу сказать, какие факторы существенны. Видимо, разумно сначала включить в уравнение (1) как можно большее количество параметров, в том числе сомнительных, в надежде на то, что мало значимые из них в процессе исследования отсеются. Для отсеивания проводится специально спланированный эксперимент и применяется соответствующий математический аппарат для его обработки (дисперсионный анализ).

1.4.4. Имитация реального процесса.

Когда имеется достоверная математическая модель процесса, то появляются широкие возможности по имитации данного процесса. Можно, не решая сложные уравнения, организовать “игру” с ними, то есть проследить поведение выходного параметра при изменении входных. Обычно такие игры осуществляются на ПЭВМ и позволяют выявить различные свойства изучаемого объекта. Например, можно проследить изменение некоторых параметров во времени и спрогнозировать развитие процесса на будущее. При этом устанавливают нежелательные тенденции его развития, можно предотвратить аварию, предусмотреть дополнительные меры для улучшения процесса и т. п. В другом случае, наоборот, необходимо установить узкие пределы изменения входных параметров, которые будут автоматически поддерживаться с целью стабилизации технологического процесса и достижения наивысшего качества продукции. Очень широко применяются имитационные модели для вычисления свойств готового продукта, прогнозирования свойств, корректировки технологического процесса и т. п.

1. 5. Типы эксперимента

Различают два типа экспериментов: пассивный и активный. Пассивные эксперименты в свою очередь проводятся по двум схемам. Первая схема характеризует полную пассивность экспериментатора при назначении входных параметров процесса. Как правило, такие эксперименты проводятся на промышленной установке, на которой процесс отлажен, все параметры процесса установлены на оптимальных или рабочих значениях и изменять их в широких пределах нельзя. Экспериментатор лишь фиксирует наблюдения, измеряя входные и выходные параметры через определенные промежутки времени, никак не воздействуя на ход процесса. Выводы, получаемые в таких экспериментах, относятся только к наблюдаемому узкому диапазону изменения рабочих параметров, и редко дают возможность выйти за этот диапазон.

Такого недостатка лишен пассивный эксперимент, проводимый по второй схеме. Возможно, его не следовало бы называть пассивным, так как действия экспериментатора здесь весьма активны. Лучше называть его лабораторным экспериментом в противовес рассмотренному выше промышленному эксперименту. В таком эксперименте все входные параметры изменяются, наоборот, в самых широких пределах, а поведение объекта исследуется не только при рабочих, но и при любых возможных, даже запредельных значениях параметров. Обычно лабораторный эксперимент такого типа проводится следующим образом. Все факторы, кроме одного x_i , фиксируют на заданных постоянных значениях, а изменяют в широком диапазоне значений только параметр x_i . Получается зависимость

$$y = f(x_i)$$

или серия зависимостей при других наборах значений фиксированных параметров. Затем строятся точно такие же зависимости от других параметров. Полученные графики наглядны, содержательны и достоверны.

Однако, такие эксперименты обладают многими недостатками. Прежде всего, они дороги: опыты проводятся в большом количестве точек, расходуется много материала на изготовление образцов, затрачивается много времени на измерения и пр. Кроме того, графическое представление информации иногда не достаточно, необходима математическая формула, то есть математическая модель, без которой, как указывалось выше, невозможно имитационное моделирование или автоматизированное управление объектом, а также поиск экстремальных условий и многое другое. Получить формулу по графикам принципиально не сложно, хотя возможны практические трудности. Она из них связана с тем, что по графикам трудно количественно оценить эффекты взаимодействия параметров. Поясним это на примере.

Пусть проводится эксперимент только с двумя входными параметрами x_1 и x_2 . По первой серии опытов построены графики зависимостей $y = f(x_1)$ при нескольких постоянных значениях параметра x_2 . Вторая серия опытов дала возможность построить семейство графиков $y = f(x_2)$ при фиксированных значениях параметра x_1 (рис 3).

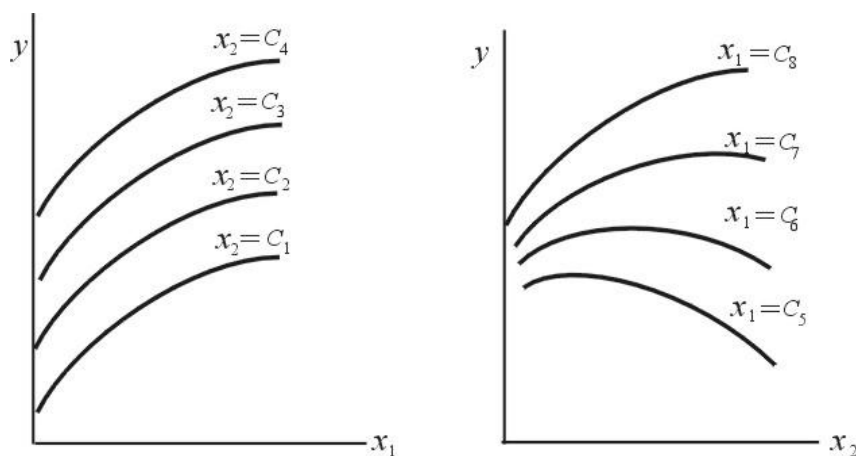


Рис. 3. Характер экспериментальных кривых при отсутствии (а) и наличии (б) взаимодействия между факторами.

Если кривые параллельны, как показано на рис 3 а), то их можно записать в виде единого уравнения вида

$$y = A + f(x_1) + f(x_2),$$

где A – некоторая константа. Функции $f(x_1)$ и $f(x_2)$ каждая зависит только от одной переменной. В этом смысле факторы x_1 и x_2 независимо друг от друга воздействуют на величину y , поэтому называются независимыми. Такие уравнения называются сепарабельными.

На графиках сепарабельные функции представляются в виде параллельных кривых. В нашем примере на рис. 3,а мы имеем семейство таких кривых в функции от параметра x_1 .

Если кривые на графиках не параллельны, то семейство кривых будет описываться уравнением вида

$$y = A + f(x_1) + f(x_2) + f(x_1) \cdot f(x_2),$$

Для семейства прямых линий получим

$$y = A + B \cdot x_1 + C \cdot x_2 + D \cdot x_1 \cdot x_2.$$

Видно, что ход зависимости от одной переменной определяется значением другой переменной. Факторы x_1 и x_2 в своем воздействии на значение y зависят друг от друга, такие переменные взаимодействуют. Член $f(x_1) \cdot f(x_2)$ (или для прямых линий $D \cdot x_1 \cdot x_2$) отражает степень их взаимного влияния на y , поэтому называется взаимодействующим или *взаимодействием*. При пассивных экспериментах взаимодействующие члены определить легко только тогда, когда количество входных факторов мало (например, как у нас, два). Но при большом количестве входных параметров и графиков это превращается в трудную задачу.

Заметим, что приведенная выше теорема Парето по сути своей утверждает, что в технико-экономических задачах многие факторы не взаимодействуют, каждый из них влияет на величину y самостоятельно. По крайней мере, все входные параметры можно разбить на небольшие группы, внутри которых взаимодействие между факторами допустимо, а между группами они отсутствуют. Отсюда понятно, что проводя эксперимент только с входными параметрами одной группы, которых не может быть много, мы не получим искаженный результат.

Активный эксперимент не имеет недостатков пассивных экспериментов. Он заранее планируется так, чтобы с минимальными затратами получить требуемую модель процесса или иную информацию. При этом определяется экономная стратегия проведения эксперимента, в соответствии с которой сначала проводят простую серию опытов. Анализируя эту серию, делают вывод о том, достаточна ли она для достижения цели. Если - нет, то планируется следующая серия опытов.

При современных дорогостоящих опытах проведение активных экспериментов становится практически обязательным. Технике проведения и обработке результатов таких экспериментов и посвящена данная книга. Однако вовсе не следует, что, овладев этой техникой, необходимо совсем отказаться от проведения пассивных экспериментов. Когда опыты не дороги, требуется

наглядная форма их представления, очень эффективен пассивный эксперимент второго вида. Он даже обладает одним преимуществом перед активным экспериментом. Активный эксперимент работает только в однородной области, когда в исследуемой области изучаемая величина y не может изменяться скачкообразно, когда свойства материала не меняются, ничто не ломается, точность измерения остается постоянной.

Если же требуется оценить какие-то быстро изменяющиеся эффекты при переходе из одной области в другую, то он бессилён. Например, при изучении какого-нибудь свойства сплава на границе перехода из одного фазового состояния в другое активный эксперимент неприменим. Наоборот, большое количество опытов пассивного эксперимента здесь может оказаться единственным способом описания всех тонкостей, связанных с фазовым переходом. Известные в металлографии дилатометрические кривые, характеризующие изменение длины образца при его остывании, строятся только методом пассивного эксперимента и обязательно по большому числу опытов, чтобы уловить точки фазовых переходов.

Контрольные вопросы.

1. Какая информация заключена в параметрической модели $y = f(x_1, x_2, x_3, \dots, x_k)$?
2. Какую информацию дает модель в виде «черного ящика»?
3. Является ли чертеж моделью объекта?
4. Каковы преимущества математического моделирования перед физическим и наоборот?
5. Почему английский философ Б. Рассел считал, что развитие математического моделирования таит опасность для человечества?
6. Чем отличаются понятия «измерения», «опыт» и «эксперимент»?
7. Чем отличаются понятия «фактор» и «уровень фактора»?
8. Зачем необходимо ранжировать входные параметры перед проведением эксперимента?
9. Какие преимущества получает исследователь, имея математическую модель процесса или закон?
10. Чем отличается модель от закона?
11. Какая функция называется сепарабельной и какие достоинства имеют процессы, описываемые такими функциями?
12. В чем достоинства активных, то есть планируемых, экспериментов по сравнению с пассивными? Имеют ли преимущества пассивные эксперименты по сравнению с активными?
13. Как по внешнему признаку на графиках узнать, являются ли два фактора взаимодействующими? Какой смысл заложен в понятие о взаимодействии факторов?

2. Основные положения теории подобия и размерностей.

2.1. Характеристика входных и выходного параметров.

Входные параметры можно классифицировать по разным признакам. По одному из них мы их уже разделили: факторы бывают значимыми и незначимыми, или, что то же самое, существенными и несущественными. Правда, количественной характеристики критерия, по которому производится такое разделение, мы пока не знаем. При изучении дисперсионного анализа критерий будет сформулирован.

Входные и выходной параметры – это величины измеряемые, следовательно, содержат в себе некоторую ошибку измерения. Будем считать, что все входные параметры измеряются точно, значения их в эксперименте равны тем, которые мы запланировали. Ошибка в измерении входных параметров равна нулю, эти параметры детерминированы. Выходной параметр y , наоборот, измеряется с ошибкой, то есть является величиной статистической, вероятностной. При одном и том же наборе значений входных параметров в каждом опыте мы можем получать разные значения, отклоняющиеся от истинного (среднего) на некоторую величину $\pm \Delta y$. Считается, что эти отклонения происходят по вине многих мелких случайных причин и подчиняются нормальному закону распределения.

Отклонения в значениях y от среднего (при нормальном законе распределения) возможны не только из-за ошибок измерения, но и по другим случайным причинам. Например, если входные параметры в опыте установлены не совсем точно на запланированных уровнях, то ошибка повлияет на значение y . Отклонения входных параметров от точных значений рассматривается как очередная случайность, повлиявшая на значение y . Таким образом, в отличие от детерминированных входных параметров, выходной параметр y является величиной вероятностной, или случайной по терминологии математической теории вероятности.

Нормальный закон распределения основан на том, что таких мелких случайностей в опытах много, и каждая понемногу воздействует на величину y . Каждая из случайностей может проявляться в одних, и не проявляться в других опытах. Не требуется доказательства нормальности распределения величины y , если она измеряется в каких-то физических единицах измерения и точность ее измерения полностью зависит от условий проведения опыта.

Если каким-то образом показано, что величина y распределена не по нормальному (или логнормальному) закону, то, можно определенно утверждать, что мы столкнулись с одним из двух случаев. Во-первых, мы можем иметь дело со случайной величиной y иной природы, чем нормально распределенная. Чтобы лучше “прочувствовать” природу величины, подчиняющуюся нормальному закону распределения, приведем пример случайной ве-

личины иной природы. Пусть подлежит исследованию время от поломки до поломки нагревательной печи в прокатном цехе. В том, что это время - величина случайная, сомнений нет, но оно не будет зависеть от точности часов, с помощью которых измеряется. Можно также пытаться стабилизировать условия проведения измерения и исключить мелкие случайности, но от этого разброс данных не изменится. В данном случае время - это случайная величина другой природы. Это так называемая марковская случайная величина, случайность которой определяется не условиями измерения, а тем процессом, который она описывает (в нашем случае, условиями эксплуатации печи). Часто величины этой природы не измеряются физическими единицами, а нумеруются в штуках, типах, днях недели и т. п. (тип станка, количество автомобилей, число поломок печи в месяц).

Во-вторых, искажение нормальности закона распределения случайной величины y может произойти по вине некоторых неучтенных в эксперименте существенных факторов, которые не случайным образом воздействуют на значение y , а систематически в каждом опыте проявляют себя. Например, если измеряется температура в печи с помощью термопары, температура холодного спая которой не фиксируется, то при постоянной температуре в печи по мере разогрева холодного спая, показания прибора будут изменяться непрерывно от одного опыта к другому. Закон изменения будет отличным от нормального. Температура холодного спая термопары - это не случайный, а систематический фактор. Такие систематические факторы ведут себя как обычные существенные параметры.

В общем случае отклонение Δy от среднего значения можно представить в следующем виде:

$$\Delta y = f_1(x_1, x_2, \dots, x_p) + f_2(x_{p+1}, x_{p+2}, \dots, x_s),$$

где x_1, x_2, \dots, x_p - систематические, а $x_{p+1}, x_{p+2}, \dots, x_s$ - случайные факторы. В первую группу параметров входят все существенные факторы, влияющие на процесс. Среди них, конечно, присутствуют все те, которые мы сознательно включили в число существенных, но есть и такие, которые, являясь существенными, в эксперимент не включены по невнимательности или незнанию. При наличии последних показания прибора, измеряющего y , при постоянных значениях всех существенных факторах демонстрируют функциональную зависимость от порядкового номера измерения (от $i = 1$ до n) - рис 4.

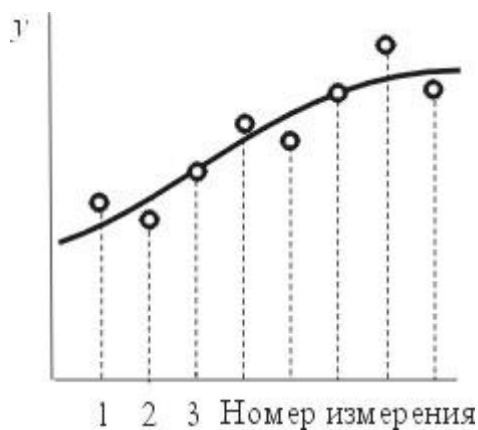


Рис.4. Показания прибора

Эта зависимость называется *трендом*. Если выделить тренд, то есть из каждого измеряемого значения y вычесть значение тренда, то останется только случайная составляющая $f_2(x_{p+1}, x_{p+2}, \dots, x_s)$, которая обязана обеспечить нормальный закон распределения. Случайной нормально распределенной величиной будут не y , а разность $(\acute{o} - \grave{o})$ между экспериментальными значениями y и значениями \grave{o} , лежащими на кривой.

В общем случае, любой другой существенный входной параметр также можно рассматривать как тренд. На анализе величины этого тренда построен, по существу, весь регрессионный и дисперсионный анализ.

Вернемся к терминологии. Ранее мы в двух значениях упомянули термин «случайный». В первом случае случайной величиной назван параметр, имеющий вероятностный характер. Это дань математической теории вероятности. Такая случайная величина может быть нормально распределенной или иметь иную природу и другой закон распределения. Во втором значении случайной величиной назван параметр, который случайным образом появился в том или ином измерении. В другом измерении он может уже не проявить себя. Это те параметры $x_{p+1}, x_{p+2}, \dots, x_s$, которые входят в приведенную выше зависимость Δy . Далее мы будем также употреблять термин «случайный» в обоих значениях, но из контекста всегда будет ясно, в каком значении этот термин применяется.

2.2. Теорема о существенных параметрах.

При постановке задачи опасно пропустить и не включить в исследование какой-нибудь существенный параметр, при этом результат исследования в целом может быть искаженным. С другой стороны, нет смысла включать в эксперимент и несущественные факторы, так как это только удорожает исследование. Теорема о существенных параметрах позволяет разрешить эту проблему.

Если для изучаемого процесса имеется математическая модель, то есть записаны математические уравнения, то в них содержится полный набор существенных параметров и отсутствуют несущественные параметры.

Эту теорему можно рассматривать как частный случай более общей теоремы информатики, которая требует, чтобы в математическом описании процесса обязательно была отражена вся существенная информация о нем (иначе нет смысла в таком описании).

Рассмотрим известную задачу о колебании маятника весом mg (здесь m - масса маятника и g - ускорение силы тяжести), подвешенного на упругой нити длиной l , отклоненного от вертикального положения на угол φ_0 (рис 5).

Уравнение движения маятника, описывающее изменение текущего угла φ во времени, известно:

$$\frac{d^2\varphi}{dt^2} = -\frac{g}{l} \cdot \sin \varphi. \quad (4)$$

Если мы в качестве выходного параметра сделаем угол φ , то есть захотим изучать характеристики движения маятника, то будем иметь параметрическое уравнение:

$$\varphi = f(t, g, l).$$

Если выходным параметром станет время, например, период колебания маятника T , равный времени прохождения двойного угла φ_0 , то получим параметрическое уравнение

$$T = f(\varphi_0, g, l). \quad (5)$$

Как видно, в обоих уравнениях масса маятника m отсутствует. Как для угла φ , так и для периода колебания T этот параметр несущественный. Получен важный результат без решения уравнения, только на основании его параметрического анализа.

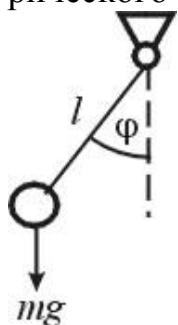


Рис.5. Колебание маятника.

Если мы будем изучать натяжение нити N , то из уравнения натяжения

$$m \cdot l \cdot \left(\frac{d\varphi}{dt} \right)^2 = N - mg \cdot \cos\varphi$$

получим параметрическое уравнение

$$N = f(m, g, l, \varphi).$$

Для натяжения нити масса маятника – параметр существенный.

Таким образом, при наличии математической модели или закона экспериментатор чувствует себя уверенно и в какой-то мере застрахован от пропуска существенных и записи несущественных параметров (конечно, при условии, что модель верна). А как быть в том случае, когда модели нет и трудно ее получить? Видимо, опыт экспериментатора, некоторые предварительные сведения и интуиция могут сыграть решающую роль в формировании набора существенных параметров процесса. Но всегда, пока не получена достоверная модель процесса, у исследователя должно оставаться чувство здорового сомнения в том, что эксперимент поставлен правильно, что все существенные параметры определены полно. Может быть, лучше подстраховаться и включить в исходную параметрическую модель (1) также те параметры, в значимости которых нет уверенности.

Необходимо заметить, что некоторые параметры могут быть существенными в одних условиях и несущественными в других. К примеру, сопротивление воздуха оказывает малое влияние на летные характеристики самолета при малых скоростях полета, но становится весьма существенным фактором при скоростях, близких к звуковым. Та область значений параметров, при которой какой-либо из параметров перестает влиять на процесс, называ-

ется областью *автомодельности*. В теории подобия область автомодельности играет особую роль, поиск ее может быть предметом самостоятельного исследования.

2.3 Размерные и безразмерные величины. Зависимые и независимые размерности.

В общем случае связь между входными и выходными параметрами выражается параметрическим уравнением (1). Все параметры, входящие в него, могут быть:

- первичными размерными,
- вторичными размерными,
- безразмерными.

Первоначально все входные и выходные параметры измеряются приборами и инструментами в виде размерных величин в стандартных единицах измерения - метрах m , секундах s , граммах g , вольтах V и т. д. Стандартные единицы измерения объединены в международную систему СИ (есть и другие системы). В эксперименте обязательно использование только стандартных единиц измерения (только метр, но не аршин или дюйм, хотя в бытовых измерениях могут быть использованы и эти единицы).

Вторичные размерные величины образуются из первичных размерных величин и представляют собой комплексы, имеющий некоторый физический смысл. Соответственно, размерности этих величин образуются как комбинация стандартных размерностей. Например, площадь с размерностью m^2 , сила - $g \cdot m/c^2$, скорость - m/c и т. д. В теории подобия доказывается теорема, - мы ее, как и все остальные теоремы примем без доказательств, - утверждающая, что любая вторичная размерность образуется из первичных путем их перемножения и возведения в степень. Если используемые нами первичные размерности в общем виде представить в виде ряда латинских букв M, L, T, P , то размерность $[A]$ любой вторичной величины A будет выражаться формулой:

$$[A] = M^a \cdot L^b \cdot T^c \cdot P^d, \quad (6)$$

где a, b, c, d - степени при первичных размерностях.

Например, если обозначено: m - M , s - L , g - T (P отсутствуют), то размерность площади Q получится при $a = 2, b = 0, c = 0$, силы F - при $a = 1, b = -2, c = 1$, скорости V - при $a = 1, b = -1, c = 0$. Можно записать:

$$\begin{aligned} [Q] &= m^2 = M^2 \cdot L^0 \cdot T^0, \\ [F] &= g \cdot m/c^2 = M^1 \cdot L^{-2} \cdot T^1, \\ [V] &= m/c = M^1 \cdot L^{-1} \cdot T^0. \end{aligned}$$

По этой теореме, как следствие, нельзя складывать или вычитать различные размерности; их можно только умножать, делить или возводить в степень

Каждая конкретная задача в первоначальном виде, то есть на стадии формирования параметрического уравнения (1), содержит определенный набор входных и выходной параметры с первичными и вторичными размерностями для данной задачи. Из этого ряда всегда можно выделить зависимые и

независимые размерности. *Независимыми* будут такие размерности, которые нельзя образовать друг из друга. Так все первичные размерности, а также часть вторичных могут быть отнесены к независимым. *Зависимые* размерности образуются из независимых по формуле (17) и всегда вторичны.

Часто параметры, имеющие независимую размерность, также называют независимыми. Однако, в задаче может быть несколько параметров с одинаковыми независимыми размерностями (например, длина, ширина и высота с одной размерностью м). Выходной параметр всегда носит зависимую размерность.

Для примера, рассмотрим некоторую задачу, имеющую 6 выходных параметров x_1, x_2, \dots, x_6 с размерностями соответственно: м, м, с, г·м/с², г/с, м/с. Независимыми размерностями признаем: м, с, г/с. Остальные размерности можно образовать из этих размерностей, поэтому эти остальные становятся зависимыми размерностями.

Имеется теорема, согласно которой из приведенного ряда в качестве независимых можно выбирать любые размерности, при этом количество независимых размерностей не изменится (в нашем случае 3).

В параметрическое уравнение (1) помимо параметров, имеющих первичные и вторичные размерности, могут входить безразмерные параметры. Их следует рассматривать как частный случай вторичных размерных величин, когда входящие в уравнение (17) размерности взаимно сокращаются (размерность числителя равна размерности знаменателя). Некоторые из таких величин по определению безразмерны, хотя в основе своей они всегда представляют собой некоторые комплексы из первичных размерных величин. Например, коэффициент трения как безразмерная величина образовался как отношение силы трения к нормальной силе. Радиан - это отношение длины дуги окружности к радиусу, поэтому является безразмерным. Кстати, угол, измеряемый в градусах (единицах измерения угла), также может рассматриваться как величина безразмерная, так как градус представляет собой отношение длины дуги к длине окружности. С этих позиций и параметры, измеряемые метрами, можно трактовать как безразмерные величины, поскольку метр определен как некая доля длины меридиана. Другие стандартные единицы измерения также являются стандартизированными порциями измеряемых величин. Видимо, четко очерченных границ между размерными и безразмерными величинами нет.

2.4. Структура связей между параметрами.

Пусть в исходном параметрическом уравнении

$$y = f(x_1, x_2, x_3, \dots, x_k)$$

все параметры размерны и первые n параметров имеют независимые размерности ($n < k$). Выясним структуру этого уравнения в предположении, что оно описывает некий закон, который не зависит от масштаба единиц измерения.

Независимые размерности обозначим следующим образом:

$$[x_1] = A_1, [x_2] = A_2, \dots, [x_n] = A_n.$$

Размерности остальных (зависимых) параметров определяются по формуле (17):

$$[y] = A_1^{m1} A_2^{m2} \dots A_n^{mn},$$

$$[x_{n+1}] = A_1^{p1} \cdot A_2^{p2} \dots A_n^{pn},$$

.....

$$[x_k] = A_1^{q1} \cdot A_2^{q2} \cdot A_n^{qn}.$$

Изменим теперь единицы измерения независимых величин соответственно в $a_1, a_2 \dots a_n$ раз. Для физического закона от этого закономерность, заложенная в исходное параметрическое уравнение (1), не изменится. В новых единицах все параметры запишутся следующим образом:

независимые:

$$x_1'' = a_1 \cdot x_1,$$

$$x_2' = a_2 \cdot x_2,$$

.....

$$x_n' = a_n \cdot x_n,$$

зависимые:

$$y' = a_1^{m1} \cdot a_2^{m2} \dots a_n^{mn} \cdot y,$$

$$x'_{n+1} = a_1^{p1} \cdot a_2^{p2} \dots a_n^{pn} \cdot x_{n+1},$$

.....

$$x'_k = a_1^{q1} \cdot a_2^{q2} \dots a_n^{qn} \cdot x_k.$$

Числа a_1, a_2, \dots, a_n можно выбрать произвольно, например, таким образом:

$$a_1 = 1/x_1, \quad a_2 = 1/x_2, \quad \dots, \quad a_n = 1/x_n.$$

Тогда в исходном параметрическом уравнении первые n параметров будут иметь значения равные 1 и превратятся в своеобразные единицы измерения для зависимых параметров. Последние же примут вид:

$$y' = \Pi = \frac{y}{x_1^{m1} \cdot x_2^{m2} \dots x_n^{mn}},$$

$$x'_{n+1} = \Pi_1 = \frac{x_{n+1}}{x_1^{p1} \cdot x_2^{p2} \dots x_n^{pn}}, \tag{7}$$

.....

$$x'_k = \Pi_{k-n} = \frac{x_k}{x_1^{q1} \cdot x_2^{q2} \dots x_n^{qn}}.$$

Видно, что числители и знаменатели величин $\Pi, \Pi_1, \Pi_2, \dots, \Pi_{k-n}$ имеют одинаковые размерности, следовательно параметры Π_k становятся безразмерными. Введя их в исходное параметрическое уравнение, получим:

$$\begin{aligned} \Pi &= f(1, 1, \dots, 1, \Pi_1, \Pi_2, \dots, \Pi_{k-n}) \quad \text{или} \\ \Pi &= f(\Pi_1, \Pi_2, \dots, \Pi_{k-n}) . \end{aligned} \quad (8)$$

2. 5. П-теорема подобия.

Уравнение (8) является математической записью П-теоремы подобия. Сформулируем ее следующим образом. Если известно параметрическое уравнение (1) в размерном виде, то его всегда можно привести к безразмерному виду (8), причем количество безразмерных параметров будет меньше, чем количество размерных параметров на количество независимых размерностей. В уравнении (1) количество входных параметров равно k и количество независимых размерностей пусть равно n , поэтому в уравнении (8) количество размерных параметров будет равно $k - n$.

П-теорема широко – используется при проведении экспериментальных исследований. Прежде всего, она дисциплинирует исследователя, утверждая, что переход от размерных параметров к безразмерным осуществим всегда. Если в каком-то случае этого сделать не удастся, то это свидетельствует только о том, что какие-то существенные параметры с недостающей размерностью в исследовании пропущены. Более того, П-теорема признает неграмотной запись такого уравнения (8), в котором наряду с безразмерными встречаются и некоторые размерные, хотя с точки зрения самого эксперимента здесь нет никакой ошибки.

Кроме того, П-теорема обязывает, чтобы перед началом эксперимента всегда все параметры исходного параметрического уравнения преобразовывались в безразмерные. При этом сам эксперимент упрощается, так как не все равно, какое количество входных параметров участвует в эксперименте – k или $k - n$. Чем меньше входных параметров, тем эксперимент дешевле.

Наконец, безразмерные параметры в более общей форме характеризуют объект, чем размерные. Например, если исследуется некоторый параллелепипед в размерных параметрах – высота H , ширина B и длина L – , то и выводы о нем будут справедливыми только в пределах этого параллелепипеда, но если ввести безразмерные параметры B/H и L/H , то с такими соотношениями размеров становятся одинаковыми все подобные параллелепипеды, и выводы, полученные в эксперименте, распространяются уже на целый класс подобных фигур.

Рассмотрим несколько примеров.

1. Вернемся вновь к маятнику (рис. 5) и с позиций П-теоремы рассмотрим задачу о периоде T его колебания. Ранее было записано параметрическое уравнение (5):

$$T = f(\varphi_0, g, l).$$

Параметр φ_0 безразмерный, и функция от него также безразмерна, поэтому можно предложить несколько форм записи параметрического уравнения, например:

$$T = f_1(\varphi_0) \cdot f_2(g, l) \quad \text{или} \quad \varphi_0 = f_3(g, l, T) \quad \text{и т. п.}$$

Остановимся на последней как более удобной.

Размерности величин в правой части таковы:

$$[T] = c, \quad [l] = m, \quad [g] = m/c^2.$$

Две первые размерности независимые, а третья зависимая, ($k = 3, n = 2$) и возможна только единственная комбинация всех трех параметров, которая будет безразмерной:

$$\varphi_0 = f_4(T^2 \cdot g / l).$$

Задача подготовлена для экспериментального определения периода колебания маятника: достаточно теперь экспериментально построить график, по одной оси которого отложить несколько произвольных значений угла φ_0 , а по другой – значение комплекса $T^2 \cdot g / l$, в котором время прохождения маятником полного периода колебаний T (двойного угла φ_0) измеряется с помощью секундомера (рис.6). Далее останется лишь вычислить значение периода колебаний из выражения $T = f_1(\varphi_0) \cdot \sqrt{g/l}$, и задача будет решена.

Обязательно следует подчеркнуть, что решение здесь достигнуто экспериментально. Для данной задачи известно и теоретическое решение, полученное путем математического решения уравнения колебания маятника (16):

$$T = 2\pi \cdot \sqrt{g/l}$$

Экспериментальное значение функции $f_1(\varphi_0)$ будет также близким к постоянному значению 6,28, как это видно рис. 6.

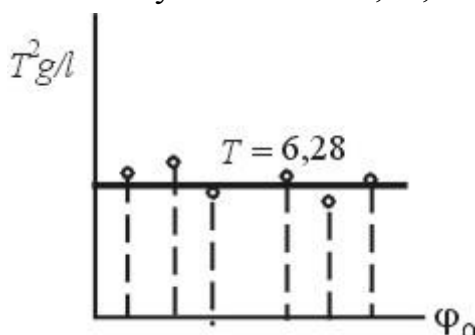


Рис. 6. Экспериментальное решение задачи о колебании маятника.

Рассматривая задачу о колебании маятника, можно подумать, что экспериментальный путь решения менее эффективен, чем математический. Для данной задачи это, видимо, так. Но для более сложных задач он может быть не только эффективным, но и единственным.

2. Проанализируем следующую задачу.

Определить силу F , необходимую для перемещения со скоростью V шара диаметром d в жидкости, имеющей плотность ρ и коэффициент вязкости μ . Можно изменить содержание задачи, задав вместо жидкости воздушную среду, а вместо шара летательный аппарат. Можно также видоизменить ее применительно к кораблю, движущемуся по поверхности моря. В такой постановке задача приобретает практический интерес, поэтому решение, которое мы собираемся осуществить экспериментальным путем на основе П- теоремы подобия, применимо к широкому классу прикладных задач.

Предположим, что поставленная задача содержит полный набор существенных входных параметров, тогда параметрическое уравнение будет иметь вид:

$$F = f(d, V, \rho, \mu). \quad (9)$$

Необходимо заметить, что, не имея уравнения движения шара в жидкости, мы не можем быть до конца уверенными, что все существенные параметры учтены. В данном случае только опыт, интуиция и знания дают уверенность в том, что параметрическое уравнение записано правильно.

Проанализируем размерности:

$$[F] = \text{кгм/с}, [d] = \text{м}, [V] = \text{м/с}, [\rho] = \text{кг/м}^3, [\mu] = \text{кг/мс}.$$

Все размерности разные. В правой части имеем число размерных параметров $k = 4$, независимые размерности – м , м/с и кг/м^3 , число их $n = 3$, поэтому число безразмерных комплексов будет $4 - 3 = 1$. Независимым размерностям соответствуют независимые параметры d , V и ρ .

Чтобы практически воспользоваться формулой (9) для формирования безразмерного параметра, поступим следующим образом. Запишем слева зависимый параметр μ , а справа его размерность кг/мс :

μ кг/мс
 Справа имеющуюся размерность разделим на кг/м^3 , чтобы сократить кг . Соответственно и слева имеющийся параметр разделим на ρ , получим:

$$\mu/\rho \qquad \text{кг/мс: кг/м}^3 = \text{м}^2/\text{с}.$$

Далее следует поделить правую часть на м/с и левую - на V , в правой части останется размерность м , которую уничтожаем делением на размерность параметра d . Окончательно в левой части сформируется безразмерная комбинация $\mu/\rho V d$, которая известна в теории подобия как критерий Рейнольдса Re (точнее, $1/Re$). Далее то же самое студентам полезно проделать с размерностью силы F и получить безразмерную комбинацию $F/\rho V^2 d^2$. Окончательное безразмерное параметрическое уравнение будет иметь вид:

$$F/\rho V^2 d^2 = f(\mu/\rho V d).$$

Так как в левой и правой частях уравнения содержится только по одной переменной (безразмерной), то легко экспериментально построить график этой зависимости, которая будет справедлива для любых шаров, любых скоростей и любых характеристик жидкости (и не только для шаров и не только для жидкостей). Задача о движении шара в жидкости свелась к экспериментальному построению графика.

В учебных целях также полезно проследить, что получится, если в качестве независимых размерностей выбрать м , м/с и кг/м , которым соответствуют параметры d , V и μ . Применяя приведенную выше процедуру, в правой части исходного уравнения (9), получим тот же критерий Рейнольдса $Re = \rho V d / \mu$. Слева же будем иметь: $F / \mu V d$. Можно экспериментально построить графическую зависимость

$$F / \mu V d = f(\rho V d / \mu),$$

и никакого противоречия с предыдущим результатом здесь нет, так как

$$(F / \mu V d) = (F / \rho V^2 d^2) \cdot Re.$$

Кстати, вторая форма построения этой зависимости проще и применяется чаще, чем первая.

2. В 1915 году Релей нашел блестящее решение задачи Буссинеска, используя только возможности П-теоремы. Приблизив к металлургической проблематике, сформулируем эту задачу следующим образом: найти количество тепла H , отнимаемое в единицу времени водоохлаждаемой трубой, находящейся внутри разогретой печи, если известны: диаметр трубы d , разность температур печного пространства и жидкости внутри трубы θ (будем считать, что в стационарном режиме работы она неизменна во времени), скорость движения жидкости в трубе V , теплоемкость жидкости C , плотность ρ и коэффициент теплопроводности μ .

Можно записать параметрическое уравнение

$$H = f(d, \theta, V, C, \lambda, \rho). \quad (10)$$

Нетрудно догадаться, что рассматриваемый процесс описывается несколькими сложными дифференциальными уравнениями, решить которые сложно даже на быстродействующих ЭВМ при современном развитии математических численных методов. Так как таких уравнений нет, то нельзя воспользоваться теоремой о существенных параметрах, и нет уверенности в том, что в правой части параметрического уравнения (10) перечислены все существенные параметры. Кстати, после Релея различные ученые долго не соглашались с предложенным набором существенных параметров. Одни считали, что их мало, и вносили в задачу свои добавки, другие, наоборот, полагали, что среди записанных есть один лишний – параметр θ , так как он, как и параметр H , может быть выражен через энергию.

В постановке Релея размерности таковы:

$$[H] = \text{дж}/\text{с}, [d] = \text{м}, [\theta] = \text{К}, [V] = \text{м}/\text{с}, [C] = \text{дж}/(\text{г}\cdot\text{К}), \\ [\lambda] = \text{дж}/(\text{м}\cdot\text{с}\cdot\text{К}), [\rho] = \text{г}/\text{м}^3.$$

В правой части параметрического уравнения пять из шести размерностей должны быть приняты как независимые и только одна будет зависимой (например, соответствующая параметру C). В правой и левой частях уравнения останется по одному безразмерному комплексу:

$$\frac{H}{\lambda \cdot \theta \cdot d} = f\left(\frac{d \cdot V \cdot C \cdot \rho}{\lambda}\right).$$

Такую зависимость можно построить в виде графика экспериментально, этот график и будет решением поставленной задачи.

Проанализировав зависимость, можно сразу сделать несколько полезных выводов. Например, видно, что количество тепла H прямо пропорционально разности температур θ . (По параметрам d и λ такого вывода сделать нельзя, так как они входят в левую и правую часть уравнения). Кроме того, параметры V, C и ρ отсутствуют в левой части, следовательно, каждый из них в отдельности не влияет на процесс, а определяющим будет только их произведение, которому можно придать некоторый физический смысл.

2.6. Критерии процесса.

Все приведенные выше примеры имеют одно общее свойство: все размерности входных параметров разные. При приведении к безразмерному виду получаем уравнение типа

$$\Pi = f(\Pi_1),$$

в котором один безразмерный выходной параметр зависит только от одного входного безразмерного параметра. В этом случае оба безразмерных параметра обязаны быть критериями.

Критерием процесса является такой безразмерный параметр, при котором связь между параметрами не изменяется при любом способе его формирования. Можно, например увеличить в два раза один из размерных параметров, входящих в числитель критерия Π_1 , и одновременно увеличить также в два раза другой размерный параметр, находящийся в его знаменатель. Значение критерия Π_1 от этого не изменится и также неизменным должно остаться значение выходного параметра Π . На маленькой модели и натурном объекте зависимость выходного критерия Π от входного Π_1 будет той же самой. Если это не так, то параметр Π_1 нельзя считать критерием, он будет простым безразмерным параметром. Критерий процесса интереснее простого безразмерного параметра еще и тем, что всегда имеет некоторый физический смысл и “ответственен” за физическую сущность процесса, являясь обобщенной характеристикой этого процесса (или некоторой части процесса).

Когда безразмерный параметр формируется из размерных единственно возможным их сочетанием, то такой безразмерный параметр автоматически становится критерием (как в приведенных выше примерах). Но, если существует несколько возможностей формирования безразмерных комбинаций из размерных, то поиск среди них критерия становится трудной и самостоятельной задачей. П-теорема подобия здесь не в состоянии помочь.

Рассмотрим такую ситуацию на примере. На рис. 7 приведен процесс прокатки прямоугольной полосы толщиной h_0 до толщины h_1 в цилиндрических валках диаметром D . Пусть изучается параметр P – давление металла на валки, то есть усилие прокатки, приходящееся на единицу площадки контакта металла с валками. В качестве входных нас интересуют приведенные на рисунке геометрические параметры:

$$p = f(h_0, h_1, D).$$

Параметр l – длина дуги контакта – сюда не входит, так как она вычисляется по другим параметрам:

$$l = \sqrt{0,5(h_0 - h_1)D}.$$

Размерность выходного параметра $[p] = н/м^2$, а всех трех входных параметров $[h_0] = [h_1] = [D] = м$. Попытаемся проанализировать задачу методами теории подобия. Видно, что в левой части параметрического уравнения невозможно сформировать безразмерную комбинацию. Однако П-теорема утверждает, что это можно сделать всегда. В таких случаях причина состоит в том, что не все существенные параметры перечислены в задаче. В нашем

уравнении отсутствует некоторый параметр с размерностью n/m^2 , им может быть предел текучести прокатываемого металла σ , и параметрическое уравнение должно иметь вид:

$$p = f(h_0, h_1, D, \sigma).$$

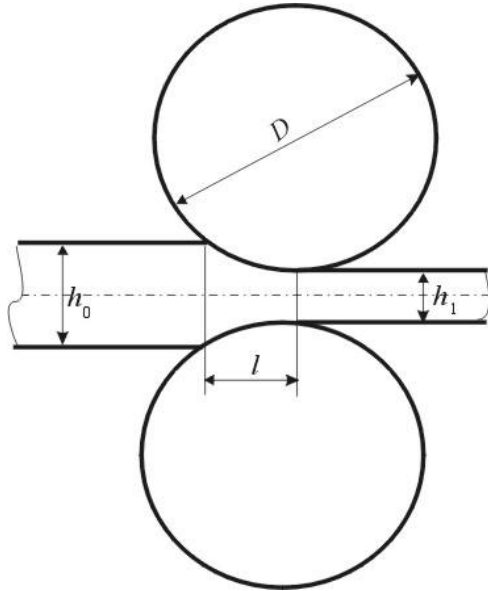


Рис. 7. Схема прокатки металла.

В правой части любой из трех параметров можно принять за независимый и, поделив на него, образовать два безразмерных комплекса. С точки зрения П-теоремы равнозначны любые безразмерные уравнения:

$$P/\sigma = f(h_0/h_1, D/h_1),$$

$$P/\sigma = f(h_1/h_0, D/h_0),$$

$$P/\sigma = f(h_0/D, h_1/D).$$

Более того, любая линейная комбинация из приведенных выше трех размерных параметров с размерностью m , имеющая некоторый физический смысл, может заменить один из них. Например, обжатие $\Delta h = h_0 - h_1$, средняя высота полосы $h_{cp} = 0,5(h_0 + h_1)$ или длина очага деформации l и т. д. Существует следствие из П-теоремы, утверждающее, что от применения любых из размерных величин в качестве независимых количество безразмерных параметров в конечном безразмерном параметрическом уравнении останется неизменным. В нашем случае их всегда будет два, и можно записать несколько безразмерных параметрических уравнений с участием вновь введенных параметров, например:

$$p/\sigma = f(\Delta h/h_0, D/h_{cp}),$$

$$p/\sigma = f(\Delta h/h_{cp}, l/h_{cp}),$$

$$p/\sigma = f(\Delta h/l, l/h_{cp}) \quad (11)$$

и так далее.

По П-теореме все формы записи равнозначны, но не все приведенные безразмерные комбинации входных параметров являются критериями. Кри-

терий имеет большую практическую ценность, чем просто безразмерный параметр, так как несет ответственность за физическую сущность процесса и зависимость p/σ от него будет неизменной при изменении и числителя, и знаменателя. Естественно, для описываемого процесса прокатки начался поиск критерия, растянувшийся почти на десяток лет. Чижиков Ю. М. провел следующий эксперимент по исследованию безразмерного параметра $\Delta h / h_0$. В первой серии опытов этот параметр изменялся только за счет знаменателя: с постоянным обжатием Δh прокатывались образцы различной исходной толщины h_0 . Получилась зависимость, показанная на рис. 8а, кривой 1. Во второй серии опытов начальная толщина полосы h_0 была постоянной, но менялось обжатие Δh за счет изменения конечной толщины полосы h_1 . Изменялся только числитель, и получилась кривая 2.

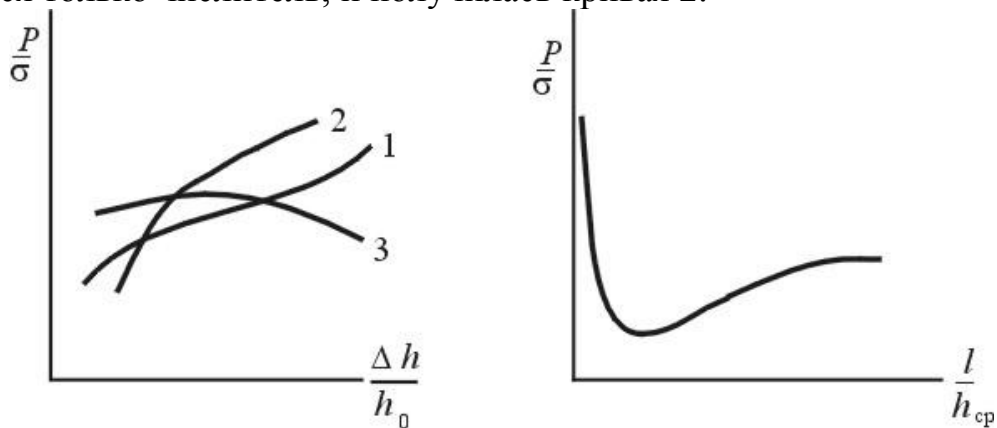


Рис. 8. Зависимость давления металла на валки при прокатке от факторов $\Delta h/h_0$ (а) и l/h_{cp} (б).

Наконец, если оставить постоянным значение h_1 , а изменять одновременно обжатие Δh и начальную толщину полосы h_0 (и числитель и знаменатель), то получается кривая 3. Вывод: параметр $\Delta h / h_0$ не является критерием процесса прокатки.

Критерием оказался параметр l / h_{cp} , экспериментально найденный позднее Тарновским И. Т. На рис. 8 б приведена зависимость P/σ от критерия l / h_{cp} , которая сохраняется неизменной при любом задании числителя и знаменателя критерия, при прокатке на любом прокатном стане.

Таким образом, из всех приведенных выше форм записи безразмерных параметрических уравнений наиболее предпочтительной оказалась последняя из серии (11). В ней помимо упомянутого критерия присутствует также критерий $\Delta h / l$. Этот критерий, называемый в теории прокатки углом захвата (точнее, половиной угла захвата), изменяется в узких пределах и слабо влияет на процесс. Его даже можно исключить из параметрического уравнения. Как указывалось выше, та область процесса, в которой один из параметров становится слабым или перестает влиять на изучаемые параметры, в теории подобия называется областью *автомодельности*. Нахождение такой области весьма желательно для экспериментального исследования, так как в

этой области количество входных параметров уменьшается и эксперимент упрощается. Давление металла на валки при прокатке автомодельно по параметру $\Delta h / l$ и практически определяется только одним критерием $l / h_{\text{ср}}$.

Практически всегда, когда находится критерий, некоторые другие параметры теряют свою значимость, поскольку критерий по природе своей вбирает в себя всю информацию о процессе.

Контрольные вопросы по главе 2.

1. Что такое тренд?
2. Как выделить случайную составляющую измерений, подчиняющуюся нормальному закону распределения?
3. Как определяется область автомодельности?
4. Какой практический результат можно извлечь из теоремы о существенных параметрах?
5. Как сформировать безразмерные параметры при известном наборе входных параметров?
6. Чем отличается критерий от простого безразмерного параметра?
7. Как выделяют критерии из параметрического уравнения?
8. Что дает П-теорема, какие недостатки она имеет?
9. Почему задачу Буссинеска удалось решить Реллею экспериментально?
10. В чем преимущества и недостатки решения задачи о маятнике теоретическим и экспериментальным путем?

3. Основы физического моделирования

3.1. Теоремы “классической” теории подобия.

Физическое моделирование процессов базируется на трех основных теоремах подобия, которые в силу исторического развития относятся к “классической” теории подобия, в отличие от последующих теорем, которые в какой-то мере лишь развивают и обобщают “классические” теоремы. В этой связи полезно формулировать их в начальных редакциях, постепенно усложняя смысловую нагрузку и демонстрируя направления развития теории подобия в целом.

Первая теорема или теорема Ньютона. “Явления, подобные в физическом и геометрическом смысле, имеют одинаковые критерии, равные по значению.” Далее в трудах В. Л. Кирпичева эта теорема получила более конкретную и приближенную к практическому использованию формулировку. “В двух подобных явлениях, то есть при равенстве всех геометрических и физических критериев, физические явления в них протекают одинаково, отличаясь только переходным масштабом”. “Две механические системы – модель и натура – геометрически подобные по размерам, будут получать геометрически подобные движения, то есть будут копировать движения, только изменив масштаб”. Количественные характеристики природы легко пересчитываются из

таких же характеристик, добытых экспериментально на модели, путем простого домножения на масштабный коэффициент.

Следует заметить, что первая теорема подобия является гениальной догадкой Ньютона о том, что в одинаковых (подобных) процессах содержится некоторый общий механизм – критерий, ответственный за ход процесса в целом. Более того, можно смело утверждать, что такой механизм (критерий) имеется всегда и во всех процессах, его надо только найти. Другого практического смысла эта теорема не имеет, хотя и это утверждение дает много.

Вторая теорема. “Для того, чтобы данные, полученные из опытов или уравнений, можно было распространить на подобные явления, эти данные необходимо обрабатывать в виде зависимости между критериями подобия, или в виде критериального уравнения

$$f(K_1, K_2, \dots, K_n) = 0, \quad (12)$$

где K_1, K_2, \dots, K_n - критерии процесса.

Эта теорема имеет два важнейших следствия. Во-первых, любой процесс можно записать в виде критериального уравнения, в котором в качестве параметров будут представлены только критерии процесса. Никаких других параметров – размерных или безразмерных некритериев – в нем не должно быть. Такое утверждение очень дисциплинирует исследователя, требуя от него выполнения преобразования параметров в критерии до конца, и только после этого, строго говоря, эксперимент следует считать подготовленным. Теоретически переход к критериальному представлению задачи выполним всегда, однако практически это удается сделать далеко не в каждой задаче.

Второе следствие утверждает, что если для исследуемого процесса записаны математические уравнения (существует математическая модель), то существуют равнозначные им по информативности критерии процесса. Можно расширить сформулированную ранее теорему о существенных параметрах утверждением, что в критериях содержится только существенная информация, в их формировании участвуют только существенные параметры и ни один из них не пропущен.

Таким образом, математическая модель позволяет анализировать процесс путем решения математических уравнений методами математического моделирования, а преобразование модели в критериальную форму (12) позволяет без потери информации осуществлять анализ того же процесса экспериментальным путем.

Первая и вторая теоремы подобия рассматривают свойства заведомо подобных процессов. Третья теорема подобия (теорема Кирпичева- Гухмана) определяет признаки, по которым можно установить, что процессы подобны. “Необходимым и достаточным условием подобия является пропорциональность сходственных величин, входящих в условие однозначности, и равенство критериев сопоставляемых явлений”. Под условием однозначности подразумеваются начальные и граничные условия задачи, то есть условия в начальный момент времени и на границах рассматриваемой области действия процесса. Следовательно, чтобы в модели и натуре процессы протекали по-

добным образом, необходимо обеспечить равенство не только критериев, но и начальных (“стартовых”) и граничных условий.

3.2. Физическое моделирование.

Теоремы подобия, прежде всего третья теорема, дают прямое указание, как следует осуществлять физическое моделирование процессов, поэтому эти теоремы стали мощным инструментом при исследовании объектов во всех прикладных науках – гидротехнике, теплотехнике, аэродинамике, механике и др. Многие практические задачи решены путем моделирования, особенно при исследовании громоздких или длительно протекающих процессов. Даже если процессы неплохо описаны математическими уравнениями и возможно полноценное математическое моделирование, полезно промоделировать их на лабораторных установках, прежде чем строить промышленные объекты, так как математические уравнения не учитывают всех тонкостей процесса и особенностей конкретного оборудования. Многие дополнительные сведения об объекте можно получить только при физическом моделировании, поэтому предварительное моделирование осуществляют при строительстве прокатных станов, самолетов, плотин ГЭС и т. д.

Однако, несмотря на кажущуюся простоту требований к подобию процессов, равенство всех критериев на модели и натуре обеспечить не всегда удается. Есть мнение, что такого нельзя достигнуть практически в большинстве исследований, и физическое моделирование осуществляется всегда приближенно. К примеру, выше отмечалось, что летные характеристики самолета можно промоделировать в аэродинамической трубе, если для модели и натуре обеспечить равенство критерия Рейнольдса $Re = Vd\rho/\mu$. Для моделирования на уменьшенных моделях приходится либо увеличивать скорость потока воздуха в трубе до предельно возможных и даже невозможных значений, либо существенно увеличивать плотность воздуха, что также имеет определенные ограничения. Приходится строить сложные закрытые трубы, в которых с предельной скоростью движется сжатый, почти плотный воздух. Нередко предпочитают использовать в качестве модели настоящий самолет.

Другой характер затруднений можно продемонстрировать при моделировании процесса горячей прокатки полосы, представленного на рис. 7. Очевидно, на модели и натуре, помимо равенства геометрического критерия l/h_{cp} , необходимо обеспечить одинаковое изменение предела текучести металла σ в процессе исследования, то есть заставить одинаково во времени остывать лабораторный и промышленный слитки металла. Это невозможно, и всегда маленький слиток остывает быстрее большого. Кроме того, при отливке крупного и маленького слитков условия их затвердевания настолько различны, что невозможно достигнуть равенства пределов текучести даже при одинаковой температуре. Как же выйти из положения? Конечно, единого рецепта нет, но можно предложить несколько решений. Например, если крупный слиток остывает очень медленно и практически не изменяет температуру во времени, то на лабораторном стане можно прокатывать в холодном

состоянии свинцовый слиток, который при комнатной температуре имеет примерно такой же предел текучести, что и стальной слиток при высокой температуре. Другой вариант: построить график изменения предела текучести металла от температуры, затем вычислить или измерить температуры в натурном и модельном слитках в момент их сравнения и по графику определить их фактические пределы текучести σ_m и σ_n , а затем результат измерения усилия P_m , полученный на модели, пересчитать в натурное значение P_n по формуле $P_n = P_m \cdot \sigma_n / \sigma_m$.

Интересная ситуация возникает при моделировании балки, представленной на рис. 9. прямоугольного сечения. Балка может быть, например, моделью моста. Тяжелая балка на двух опорах высотой H , шириной B и длиной L , изготовленная из однородного материала с модулем упругости E ($н/м^2$) и плотностью ρ ($н/м^3$), нагружена внешней силой P ($н$), а также силой собственной тяжести Q .

Последняя определяется удельным весом материала $\gamma = g \cdot \rho$ (размерность $н/м^3$, так как $н = кгм/с^2$) и равна

$$Q = g \cdot \rho \cdot H \cdot B \cdot L.$$

Пусть в задаче требуется определить некоторую прочностную характеристику балки y , и пусть эта величина безразмерна. Если перечень существенных параметров определен, то можно записать параметрическое уравнение:

$$y = f(H, B, L, P, Q, E).$$

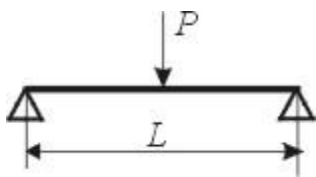


Рис.9. Моделирование балки на двух опорах

В качестве независимых размерностей выберем стандартные $н$ и $м$, тогда согласно П-теореме число безразмерных параметров будет равным $6 - 2 = 4$, и в безразмерном виде параметрическое уравнение будет иметь вид:

$$y = f\left(\frac{B}{L}, \frac{H}{L}, \frac{E}{\rho g L}, \frac{P}{EL^2}\right)$$

Заметим, что геометрические параметры H , B и L с размерностью $м$ никакими другими линейными параметрами заменены быть не могут. Это резко отличает нашу задачу от задачи по прокатке, в которой также участвовало несколько параметров с одинаковыми размерностями, но была возможность записать несколько равнозначных комбинаций. Поэтому в задаче о балке геометрические параметры в представленных отношениях являются критериями. Итак, два из приведенных в уравнении критериев являются геометрическими, третий характеризует массовые силы тяжести и последний отражает обобщенную характеристику внешней силы. Допустим, длина модельной балки в 10 раз меньше натурной. Тогда для равенства первых двух критериев высота H и ширина B модели также должны быть уменьшены в 10 раз. Внешняя сила P на модели должна быть в 100 раз меньше, что при равенстве

модулей упругости обеспечит равенство четвертого критерия. Все эти условия легко выполнить. Но равенство третьего критерия выполнить трудно. Если выполнить модель и натуральный мост из одинакового материала (при равенстве E и ρ), то на модели необходимо увеличить в 10 раз ускорение силы тяжести g , то есть поместить модельную балку в центрифугу.

Уже приведенных примеров достаточно, чтобы убедиться в том, что условия подобия на практике не всегда выполнимы и требуют порой ухищрений и инженерной выдумки. В подавляющем большинстве модельных экспериментов подобие выполняется лишь приближенно.

3.3. Фракционный анализ.

“Фракционным анализом называется метод нахождения информации о решении задачи, когда нет времени, способов или необходимости использования более мощных численных математических методов анализа, ставящих своей целью получение полной информации о решении” [6]. Метод фракционного анализа обеспечивает приведение исходных математических уравнений к критериальному виду и извлечение информации путем выделения и анализа критериев, а также физического моделирования на их основе. Это сравнительно молодое направление исследования уравнений математической физики и процессов, описываемых ими. Оно интенсивно развивается с привлечением математического аппарата теории групп и других разделов математики.

Методы фракционного анализа связаны с “классической” теорией подобия, так как в основе своей преследуют те же цели - приведение задачи к критериальному виду и подготовка ее к физическому моделированию. Однако, если “классический” метод начинает анализ с формирования параметрического уравнения (1), то фракционный анализ требует вначале четкой формулировки краевой задачи. В математической физике краевой называется такая задача, для которой сформулированы дифференциальные (или иные) уравнения, полно описывающие процесс, а также условия однозначности их решения, то есть начальные и граничные условия. Такие задачи решаются в аэро- и гидродинамике, гидравлике, теории упругости и пластичности, механике, теплотехнике и других технических науках. Для фракционного анализа недоступны задачи типа “черного ящика”. Обычная теория подобия позволяет находить критерии подобия и для таких задач, для которых математических уравнений нет. (см. выше задачи по формированию критериев Рейнольдса и Релея). Однако, когда они есть, она сводит хорошо поставленную краевую задачу к более бедному по содержанию параметрическому уравнению (1), тем самым теряя часть информации, заложенной в уравнениях.

Фракционный анализ извлекает из этих уравнений больше информации. Эта информация не столь полна по сравнению с той, что добывается путем математического решения краевой задачи, но часто ее вполне достаточно для инженерной практики. Кроме того, математические методы иногда очень сложны, требуют соответствующих знаний и квалификации. При сложных

граничных условиях часто решение становится невозможным или сомнительным. Фракционный анализ не требует фундаментальных математических знаний, и сравнительно прост даже при сложных граничных условиях. К тому же, он направлен на экспериментальные методы постановки и решения задачи, без которых практически не обойтись при исследовании любого процесса, даже если задача хорошо решается математически.

Фракционный анализ требует таких же условий в постановке задачи, что и методы математической физики. Корректной считается такая задача, в которой решение существует и является единственным. Во многих задачах требуются сложные доказательства того, что такие условия отражены в ее постановке, хотя большинство практически важных технических задач логически подразумевают наличие и единственность решения, которые не требуют доказательства. К примеру, никаких сомнений не возникает в том, что в определенных условиях нагрева тела какое-то распределение температур по его сечению всегда существует и является единственным.

Фракционный анализ, как и математические методы, требуют того, чтобы математические функции, применяемые для описания процесса, были аналитическими, то есть непрерывными, однозначными и дифференцируемыми. Для однофакторного случая легко продемонстрировать случаи неаналитического поведения функции. На рис. 10 приведена аналитическая функция одного переменного (рис. 10 а), и далее показаны неаналитические функции: на рис 10 б функция терпит разрыв и в точке разрыва неаналитична, на рис. 10 в функция недифференцируема в точках перелома, на рис 10 г аналитичность теряется в точке, в которой функция уходит в бесконечность.

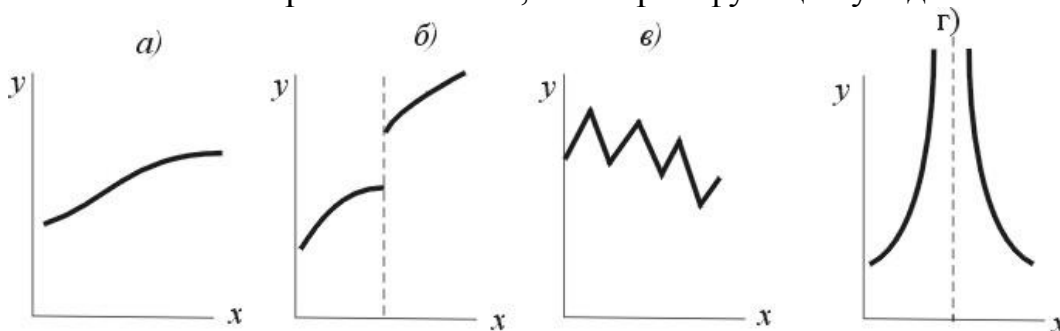


Рис. 10. Аналитическая (а) и неаналитические функции (б – г)

Граничные условия для функций краевой задачи также должны быть конечными и однозначными. Например, непригодны для анализа задачи в областях с границами, уходящими в бесконечность.

Логически легко экстраполировать эти случаи также на многомерное пространство.

3.4. Процедуры фракционного анализа.

Рассмотрим простую краевую задачу.

Допустим, некоторое объемное тело в форме параллелепипеда со сторонами M , N и L посажено в горячую печь с постоянной температурой T_a (рис. 11). Требуется проанализировать процесс нагрева тела, в частности ус-

тановить, через какое время центральная точка тела нагреется до температуры, например, равной $T_c = 0,8T_a$. Это время будем считать временем полного прогрева тела до требуемых нам температур., обозначим его t_c . Важно подчеркнуть, что температуру прогрева ($0,8 T_a$) мы задаем сами.

Основное дифференциальное уравнение теплопроводности известно:

$$\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2} = \frac{1}{a} \cdot \frac{\partial T}{\partial t}. \quad (13)$$

Здесь T – текущая температура тела, зависящая от координат и времени нагрева:

$$T = f(x, y, z, t),$$

a – коэффициент температуропроводности с размерностью $[a] = m^2 / c$.

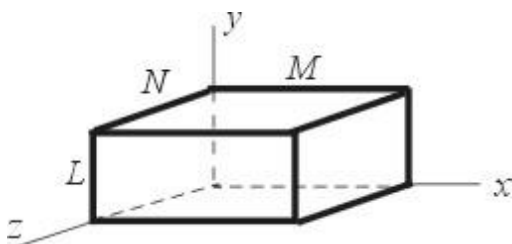


Рис. 11. Размеры нагреваемого тела.

Граничные условия определяются тепловым потоком в направлении нормали к поверхности металла, и, очевидно, таким направлением следует считать наименьший из трех размеров тела. Пусть таким является размер L . Тогда граничное условие записывается в виде уравнения:

$$\lambda \left(\frac{\partial T}{\partial L} \right) = \alpha \cdot (T_a - T_n),$$

где λ – коэффициент теплопроводности, характеризующий способность металла проводить тепло, размерность $вт/(m \cdot K)$ (K – градусы Кельвина),

α – коэффициент теплопередачи, характеризующий интенсивность теплообмена, размерность $вт/(m^2 \cdot K)$,

T_n – температура поверхности нагреваемого тела.

Сначала проанализируем эту задачу традиционными методами теории подобия. Из основного уравнения теплопроводности получаем параметрическое уравнение:

$$T = f(M, N, L, t, a),$$

из граничных условий параметрическое уравнение будет иметь вид:

$$T/(T_a - T_n) = f(L, \alpha, \lambda).$$

Проанализировав размерности уравнения теплопроводности m, c и m^2/c и приняв за независимый параметр L , по П-теореме будем иметь:

$$T = f\left(\frac{L}{M}, \frac{L}{N}, \frac{L^2}{\alpha \cdot t}\right).$$

Первые два критерия геометрические, последний известен в теплотехнике как критерий Фурье $Fu = a \cdot t / L^2$.

Аналогично из граничных условий получается критерий Био:

$$Bi = \alpha \cdot L / \lambda.$$

Кстати, в теории подобия произведение (или частное) двух критериев также является критерием, и в теплотехнике часто используется критерий Био - Фурье, равный $Bi \cdot Fu$.

Теперь те же преобразования осуществим методами фракционного анализа, применив процедуры фракционного анализа к основному дифференциальному уравнению теплопроводности (13). Эти процедуры называются нормализацией и выполняются в два приема. Первый состоит в том, что уравнение приводятся к безразмерным координатам. С помощью второго приема делают безразмерными также коэффициенты при координатах, которые и определяют критерии процесса.

Введем безразмерные координаты следующим образом:

$$\bar{T} = \frac{T}{T_a}, \bar{x} = \frac{x}{M}, \bar{y} = \frac{y}{L}, \bar{z} = \frac{z}{N}, \bar{t} = \frac{t}{t_c}.$$

Следует подчеркнуть, что безразмерные координаты обязательно выбираются так, чтобы пределы их изменения находились в интервале от 0 до 1, так как это дает дополнительную информацию для анализа. В нашем случае все безразмерные координаты удовлетворяют этому условию.

Ход преобразований проследим на первом слагаемом уравнения (13). Сначала запишем в новых координатах первую, а затем вторую производные. Заметим, что

$$T = T_a \cdot \bar{T}, \quad \frac{\partial T}{\partial x} = T_a \frac{\partial \bar{T}}{\partial x}. \quad \text{Поэтому}$$

$$\frac{\partial T}{\partial x} = T_a \cdot \frac{\partial \bar{T}}{\partial \bar{x}} = T_a \cdot \frac{\partial \bar{T}}{\partial \bar{x}} \cdot \frac{\partial \bar{x}}{\partial x} = \frac{T_a}{M} \cdot \frac{\partial \bar{T}}{\partial \bar{x}},$$

$$\frac{\partial^2 T}{\partial x^2} = \frac{\partial}{\partial x} \left(\frac{T_a}{M} \cdot \frac{\partial \bar{T}}{\partial \bar{x}} \right) = \frac{T_a}{M} \cdot \frac{\partial}{\partial \bar{x}} \left(\frac{\partial \bar{T}}{\partial \bar{x}} \right) \cdot \frac{\partial \bar{x}}{\partial x} = \frac{T_a}{M^2} \cdot \frac{\partial^2 \bar{T}}{\partial \bar{x}^2}.$$

Аналогично преобразуются остальные переменные, и окончательно получим:

$$\frac{T_a}{M^2} \cdot \frac{\partial^2 \bar{T}}{\partial \bar{x}^2} + \frac{T_a}{L^2} \cdot \frac{\partial^2 \bar{T}}{\partial \bar{y}^2} + \frac{T_a}{N^2} \cdot \frac{\partial^2 \bar{T}}{\partial \bar{z}^2} = \frac{T_a}{a \cdot t_c} \cdot \frac{\partial \bar{T}}{\partial \bar{t}}.$$

Первая операция нормализации завершена, но при безразмерных координатах имеются размерные множители. Вторая операция нормализации направлена на преобразование их также к безразмерному виду, для этого достаточно все слагаемые умножить на L^2 / T_a . Получим:

$$\frac{L^2}{M^2} \cdot \frac{\partial^2 \bar{T}}{\partial \bar{x}^2} + \frac{\partial^2 \bar{T}}{\partial \bar{y}^2} + \frac{L^2}{N^2} \cdot \frac{\partial^2 \bar{T}}{\partial \bar{z}^2} = \frac{L^2}{a \cdot t_c} \cdot \frac{\partial \bar{T}}{\partial \bar{t}}.$$

Получившиеся безразмерные коэффициенты являются критериями процесса:

$$\frac{L}{M}, \frac{L}{N}, \frac{L^2}{a \cdot t_c}.$$

Ранее эти же критерии были получены традиционным путем с помощью П-теоремы. Так в чем же достоинства фракционного анализа?

Во-первых, помимо критериев в задаче фигурируют безразмерные координаты \bar{x} , \bar{y} , \bar{z} и \bar{t} и безразмерный выходной параметр \bar{T} . Это расширяет возможности моделирования процессов. Прежде всего, можно развить третью теорему подобия и сформулировать четвертую теорему: если в модели и в натуре процессы подобны, то в них равны не только критерии, но и безразмерные координаты.

Следовательно, при равных критериях и одинаковых безразмерных входных параметрах (безразмерных координатах) в модели и в натуре одинаковые значения принимают выходные безразмерные параметры. В нашей задаче при равенстве критериев процесса и при одинаковых значениях координат \bar{x} , \bar{y} , \bar{z} и \bar{t} безразмерная температура \bar{T} в модели и в натуре будет одинаковой. Кривые нагрева двух подобных образцов в точках с одинаковыми безразмерными координатами \bar{x} , \bar{y} , \bar{z} через равное безразмерное время \bar{t} совпадают, если температуру их измерять в безразмерных параметрах \bar{T} . Такой результат непосредственно не следовал из традиционной теории подобия. На основе 4-й теоремы подобия при необходимости можно строить по иному процесс моделирования объектов.

Во-вторых, традиционная теория подобия дает критерии в общем виде, а после нормализации в фракционном анализе те же критерии имеют почти конкретное значение. Доказано, что при правильном задании безразмерных координат, когда они изменяются в пределах от 0 до 1, критерии принимают постоянное значение C , в большинстве случаев близкое к 1. Это позволяет проводить приближенные вычисления. В рассмотренной задаче о нагреве общая теория подобия дает критерий Фурье в виде $Fu = L^2 / at$, тогда как во фракционном анализе получен критерий в форме $Fu = L^2 / at_c$, причем значение его близко к 1. Это позволяет приближенно определить время прогрева тела t_c : если принять $L^2 / at_c = 1$, то $t_c = L^2 / a$. Чаще всего, только ради определения времени прогрева тела в печи и ставится задача теплопроводности. В данном случае она решена приближенно методом фракционного анализа без математического решения и счетно-решающей техники, и даже без проведения эксперимента. В этом смысле фракционный анализ дает больше информации, чем П-теорема подобия. Фракционный анализ очень полезен при приближенных вычислениях.

3.5. Возможности фракционного анализа.

Фракционный анализ особенно эффективен при сложных граничных условиях, когда математическое решение задачи резко усложняется или становится практически невозможным.

Рассмотрим задачу о нагреве того же параллелепипеда при условии, что температура печи не постоянна, а меняется во времени по синусоидальному закону вокруг среднего значения T_a с амплитудой T_0 (рис. 12) по уравнению:

$$T = T_a + T_0 \cdot \text{Sin} (t / \tau) ,$$

где t - текущее время, t_1 - время полного периода колебаний и τ - обратная частота, $\tau = t_1/2\pi$.

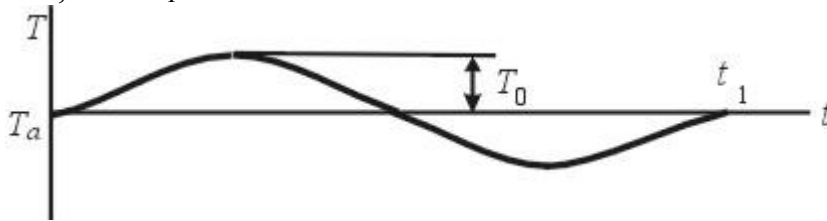


Рис.12. Граничные условия при синусоидальном колебании температуры печи.

По-прежнему будем считать тело полностью прогретым, когда центральная точка достигнет температуры $0,8 \cdot T_a$.

Граничные условия в безразмерных координатах имеет вид:

$$T = 1 + \frac{T_0}{T_a} \cdot \sin\left(\frac{t_c \cdot \bar{t}}{\tau}\right) = 1 + C \cdot \sin\left(\frac{t_c \cdot \bar{t}}{\tau}\right),$$

где C - константа. Видно, что граничное условие такого вида содержит новый критерий $\Pi = t_c/\tau$. Появляется возможность его анализа.

Допустим, $t_c < \tau/4$, то есть $\Pi < 0,25$. Тело прогревается раньше, чем температура печи совершит четверть колебания, колебания печи очень медленны. В этом случае температура всего тела, включая центральную точку, будет повторять почти ту же синусоиду (14), что и температура печи.

Представим обратный случай, когда $t_c/\tau > 1$, то есть колебания температуры печи быстры, и не все слои металла успевают за ними. Колебания температуры сказываются только в поверхностных слоях тела. Пусть на самой поверхности в слое в несколько молекул температура тела точно равна температуре печи, колебания поверхности металла происходят с той же амплитудой, что и печи. Чем дальше от поверхности, тем амплитуда колебания температуры металла будет меньше и в некотором центральном ядре толщиной δ амплитуда равна нулю; внутри ядра тело нагревается так же, как и при постоянной температуре печи T_a , не реагируя на колебания в печи. Можно оценить толщину ядра δ , если задать значение критерия Фурье в ядре, например, $Fu = \delta^2 / ta \geq 3$. Тогда $\delta \geq \sqrt{3ta}$, и если время t принять равным четверти периода $t = \tau/4 = t_1/8\pi$, то толщина ядра определяется однозначно. Конечно, все эти расчеты и рассуждения не претендуют на высокую точность, но как прикидочные они вполне имеют практическое значение.

Фракционный анализ широко применяется для приближенных вычислений. Без решения краевой задачи и без экспериментальных исследований можно получать достаточно полную информацию о процессе. Задача о нагреве параллелепипеда уже дает представление о том, как фракционный анализ используется для приближенной оценки значений параметров, входящих в критерий.

Рассмотрим еще один пример эффективного использования фракционного анализа для приближенных вычислений.

Пусть балка прямоугольного сечения высотой h , шириной b и длиной L , консольно закрепленная одним концом, имеющая массу ρ одного пог. м., отклонена от горизонтали на величину δ и затем совершает свободные незатухающие колебания (рис. 13). Требуется приближенно оценить частоту колебания балки методами фракционного анализа.

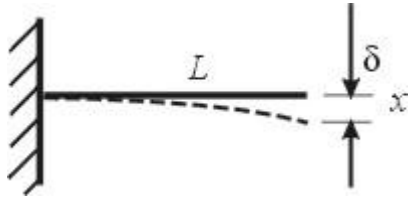


Рис. 13. Задача о колебании балки.

В теории упругости выводится дифференциальное уравнение колебания балки:

$$EI \cdot \frac{\partial^4 \bar{y}}{\partial \bar{x}^4} + \rho \cdot A \cdot \frac{\partial^2 y}{\partial t^2} = 0. \quad (14)$$

Здесь A - площадь поперечного сечения, $A = bh$; E - модуль упругости материала, н/м; I - момент инерции балки, $I = bh^3/12$. Отметим, что в исходном уравнении отсутствует параметр δ как несущественный.

Граничные и начальные условия имеют вид:

- так как на защемленном конце балки перемещения и скорости перемещения равны нулю, то при $x = 0$

$$y = 0 \quad \text{и} \quad \frac{\partial y}{\partial x} = 0;$$

- поперечное сечение свободного конца балки всегда остается плоским, следовательно

$$\text{при } x = L \quad \frac{\partial^2 y}{\partial x^2} = 0;$$

- вертикальные перемещения за счет сил инерции равны нулю, поэтому

$$\text{при } x = L \quad \frac{\partial^2}{\partial x^2} \left(EI \cdot \frac{\partial y}{\partial x} \right) = 0;$$

- при $x = L$ и $t = 0$ $y = \delta$ и $\frac{\partial y}{\partial x} = 0$ (скорость равна нулю).

Точное решение дифференциального уравнения, конечно, дало бы полную информацию о синусоидальных колебаниях каждой точки балки, но в данном случае эта информация избыточна, так как требуется приближенно определить только период колебания. Нет смысла прибегать сразу к мощным математическим средствам, если есть более дешевые способы добыть нужную информацию, к которым относится и фракционный анализ.

Рассмотрим 1/4 -ю часть периода колебаний. Введем новые безразмерные координаты так, чтобы они изменялись в интервале от 0 до 1:

$$\bar{x} = \frac{x}{L}, \quad \bar{y} = \frac{y}{\delta}, \quad \bar{t} = \frac{t}{t_1 / 4} = \omega \cdot t,$$

где ω – обратная величина времени, необходимого для совершения 1/4-периода колебания, t_l – время полного периода (см. рис. 12). Подставим новые координаты в основное уравнение, получим:

$$\frac{\delta \cdot EI}{L^4} \left(\frac{\partial^4 \bar{y}}{\partial \bar{x}^4} \right) + \rho \cdot A \cdot \delta \cdot \omega^2 \cdot \frac{\partial^2 \bar{y}}{\partial \bar{t}^2} = 0 .$$

После первой операции нормализации при безразмерных координатах параметры размерны, поэтому приведем их к безразмерному виду, поделив на $\delta \cdot EI / L^4$:

$$\frac{\partial^4 \bar{y}}{\partial \bar{x}^4} + \frac{\rho AL^4 \omega^2}{EI} \cdot \frac{\partial^2 \bar{y}}{\partial \bar{t}^2} = 0. \quad (15)$$

Из этого уравнения получается критерий

$$\Pi_1 = \frac{\rho AL^4 \omega^2}{EI} = \left(\frac{L^2 \omega}{RC} \right)^2 ,$$

где $R = \sqrt{I / A}$ – радиус инерции сечения балки и

$C = \sqrt{E / \rho}$ – скорость звука в материале балки.

Заметим, что в критерии, как и в уравнении, отсутствует параметр δ .

Фракционный анализ интересен тем, что значение полученного критерия близко к 1, поэтому сразу получаем ответ на поставленную задачу: период колебания балки приближенно равен

$$\omega = RC / L^2 .$$

Точное математическое решение исходного дифференциального решения дает значение $\omega = 0,879 \cdot RC / L^2$. Ошибка на 12% вполне приемлема для инженерных расчетов.

Анализ критерия позволяет сделать еще несколько практически важных выводов. Например, при моделировании натурной балки ее моделью не обязательно соблюдать их геометрическое подобие, следует обеспечить лишь равенство критериев Π_1 , можно моделировать балки одной формы поперечного сечения с помощью балок другой формы.

Граничные условия данной задачи анализировать не интересно, так как они просты и при приведении к безразмерным координатам не дают критериев. Фракционный анализ интересен при сложных граничных условиях. В этом случае математическое решение резко усложняется или становится невозможным, но фракционный анализ сравнительно просто приводит к приближенному решению задачи.

Усложним граничные условия задачи о колебании балки. Пусть на ее свободном конце закреплен тяжелый груз (фланец). При этом уравнение колебания балки (14) не изменится, следовательно не изменится критерий Π_1 , но резко усложнится запись граничных условий. Первое и последнее из приведенных ранее условий сохранятся:

$$- \text{ при } x = 0 \quad y = 0 \quad \text{и} \quad \frac{\partial y}{\partial x} = 0 ;$$

- при $x = L$ и $t = 0$ $y = \delta$ и $\frac{\partial y}{\partial x} = 0$.

Остальные условия изменятся:

- при $x = L$ угловое перемещение фланца описывается уравнением:

$$\frac{\partial^2 y}{\partial x^2} = \frac{J_0}{EI} \cdot \frac{\partial \theta_0^2}{\partial t^2},$$

где J_0 - полярный момент инерции балки и θ_0 - угол поворота фланца;

- при $x = L$ дополнительные вертикальные перемещения фланца за счет сил инерции равны

$$\frac{\partial^3 y}{\partial x^3} = \frac{M_0}{EI} \cdot \frac{\partial^2 y_0}{\partial t^2},$$

где M_0 – масса фланца и y_0 – вертикальное отклонение фланца.

Чтобы получить полную систему граничных условий, необходимо добавить условие совместимости движения фланца и свободного конца балки:

- по вертикальным перемещениям $y_{x=L} = y_0 = \delta$;

- по повороту сечения $\left(\frac{\partial y}{\partial x}\right)_{x=L} = \operatorname{tg} \theta_0 = \frac{\delta}{L}$.

Введем безразмерные переменные $\bar{x}, \bar{y}, \bar{t}$, а также $\bar{\theta} = \partial \bar{y} / \partial \bar{x}$ и $\bar{y}_0 = y_0 / \delta = 1$. Получим граничные условия в безразмерных координатах:

- при $\bar{x} = 0$ $\bar{y} = 0$ и $\partial \bar{y} / \partial \bar{x} = 0$;

- при $\bar{x} = 1$ и $\bar{t} = 0$ $\bar{y} = 1$ и $\partial \bar{y} / \partial \bar{x} = 0$;

- при $\bar{x} = 1$ $\frac{\delta}{L^2} \cdot \frac{\partial^2 \bar{y}}{\partial \bar{x}^2} = \frac{J_0 \omega^2 \delta}{EIL} \cdot \frac{\partial^2 \bar{\theta}_0}{\partial \bar{t}^2}$ и
 $\frac{\delta}{L^3} \cdot \frac{\partial^3 \bar{y}}{\partial \bar{x}^3} = \frac{M_0 \omega^2 \delta}{EI} \cdot \frac{\partial^2 \bar{y}_0}{\partial \bar{t}^2}$.

Если разделить два последних выражения соответственно на δ / L^2 и δ / L^3 , то получим два критерия:

$$\Pi_2 = \frac{LJ_0\omega^2}{EI} \quad \text{и} \quad \Pi_3 = \frac{L^3 M_0 \omega^2}{EI}.$$

Отношение двух критериев также является критерием, поэтому, учитывая, что $\rho AL = M$ - масса балки и $J_0 = M_0 \cdot r_0^2$, где r_0 – радиус инерции фланца, запишем:

$$\Pi_4 = \Pi_2 / \Pi_1 = \frac{J_0}{\rho AL^3} = \frac{M_0}{M} \cdot \frac{r_0^2}{L^2} \quad \text{и}$$

$$\Pi_5 = \Pi_3 / \Pi_1 = \frac{M_0}{\rho AL} = \frac{M_0}{M}.$$

Таким образом, задача в целом определяется тремя критериями: либо в наборе Π_1, Π_2 и Π_3 , либо в наборе Π_1, Π_4 и Π_5 . Последний лучше, так как для

конкретных балки и фланца критерии Π_4 и Π_5 легко вычисляются: массы балки и фланца, а также радиус инерции последнего известны; поэтому критерии Π_4 и Π_5 превращаются в константы, не зависящие от характеристик колебания. Анализ параметров колебания и их моделирование можно осуществлять на балке без фланца, скорректировав результат некоторыми постоянными числами Π_4 и Π_5 .

3.6. Естественные координаты системы .

Идея использования безразмерных координат привела к фракционному анализу, при котором в таких координатах и параметрах (критериях) сосредоточена обобщенная информация о процессе. Можно и дальше развивать эту идею, объединив безразмерную координату и критерий в единую обобщенную координату, в которой будет сосредоточена вся информация об объекте.

Такая координата носит название *естественной*. Применение естественных координат открывает новые возможности для физического моделирования объектов и процессов.

Рассмотрим колебания груза массой m на пружине с жесткостью k , отклоненного от спокойного положения на расстояние δ (рис. 14). Дифференциальное уравнение колебания груза известно:

$$m \cdot \frac{d^2 x}{dt^2} + kx = 0.$$

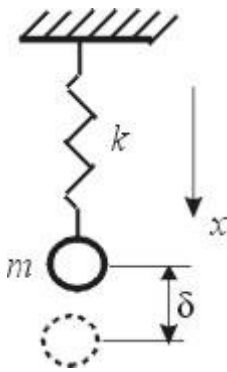


Рис. 14. Задача о колебании груза, подвешенного на пружине.

Начальные (и граничные) условия:
при $t = 0$ $x = \delta$ и скорость $dx/dt = 0$.

Введем безразмерные координаты:

$$\bar{x} = x/\delta \quad \bar{t} = t/\tau,$$

где τ – время 1/4 периода колебания.

Получим (после деления на m):

$$\frac{d^2 \bar{x}}{d\bar{t}^2} + \frac{k\tau^2}{m} \cdot \bar{x} = 0 \quad (16)$$

при начальных условиях:

$$\text{при } \bar{t} = 0 \quad \bar{x} = 1 \quad \text{и} \quad d\bar{x} / d\bar{t} = 0.$$

Здесь $k\bar{t}^2 / m$ - критерий, близкий по значению к 1, откуда сразу можно приближенно вычислить время τ и период колебания груза $\omega = 4 / \tau$.

Далее, попытаемся “упрятать” критерий в естественную координату \hat{t} , которую введем путем замены:

$$(\hat{t})^2 = \frac{\bar{t}^2 k \tau^2}{m}, \quad (17)$$

$$\text{откуда } d\hat{t} = d\bar{t} \cdot \sqrt{k\tau^2 / m}, \quad (d\hat{t})^2 = (d\bar{t})^2 \cdot \frac{k\tau^2}{m}.$$

Подставив в (26), получим:

$$\frac{d^2 \bar{x}}{d\hat{t}^2} + \bar{x} = 0. \quad (18)$$

В полученном уравнении нет никаких параметров пружины и груза, имеются только координаты, причем по координате \hat{t} измерение времени производится не в секундах и не в безразмерных единицах, а в критериях. Такая единица измерения времени естественна для данной задачи. Получено самое общее уравнение для колебания любого груза на любой пружине. Отсюда появляется возможность моделирования процессов по естественным координатам. В нашем случае, для любых грузов и пружин при равенстве естественных координат будут одинаковые характеристики колебаний.

Можно сформулировать 5-ю теорему подобия, которая обобщает (но не отменяет) 3-ю и 4-ю:

если для какой-то группы однотипных процессов записано одинаковое уравнение в естественных координатах, то необходимым и достаточным условием подобия таких процессов служит равенство их естественных координат при пропорциональности сходственных величин, входящих в начальные и граничные условия;

или короче, при одинаковых начальных и граничных условиях в точках с одинаковыми естественными координатами процессы протекают одинаково.

Из выражения (17) наглядно видна разница между 4-й и 5-й теоремами: если 4-я требует, чтобы для подобных процессов были равными и безразмерное время \bar{t} , и критерий $k\bar{t}^2 / m$, то 5-я утверждает, что для подобия достаточно равенства их произведения, образующего естественную координату.

По естественным координатам можно строить так называемые искаженные модели, в которых критерии не равны, но равны естественные координаты. Так при моделировании колебания большого груза на лабораторной установке равенство критериев можно обеспечить только путем соответствующего подбора жесткости пружины и веса груза, тогда как равенство естественных координат можно обеспечить также за счет варьирования иных характеристик колебания, например, содержащих время.

Вернемся к задаче о колебании консольной балки (рис. 13). Там было записано уравнение (15) в безразмерных координатах:

$$\frac{\partial^4 \bar{y}}{\partial \bar{x}^4} + \left(\frac{L^2 \omega}{RC} \right) \cdot \frac{\partial^2 \bar{y}}{\partial \bar{t}^2} = 0.$$

Введем новую переменную

$$\hat{x} = \bar{x} \cdot \sqrt{\frac{L^2 \omega}{RC}}$$

и вычислим новую производную:

$$\frac{\partial^4 \bar{y}}{\partial \bar{x}^4} = \frac{\partial^2}{\partial \bar{x}^2} \left(\frac{\partial^2 \bar{y}}{\partial \bar{x}^2} \cdot \frac{\partial \bar{x}}{\partial \hat{x}} \right) = \frac{L^2 \omega}{RC} \cdot \frac{\partial^2}{\partial \bar{x}^2} \left(\frac{\partial^2 \bar{y}}{\partial \bar{x}^2} \right) = \left(\frac{L^2 \omega}{RC} \right)^2 \cdot \frac{\partial^4 \bar{y}}{\partial \hat{x}^4}.$$

Подставив ее в исходное уравнение, получим

$$\frac{\partial^4 \bar{y}}{\partial \hat{x}^4} + \frac{\partial^2 \bar{y}}{\partial \bar{t}^2} = 0.$$

Уравнение содержит только естественные координаты, внутрь которых как новая единица измерения длины вошел критерий $\Pi = \sqrt{RC / L\omega}$. Полученное уравнение справедливо для любой консольной балки. Можно моделировать колебания круглой или более сложной по сечению натурной балки моделью с прямоугольным сечением. Возможности моделирования по естественным координатам расширяются по сравнению с моделированием по критерию.

Еще раз рассмотрим колебания маятника (рис. 5). При малых углах $\sin \varphi = \varphi$. Введем безразмерные координаты:

$$\bar{\varphi} = \varphi / \varphi_0 \quad \bar{t} = t / T.$$

Подставив их в уравнение (4), получим

$$\frac{d^2 \bar{\varphi}}{d\bar{t}^2} + \frac{T^2 g}{l} \cdot \bar{\varphi} = 0.$$

Как видно, получен известный уже критерий $T^2 g/l$, приравняв который постоянной величине C , можно экспериментально найти период колебания T . Раньше еще не было уверенности в том, что значение критерия будет постоянным. Это было определено только после построения экспериментального графика (рис. 6), но фракционный анализ дает такую уверенность без проведения эксперимента.

Введем новую переменную

$$\hat{t}^2 = \frac{\bar{t}^2 T^2 g}{l},$$

получим уравнение маятника в естественных координатах:

$$\frac{d^2 \bar{\varphi}}{d\hat{t}^2} + \bar{\varphi} = 0 \quad (19)$$

Критерий вошел в естественную координату. Если сравнить уравнения (18) и (19), то увидим, что они полностью совпадают. В естественных координатах процессы колебания груза на пружине и колебания маятника на нити одно-

типны. Естественные координаты помогли объединить два, казалось бы, разных процесса в один класс. Появилась возможность моделировать колебания груза с помощью маятника и наоборот. Возможности моделирования в естественных координатах еще больше расширяются.

3.7. Общие выводы по теории подобия и фракционному анализу.

Таким образом, теория размерностей, теория подобия и фракционный анализ можно рассматривать, с одной стороны, как самостоятельный аппарат, позволяющий правильно организовать процесс физического моделирования процесса, с другой стороны, как подготовительный этап при проведении многофакторного эксперимента.

Теория размерностей учит правильно организовать входные и выходные параметры процесса. Недопустимо несовпадение размерностей левой и правой частей уравнения. П-теорема подобия позволяет записать параметрическое уравнение (1) в виде связей между безразмерными параметрами, что значительно уменьшает количество входных факторов и упрощает эксперимент. Еще большая информация заключена в таком уравнении, если безразмерные параметры являются критериями процесса. Во-первых, критериев может быть меньше, чем безразмерных параметров, поскольку критерии забирают на себя всю информацию о какой-то стороне процесса, следовательно, другие параметры попадают в область автомодельности. Во-вторых, если процесс зависит от нескольких критериев, то практически всегда эти критерии не взаимодействуют между собой, то есть функция, которая описывает этот процесс, будет сепарабельной. Исследование таких функций, как отмечалось ранее, значительно упрощается как при активных, так и при пассивных экспериментах. На подготовительном этапе эксперимента приведение исходного размерного параметрического уравнения (1) к безразмерному, а лучше, к критериальному виду, практически обязательно.

Теория подобия и фракционный анализ широко применяются также как самостоятельный аппарат экспериментального исследования объектов. Они позволяют правильно поставить эксперимент по физическому моделированию объекта, правильно подобрать параметры модели, которая будет соответствовать натурному объекту. В ряде случаев удается получить точный или приблизительный ответ на поставленные вопросы вообще без проведения эксперимента. Несмотря на то, что, как отмечалось, моделирование на упрощенных моделях практически никогда не дает полной информации о натурном объекте, такое моделирование всегда проводится, даже является обязательным. И, конечно, если методами фракционного анализа или как-то иначе получена некоторая предварительная информация об ожидаемом результате, то это значительно облегчает проведение самого эксперимента.

В заключение, полезно “сверху” посмотреть на развитие теории подобия и фракционного анализа. Видно, что развитие этих наук идет по пути концентрации информации внутри параметров, участвующих в задаче. От пара-

метрического уравнения (1), содержащего большое количество малоинформативных размерных параметров мы сначала перешли к меньшему количеству более информативных безразмерных факторов, а затем – к критериям. Еще большая информация в концентрированном виде заключается в естественных координатах.

Рассматриваемые ниже дисперсионный, корреляционный и регрессионный анализы принципиально одинаково работают с любым уравнением связи входных и выходных параметров. В общем виде, параметрическое уравнение (1) можно применять всегда в однотипной записи, подразумевая под входными параметрами либо размерные, либо безразмерные параметры, либо критерии или естественные координаты. Очевидно, прежде чем приступить к построению регрессионных или дисперсионных экспериментов, необходимо проделать все необходимые процедуры с параметрами задачи методами теории подобия и фракционного анализа. При этом, возможно, требуемую информацию удастся добыть без эксперимента, а если эксперимент окажется необходимым, то он упростится, так как сократится количество факторов при повышении информативности каждого фактора.

Контрольные вопросы по главе 3.

1. Какие возможности обеспечивает 2-я теорема подобия?
2. Как необходимо проводить эксперимент по требованиям 3-й теоремы подобия?
3. Почему существует мнение, что физическое моделирование в принципе невозможно?
4. С какими функциями работает фракционный анализ?
5. Как ставится краевая задача?
6. С помощью каких процедур выделяются критерии при фракционном анализе?
7. Чем критерий, полученный фракционным анализом, отличается от критерия, полученного из параметрического уравнения?
8. Какие дополнительные возможности моделирования обеспечивает 4-я теорема подобия?
9. Какие возможности имеет фракционный анализ?
10. Что такое естественные координаты?
11. Как записать уравнение в естественных координатах?
12. Какие возможности моделирования связаны с естественными координатами?
13. Почему перед проведением эксперимента необходимо приводить параметры к безразмерному виду, критериям и естественным координатам?

4. Статистическая обработка результатов измерений.

4.1. Нормальный закон распределения

Итак, можно достаточно уверенно считать, что при отсутствии тренда (или после его выделения) измеряемая величина той природы, которую мы рассматриваем, то есть измеряемая в обычных физических единицах измерения, всегда имеет нормальное распределение. Некоторые величины становятся нормально распределенными после их логарифмирования (логнормальный закон). И если измеряемая величина распределена не по нормальному закону, то имеют место только следующие причины: а) не полностью выделен тренд или пропущен некоторый существенный фактор из первой группы, создающий тренд, и б) природа измеряемой величины y иная, и анализ ее следует осуществлять по другим статистическим формулам.

Многие экспериментаторы начинают исследования с того, что доказывают, что они имеют дело с нормально распределенной величиной y , однако полезно заметить, что такое доказательство правомерно только при большом количестве измерений (больше 50). Если их мало, то любой вывод при доказательстве становится недостоверным. При малом количестве опытов лучше и грамотнее принять закон нормальным без доказательства, ссылаясь на природу измеряемой величины.

Во всех случаях, когда измеряемая величина распределена не нормально, строго говоря, нельзя пользоваться математическим аппаратом дисперсионного и регрессионного анализов, описываемых ниже. Однако существует теорема, утверждающая, что при его использовании и в этом случае ошибки в выводах не бывают большими. В связи с этим часто пользуются формулами статистического анализа измерений, не задумываясь по поводу закона распределения измеряемой величины, хотя всегда полезно иметь в виду приведенные выше замечания.

Заметим также, что нормально распределенной величиной в нашей постановке задачи является только выходная величина y . Напомним студенту основные знания по теории вероятности и математической статистики, которые были получены им на младших курсах.

Чтобы получить закон распределения любой случайной величины y , необходимо ее многократно измерить. Пусть проведено очень большое число n измерений этой величины ($n \rightarrow \infty$, i - счетчик числа измерений, $i = 1, \dots, n$). По горизонтальной оси будем откладывать значения y , а по вертикальной - относительное количество m_k/n опытов, в которых измеряемые значения y оказались меньше заданного значения y_k .

Получим кривую, показанную на рис. 15а, которая называется *функцией распределения* $F(y)$. Ясно, что $m_k/n = 0$ при $y_k = -\infty$ и $m_k/n = 1$ при $y_k = +\infty$. По теореме Якоби отношение $F(y_k) = m_k/n$ есть вероятность того, что любое очередное i -е измерение принимает значение $y_i \leq y_k$:

$$p_i(y_i \leq y_k) = F(y_k).$$

Вероятность того, что измеряемое значение y_i будет лежать в интервале от y_1 до y_2 , равно разности значений функции распределения на концах этого интервала:

$$p(y_1 \leq y_i \leq y_2) = F(y_2) - F(y_1).$$

Наглядной характеристикой любого закона распределения, в том числе нормального, служит *плотность распределения* $f(y)$, являющаяся производной от $F(y)$ и связанная с ней равенством:

$$F(y_i) = \int_{-\infty}^{y_i} f(y) dy. \quad (20)$$

Вероятность $p_i(y_i \leq y_k) = F(y_i)$ определяется площадью фигуры, заключенной между осью y и кривой $f(y)$ и находящейся левее значения $y = y_k$.

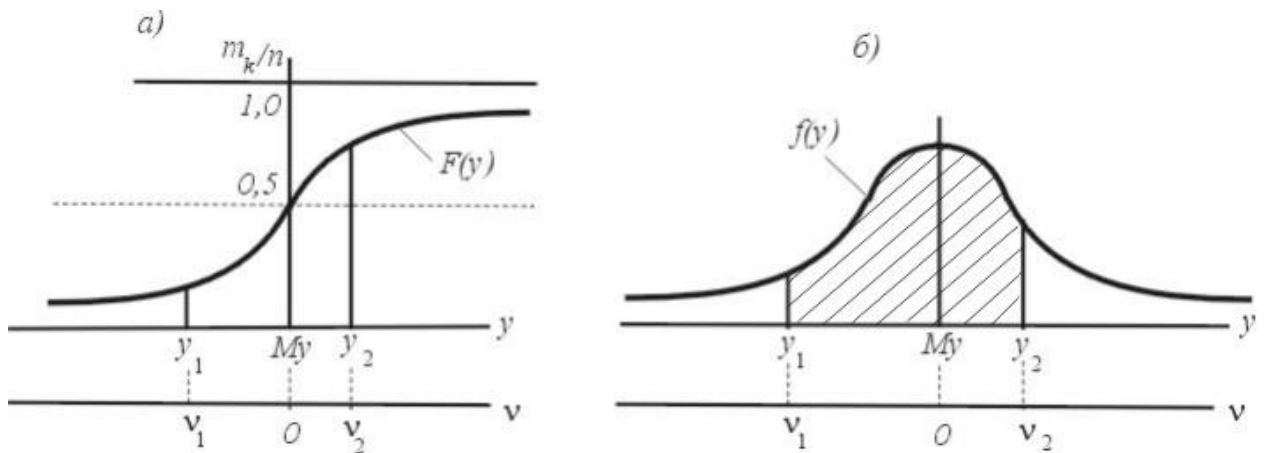


Рис. 15. Функция распределения (а) и плотность распределения (б) при нормальном распределении

Соответственно, площадь под кривой в интервале от y_1 до y_2 (заштрихованная на рис. 9б) определяет вероятность попадания очередного измерения y_i в заданный интервал:

$$p(y_1 \leq y_i \leq y_2) = \int_{y_1}^{y_2} f(y) dy.$$

Кривая плотности распределения для нормального закона распределения интересна тем, что для нее имеется уравнение (названное именем Гаусса, хотя оно было известно французским математикам задолго до Гаусса):

$$f(y) = \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot \exp\left(-\frac{(y - My)^2}{2 \cdot \sigma^2}\right). \quad (21)$$

Кривая простирается от $-\infty$ до $+\infty$, хотя применяется для конечных интервалов изменения реальных измеряемых величин. Кривая Гаусса определяется двумя величинами:

- математическим ожиданием My , которое представляет собой наиболее вероятное значение измеряемой величины y и при бесконечном количестве измерений равно точному значению этой величины, и

- дисперсией σ^2 , которая характеризует кучность расположения данных (разброс) относительно My . Чем меньше разброс данных, тем меньше дисперсия, тем острее и выше купол кривой Гаусса, но полная площадь под ней всегда равна 1.

Указанные величины вычисляются по формулам:

$$My = \frac{1}{n} \cdot \sum_{i=1}^n y_i = \bar{y},$$

$$\sigma^2 = \frac{1}{n} \cdot \sum_{i=1}^n (y_i - My)^2. \quad (22)$$

Введем новую переменную, которая называется *квантилем*:

$$v = \frac{y - My}{\sigma}, \text{ откуда } dv = dy / \sigma, \quad (23)$$

и подставим в уравнение (21) и затем в (20), получим:

$$F(v) = \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{v_i} \exp(-v^2 / 2) \cdot dv.$$

Функция распределения $F(v)$ имеет так называемый нормированный вид и зависит только от одного переменного - квантиля v , поэтому ее можно табулировать (представить в виде таблицы), и полученная таблица будет справедливой для любой нормально распределенной величины. Предварительно заметим, что если ось v переместить по вертикали на значение 0,5, то относительно новой оси v' кривая $F(v')$ будет антисимметричной (рис. 15а) и для новой функции распределения будет справедливо равенство:

$$F(-v'_i) = -F(v'_i),$$

поэтому протабулировать достаточно только положительную ветвь этой кривой, которая носит название функция Лапласа $\Phi(v)$ – см. табл. III приложения.

Из рис. 15 а следует:

$$F(v_k) = 0,5 + \Phi(v_k) = p(v_k).$$

Таблица функции Лапласа позволяет рассчитывать вероятность попадания очередного измерения y_i в заданный интервал значений $y_1 - y_2$. Для этого необходимо:

- произвести серию из достаточно большого числа измерений и вычислить по ней статистические характеристики $My = \bar{y}$ и σ^2 измеряемой величины;
- для очередного измерения y_i и заданных значений интервала y_1 и y_2 вычислить квантили v_i, v_1 и v_2 ;
- вычислить вероятность попадания измерения y_i в требуемый интервал по таблице Лапласа (Приложение П1):

$$p(y_1 \leq y_i \leq y_2) = p(v_1 \leq v_i \leq v_2) = \Phi(v_2) - \Phi(v_1).$$

В технических измерениях часто применяются симметричные интервалы, когда $y_1 = -y_2$ (следовательно, $v_1 = -v_2$). В этом случае

$$p(y_1 \leq y_i \leq y_2) = 2 \Phi(v_2).$$

Решим задачу: определить границы симметричного интервала, в котором окажется очередное измерение y_i с вероятностью (надежностью) 95 %. По таблице функций Лапласа получаем $v_2 = -v_1 = 1,96$.

Итак, с вероятностью 95 %

$$1,96 \leq v_i \leq 1,96,$$

а возвращаясь от квантиля к измеряемой переменной, используя (23), получим:

$$\begin{aligned} My - 1,96 \cdot \sigma \leq y_i \leq My + 1,96 \cdot \sigma \quad \text{или} \\ y_k = My \pm 1,96 \cdot \sigma. \end{aligned} \quad (24)$$

4.2. Полигон распределения и гистограмма.

Если количество измерений величины y не бесконечно, то мы имеем некоторую выборку из бесконечной (генеральной) совокупности измерений. Все сказанное выше о генеральной совокупности требует корректировки при переходе к выборке.

Как и раньше, построим функцию распределения, которая для выборки называется *полигоном распределения*, и функцию плотности распределения, которая теперь выполняется в виде *гистограммы* (рис. 16). Для этого в выборке из n измерений выберем максимальное и минимальное значения измеряемой величины y , определяющих реальный диапазон ее изменения. Разобьем этот диапазон на $k = 10-20$ интервалов равной длины Δy и по вертикальной оси гистограммы отложим относительное количество n_j/n значений y , находящихся в j -м интервале (j – счетчик интервалов, $j = 1, \dots, k$). Для полигона распределения по вертикали отложим $\sum n_j/n$, то есть относительное количество измерений, лежащих в j -м и всех предшествующих интервалах. По степени несовпадения полигона распределения и “теоретической” кривой, перенесенной с рис. 15 (штриховые линии), можно судить об отклонении фактического закона распределения от нормального.

Статистическими характеристиками выборки являются среднее \bar{y} и выборочная дисперсия S^2 , которые определяются по формулам:

$$\bar{y} = \frac{1}{n} \cdot \sum_{i=1}^n y_i; \quad S^2 = \frac{1}{n-1} \cdot \sum_{i=1}^n (y_i - \bar{y})^2. \quad (25)$$

В общем случае выборочное среднее не равно математическому ожиданию My , а выборочная дисперсия S^2 не равна истинной дисперсии σ^2 , и только с увеличением числа измерений в выборке их значения сближаются:

$$\text{при } n \rightarrow \infty \quad \bar{y} = My \quad \text{и} \quad S^2 = \sigma^2.$$

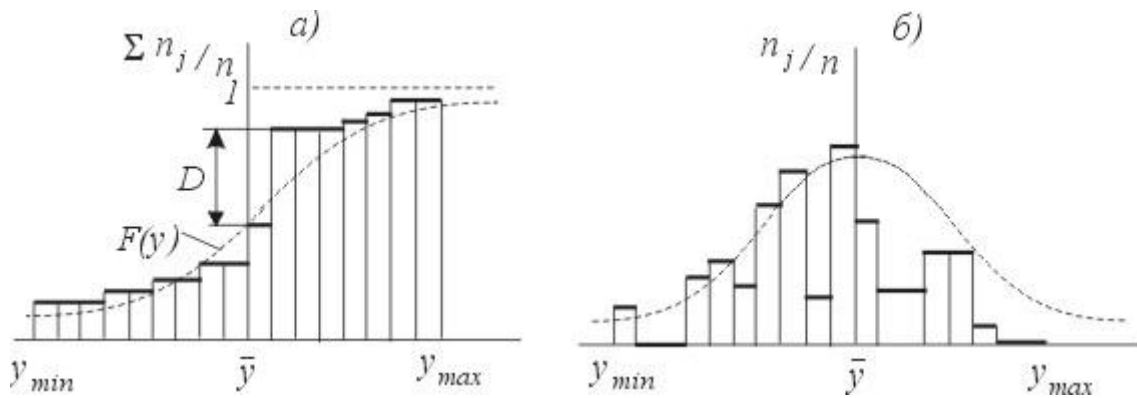


Рис. 16. Полигон распределения (а) и гистограмма (б).

В знаменателе выражения для выборочной дисперсии число $(n - 1)$ называется *числом степеней свободы* f , а числитель этого выражения носит название *суммы квадратов отклонений* SS . Таким образом

$$S^2 = SS / f . \quad (26)$$

Число $S = \sqrt{S^2}$ называют среднеквадратичной погрешностью или *стандарт*ом единичного измерения.

Сумму квадратов отклонений SS предпочитают рассчитывать по более удобной формуле, чем по (25). Проведем следующие преобразования (учитывая, что \bar{y} - число) :

$$\begin{aligned} SS &= \sum_i^n (y_i - \bar{y})^2 = \sum_i^n y_i^2 - 2\bar{y} \cdot \left(\sum_i^n y_i \right) \cdot \frac{n}{n} + \sum_i^n \bar{y}^2 = \\ &= \sum_i^n y_i^2 - 2 \cdot \frac{\sum_i^n y_i}{n} \cdot \frac{\sum_i^n y_i}{n} \cdot n + n \cdot \frac{\left(\sum_i^n y_i \right)^2}{n^2} . \end{aligned}$$

Отсюда, если ввести обозначение

$$y \cdot = \sum_i^n y_i , \quad (27)$$

получаем окончательно:

$$SS = \sum_i^n y_i^2 - \frac{y \cdot^2}{n} \quad (28)$$

Для практических расчетов иногда удобно выбрать некоторое произвольное число a_0 , при котором разность $(a_0 - y_i)$ содержит мало значащих цифр. Тогда

$$\bar{y} = a_0 + \frac{1}{n} \cdot \sum_i^n (y_i - a_0); \quad SS = \sum_i^n (y_i - a_0)^2 - n(\bar{y} - a_0)^2 . \quad (29)$$

Если раскрыть скобки, то можно убедиться в том, что равенства (29) выполняются при любом a_0 .

При большом количестве измерений ($n > 100$) удобно весь диапазон изменения величины y разбить на $\kappa = 8-20$ равных интервалов протяженностью

Δy . Всем n_j измерениям, укладываемым в один j -й интервал, присваиваем одинаковые значения y_j по середине интервала. Тогда

$$\bar{y} = \frac{1}{n} \cdot \sum_j^k n_j \cdot y_j; S^2 = \frac{1}{f} \cdot \left[\sum_j^k (n_j \cdot y_j^2) - \frac{(\sum n_j y_j)^2}{n} \right] - \frac{(\Delta y)^2}{12}. \quad (30)$$

Ниже мы будем иметь возможность посчитать статистические характеристики по всем приведенным формулам.

4.3. Проверка нормальности закона распределения.

Выше отмечалось, что проверку нормального закона распределения случайной величины y обычно не производят. Но если такая необходимость возникнет, то это производится следующим образом.

Существует два числовых критерия, по которым оценивается совпадение функции распределения $F(y)$ и полигона распределения. Первый из них – критерий Колмогорова – осуществляет оценку по максимальному отклонению D (рис. 16 а):

$$D = \max [| \sum n_j / n - \sum p_j |] = \max [| \sum n_j - \sum (n \cdot p_j) |] / n ,$$

где суммирование производится по всем интервалам, предшествующим тому, в котором наблюдается максимальное отклонение D (включая и этот интервал). С вероятностью 80 % ($p = 0,8$) закон распределения считается нормальным, если $\lambda = D \cdot \sqrt{n}$ не превышает значения 1,07.

Второй критерий – критерий Пирсона – в какой-то мере оценивает среднее отклонение полигона распределения от функции $F(y)$. Если объединить три первых и три последних интервала, то величина

$$\Pi = \sum_{j=3}^{k-2} \frac{(n_j - np_j)^2}{np_j}$$

будет иметь изученный и также табулированный закон распределения – χ^2 -распределение. Если $\Pi < \chi^2_{\tau}$, где χ^2_{τ} – табличное значение критерия при заданной вероятности p и числе степеней свободы $f = k - 3$, то закон распределения измеряемой величины признается нормальным (с вероятностью p).

Отметим еще раз, что достоверность выводов по обоим критериям зависит от количества измерений n , и если оно меньше 50, то любое заключение имеет низкую достоверность.

Пример. (Здесь и далее все примеры взяты из различных источников, однако автор позволил себе изменить технологическую сущность некоторых из них применительно к задачам металлургии и обработки металлов давлением). Необходимо доказать нормальность закона распределения и найти статистические характеристики выборки из 200 измерений диаметра трубы. Все измерения сгруппированы в $k = 10$ интервалов ($j = 1, \dots, 10$) с шагом 0,2 мм. В первых трех столбцах табл. 1 приведены данные по количеству n_j значений, укладываемых в заданные j -е интервалы. Следующие 4 столбца де-

монстрируют технику расчета критерия Колмогорова и последние 2 – критерия Пирсона.

Выборочная дисперсия и среднее вычислены по формулам (30):

$$S^2 = [1 \cdot 3,2^2 + 5 \cdot 3,4^2 + \dots + 1 \cdot 5,0^2 - (1 \cdot 3,2 + 5 \cdot 3,4 + \dots + 1 \cdot 5,0)^2 / 200] / 199 - 0,2^2 / 12 = 0,06 ;$$

$$\bar{y} = (1 \cdot 3,2 + 5 \cdot 3,4 + \dots + 1 \cdot 5,0) / 200 = 4,08.$$

Рассмотрим пример расчета p_j для 4-го интервала с границами $y_1 = 3,7$ и $y_2 = 3,9$, при которых

$$v_2 = (3,9 - 4,08) / 0,25 = -0,72,$$

$$v_1 = (3,7 - 4,08) / 0,25 = -1,52 .$$

По табл. П1 приложения

$$p_4 = \Phi(-0,72) - \Phi(-1,52) = -0,2643 + 0,4357 = 0,1715,$$

отсюда $np_4 = 200 \cdot 0,1715 = 34,3$.

В 5-м столбце каждое j -е значение получается суммированием данных предыдущего столбца от 1-го до j -го. Аналогично получаем элементы 6-го столбца по данным 3-го. В 7-м столбце дана разность значений двух предыдущих столбцов, умноженная на n . Наибольшая из них равна $D \cdot n = 19,2$, откуда

$$\lambda = 19,2 / \sqrt{200} = 1,36,$$

что больше 1,07, поэтому делаем вывод о том, что закон распределения измерений отличается от нормального (с вероятностью 80 %).

Таблица 1.

j	y_j	n_j	По Колмогорову				По Пирсону		
			np_j	Σnp_j	Σn_j	$D \cdot n$	$A = (n_j - np_j)^2$	A / np_j	
1	3,2	1	0,2	0,2	1	0,8			
2	3,4	5	1,9	2,1	6	3,9	8,41	0,65	
3	3,6	4	10,9	12,9	10	2,9			
4	3,8	18	34,3	47,2	28	19,2	265,69	7,75	
5	4,0	86	59,2	106,4	114	7,6	718,24	12,13	
6	4,2	62	55,7	162,1	176	13,6	39,70	0,71	
7	4,4	14	28,6	190,7	190	0,7	213,16	7,45	
8	4,6	6	8,0	198,7	196	2,7			
9	4,8	3	1,2	199,9	199	0,9	0,49	0,05	
10	5,0	1	0,1	200,0	200	0			
		$\Sigma n_j = 200$						П= 28,75	

Для определения критерия Пирсона три первых и три последних строки объединяются. Для первой суммы получим:

$$A = (n_j - np_j)^2 = [(1 + 5 + 4) - (0,2 + 1,9 + 10,8)]^2 = 8,41.$$

Отсюда в последнем столбце $A / np_j = 8,41 / (0,2 + 1,9 + 10,8) = 0,65$.

Сумма всех значений этого столбца равна $\Pi = 28,75$, что больше приведенного в табл. П3 приложения значения $\chi^2 = 7,8$ (при $p = 0,95$, $f = 6 - 3 = 3$), следовательно по критерию Пирсона гипотеза о нормальности закона рас-

пределения также отвергается. Видимо, измерения содержат невыделенный тренд.

4.4. Оценки статистических характеристик.

При бесконечном числе опытов каждое очередное измерение y_i находится в интервале $\bar{y} \pm 1,96 \cdot \sigma$ с вероятностью 95 %, но при уменьшении числа измерений вероятность попадания в тот же интервал уменьшается или при той же вероятности должен быть увеличен интервал попадания, то есть увеличен коэффициент, который носит название коэффициент Стьюдента:

$$\bar{y} \pm t_p \cdot S.$$

В общем случае $t_p = f(n, p)$. На рис. 17 приведена зависимость коэффициента Стьюдента от числа опытов n . при заданной вероятности 0,95. При $n > 30$ коэффициент Стьюдента примерно равен 2, поэтому доверительный интервал ($\pm t_p \cdot S$) часто оценивают по правилу “двух сигм”, При числе опытов $n > 4$ применимо уже правило “трех сигм”. Табличные значения коэффициентов Стьюдента приведены в табл. П2 приложения.

Заметим, что в литературе по теории надежности вероятность попадания случайной величины y_i в доверительный интервал ($\pm t_p \cdot S$) называется *надежностью*. В этой теории четко разделены эти два понятия. В нашем же случае мы будем употреблять эти два термина в одном значении.

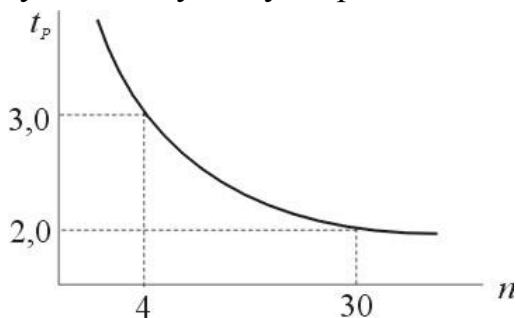


Рис. 17. Зависимость коэффициента Стьюдента от числа опытов при заданной вероятности 0,95.

Среднее значение \bar{y} само является величиной случайной и в каждой j -й серии из n_j измерений принимает значения \bar{y}_j , которое с некоторым приближением оценивает точное значение измеряемой величины M_y . По аналогии с единичным измерением можно записать:

$$\bar{y}_j = M_y \pm t_p \cdot S_y,$$

где S_y^2 – дисперсия среднего \bar{y} , которая связана с дисперсией единичного измерения соотношением:

$$S_y^2 = S^2 / n_j.$$

Дисперсия среднего принадлежит всем n_j измерениям, поэтому в n_j раз меньше дисперсии единичного измерения. Индекс j в приведенных выражениях можно опустить, если измерения сгруппированы в одну выборку. Таким образом, при конечном числе измерений точное значение измеряемой величины M_y мы не получаем, однако устанавливаем границы, в пределах которых с надежностью p оно находится.

Аналогично, дисперсия единичного измерения выборки S^2 является величиной случайной и также с некоторым приближением оценивает точное значение дисперсии σ^2 . Чтобы определить для нее доверительный интервал, воспользуемся величиной «хи-квадрат», определяемой равенством:

$$\chi^2 = \sum_i^n v^2 = \left(\frac{y_i - \bar{y}}{\sigma} \right)^2.$$

Учитывая (25), получим:

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

Как видно, χ^2 связывает истинную и выборочную дисперсии. Как случайная величина она хорошо изучена. Вероятность нахождения величины χ^2 в интервале от χ_1^2 до χ_2^2 также, как и при нормальном законе, определяется разностью соответствующих значений функции распределения $F(\chi^2)$ или площадью под кривой плотности распределения $f(\chi^2)$ (рис. 18) и зависит от числа измерений n (точнее, от числа степеней свободы $f = n - 1$) и заданной вероятности p :

$$\chi_1^2 \leq f \cdot \frac{S^2}{\sigma^2} \leq \chi_2^2.$$

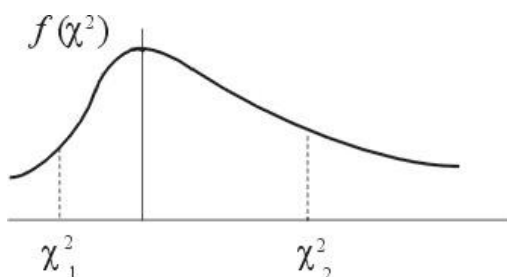


Рис. 18. χ^2 -распределение случайной величины

Обозначив $f / \chi_1^2 = \gamma_2^2$ и $f / \chi_2^2 = \gamma_1^2$, получим неравенство:

$$\gamma_2^2 \cdot S^2 \geq \sigma^2 \geq \gamma_1^2 \cdot S^2, \quad (31)$$

С его помощью определяют доверительный интервал, в котором с надежностью p находится точная дисперсия σ^2 . Табулированные значения χ^2 и γ^2 можно найти в табл. ПЗ приложения.

4.5. Статистические операции.

Рассмотрим некоторые наиболее важные операции, позволяющие решать многие задачи, связанные с обработкой экспериментальных данных.

Исключение промахов

Среди измерений могут попадаться такие, которые не укладываются в ряд с другими данными и подозреваются как промахи. Конечно, самое верное средство выявления промахов – это проверить исправность приборов, мерительного инструмента, установки и условий проведения измерений. И, если

неисправностей и недостатков не обнаружено, то следует использовать статистические методы выявления промахов. В любом случае обнаруженные промахи следует исключать из выборки, так как они очень сильно искажают статистические характеристики измеряемой величины.

Если число измерений очень велико ($n \rightarrow \infty$), то всякое i -е измерение, выходящее за пределы доверительного интервала ($\bar{y} \pm t_p \cdot S$) с вероятностью 95 % является промахом. С уменьшением числа опытов уменьшается уверенность в том, что выпадающее из доверительного интервала значение будет промахом. В этом случае он обнаруживается с помощью критерия совместности θ :

$$\theta = (|y_i - \bar{y}|) / \sqrt{(n-1) \cdot S^2 / n} ,$$

где y_i в данном случае значение, подозреваемое как промах, это либо минимальное, либо максимальное значение в ряду измерений. Предельное значение θ зависит от числа измерений n и заданной вероятности (надежности) p . При $p = 0,95$ значения θ_{\max} приведено в табл. П4 приложения. Если $\theta > \theta_{\max}$, то с вероятностью p подозреваемое значение y_i является промахом.

Пример. В табл. 2 приведено 6 измерений ($n = 6$), среди которых значение 164,3 подозревается как промах.

Таблица 2.

i	1	2	3	4	5	6
y_i	153,4	154,6	154,7	155,0	164,3	154,5

Статистические характеристики удобно посчитать по формулам (29), приняв $a_0 = 155,0$:

$$\bar{y} = 155,0 + (-1,6 - 0,4 + 0,3 + 0 + 9,3 - 0,5) / 6 = 156,08,$$

$$SS = (1,6^2 + 0,4^2 + 0,3^2 + 9,3^2 + 0,5^2) - 6 \cdot (156,08 - 155,0)^2 = 82,8 .$$

Обратим внимание на то, что подозреваемое значение 164,3 участвует в вычислениях статистических характеристик при определении промаха.

Отсюда $S^2 = 82,8 / 5 = 16,56$ и $\theta = (164,3 - 156,08) / \sqrt{5 \cdot 16,56 / 6} = 2,22$.

По табл. П4 приложения при надежности 0,95 и $n = 6$ $\theta_{\max} = 2,0$, поэтому значение 164,3 следует признать промахом (возможно, опiskeй вместо положенного 154,3). Промах обязательно исключается из серии измерений, и после этого обязательно пересчитываются среднее и дисперсия. Получим:

$$\bar{y} = 155,0 + (-1,6 - 0,4 - 0,3 + 0 - 0,5) / 6 = 154,4 ,$$

$$SS = (1,6^2 + 0,4^2 + 0,3^2 + 0,5^2) - 6 \cdot (154,4 - 155,0)^2 = 1,51 .$$

Видно, что статистические характеристики резко изменились, особенно дисперсия (более чем в 50 раз).

Чтобы закончить статистический анализ оставшейся после исключения промаха серии измерений, необходимо определить доверительные интервалы полученных статистических характеристик. Будем иметь:

- дисперсию единичного измерения $S^2 = 1,51 / 4 = 0,38$,

- дисперсию среднего из 5 измерений $S_y^2 = 0,38 / 5 = 0,076$,
- по табл. П2 коэффициент Стьюдента $t_p (f = 4; p = 0,95) = 2,78$ и доверительный интервал среднего, который оценивает точное значение измеряемой величины M_y , $\Delta y = \pm t_p \cdot S_y = 2,78 \cdot \sqrt{0,076} = 0,8$; измеряемая величина с вероятностью 95 % лежит в интервале $154,4 \pm 0,8$;

- по табл. П3 приложения коэффициенты $\gamma_1^2 = 0,359$ и $\gamma_2^2 = 8,26$, поэтому доверительный интервал для дисперсии среднего равен $0,359 \cdot 0,076 \leq S_y^2 \leq 8,26 \cdot 0,076$ или $0,027 \leq S_y^2 \leq 0,628$.

Как видно, доверительные интервалы слишком широки, потому что количество измерений ($n = 5$) очень мало.

Порядок статистической обработки измерений.

На приведенном примере полезно проследить порядок статистического анализа серии многократных измерений величины, имеющей нормальный закон распределения вокруг некоторого среднего значения:

- сначала намечаются и анализируются возможные промахи по критерию совместимости θ ,

- после исключения действительно промахов по оставшимся n опытам вычисляются статистические характеристики измерений: среднее \bar{y} и выборочные дисперсии единичного измерения S^2 и среднего S_y^2 ,

- вычисляются доверительные интервалы для каждой из статистических характеристик: для среднего $\pm \Delta y = t_p \cdot S_y$ и для дисперсии по формуле (31).

При количестве измерений больше 50 можно также доказать, что измерения подчиняются нормальному закону распределения. Если это не так, то необходимо выделить тренд или признать, что в опытах участвует величина иной природы, чем нормально распределенная. В последнем случае пользоваться приведенными статистическими формулами, строго говоря, нельзя, но с определенной осторожностью по ним также можно получить правильные выводы.

Погрешность косвенных измерений

Постановка задачи: необходимо оценить погрешность косвенной величины Z , которая находится в функциональной связи $Z = f(x)$ с измеряемой величиной x . Пусть косвенная и измеряемая величины имеют оценки:

$$Z = \bar{Z} \pm \Delta Z \quad \text{и} \quad x = \bar{x} \pm \Delta x .$$

Так как Δx обычно мало по сравнению с x , то среднеквадратичная ошибка косвенной величины вычисляется следующим образом:

$$\Delta Z = \sqrt{(Z - \bar{Z})^2} = \sqrt{[f(x) - f(\bar{x})]^2} = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 (\Delta x)^2} = f'(x) \cdot \Delta x .$$

Если величина Z является функцией нескольких измеряемых переменных

$$Z = f(x_1, x_2, \dots, x_n) , \text{ то}$$

$$\Delta Z = \sqrt{\left(\frac{\partial f}{\partial x_1}\right)^2 \cdot (\Delta x_1)^2 + \left(\frac{\partial f}{\partial x_2}\right)^2 \cdot (\Delta x_2)^2 + \dots + \left(\frac{\partial f}{\partial x_n}\right)^2 \cdot (\Delta x_n)^2} .$$

Относительная погрешность равна $\varepsilon = \Delta Z / \bar{Z}$, где $\bar{Z} = f(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)$.

Например, требуется определить погрешность определения объема V цилиндра, когда с некоторой точностью измеряются высота h и диаметр d цилиндра. Известно, что $V = \pi d^2 h / 4$.

Составляем таблицу из N измерений h и d и обычными методами для каждой из них вычисляем средние \bar{h} и \bar{d} и доверительные интервалы Δh и Δd . Затем вычисляем погрешность определения объема

$$\Delta V = \sqrt{\left(\frac{\partial V}{\partial h}\right)^2 \cdot (\Delta h)^2 + \left(\frac{\partial V}{\partial d}\right)^2 \cdot (\Delta d)^2}.$$

Производные равны

$$\frac{\partial V}{\partial h} = \frac{\pi}{4} \bar{d}^2 \quad \text{и} \quad \frac{\partial V}{\partial d} = \frac{\pi}{2} \bar{h} \bar{d}.$$

Относительная погрешность

$$\varepsilon = \frac{\Delta V}{V} = \frac{4\Delta V}{\pi \bar{h} \bar{d}^2}.$$

Подчеркнем еще раз, что приведенные вычисления справедливы при функциональной связи величин. Если эта связь вероятностная, то подобные оценки следует осуществлять методами корреляционного анализа.

Сравнение дисперсий и средних.

Во многих статистических задачах приходится сравнивать между собой дисперсии двух или нескольких выборок. Например, при сравнении точности (доверительных интервалов) при изготовлении деталей по двум разным технологиям необходимо вычислить и сравнить дисперсии представительных выборок деталей, произведенных по этим технологиям. Сравнение дисперсий также широко применяется в дисперсионном и регрессионном анализе и планировании многофакторных экспериментах, излагаемых ниже.

При бесконечном количестве измерений в каждой выборке равенство их дисперсий признается тогда, когда критерий Фишера (F -отношение)

$$F = S_1^2 / S_2^2 = \sigma_1^2 / \sigma_2^2 = 1.$$

Договоримся, что при их неравенстве большая дисперсия всегда будет располагаться в числителе, поэтому F -отношение будет больше или равно 1. Если уменьшать количество опытов в выборках числителя и знаменателя, то достоверность приведенного утверждения будет уменьшаться. С достаточной

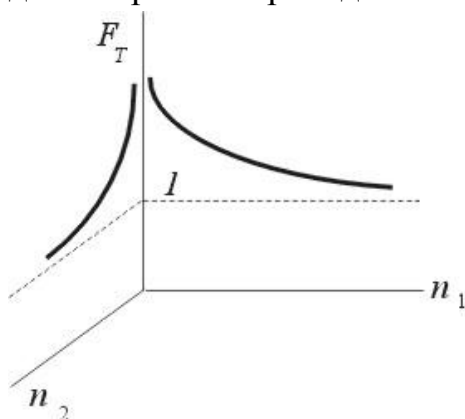


Рис. 19. Зависимость табличного значения F -отношения от числа опытов

надежностью можно считать дисперсии отличными только при значительном их неравенстве, когда F -отношение будет существенно больше 1. Можно построить график предельных значений F_T (рис. 19). Чем меньше опытов в числителе n_1 и в знаменателе n_2 , тем выше значение F_T . В приложении в табл. П6 приведены табличные значения F_T , которые определяются заданной надежностью p (в таблице $p = 0,95$) и числом степеней свободы числителя $f_1 = n_1 - 1$ и знаменателя $f_2 = n_2 - 1$.

Если критерий Фишера выше F_T , то с вероятностью p дисперсии числителя и знаменателя отличаются. И, наоборот, при $F < F_T$ нет оснований считать, что сравниваемые дисперсии статистически различимы. Это вовсе не значит, что они равны, хотя в статистическом смысле во многих статистических задачах неравенство $F < F_T$ свидетельствует о равнозначности сравниваемых дисперсий.

Пример. По первой технологии отклонение детали от номинала определяется рядом измерений: 0,02 ; 0,08 ; 0,03 ; 0,04 ; 0,03 ,
а по второй - рядом: 0,02 ; 0,01 ; 0,01 ; - 0,01.

Спрашивается, можно ли с надежностью 95 % утверждать, что вторая технология обеспечивает большую точность, чем первая, хотя, на первый взгляд, это очевидно? Рассчитаем выборочные дисперсии обеих выборок по формулам (28):

$$S_1^2 = 10^{-4} \cdot [2^2 + 8^2 + 3^2 + 4^2 + 3^2 - (2 + 8 + 3 + 4 + 3)^2 / 5] / 4 = 5,5 \cdot 10^{-4} .$$

$$S_2^2 = 10^{-4} \cdot [2^2 + 1^2 + 1^2 + 1^2 - (2 + 1 + 1 - 1)^2 / 4] / 3 = 1,58 \cdot 10^{-4} .$$

Критерий Фишера $F = 5,5 / 1,58 = 3,48$. Табличное значение F_T ($p = 0,95$; $f_1 = 4$; $f_2 = 3$) = 9,1. С вероятностью 95 % нельзя утверждать, что дисперсии выборок различны, что одна технология лучше другой. Для требуемой уверенности не достаточно количества проведенных измерений.

Когда приходится сравнивать несколько дисперсий между собой, то пользуются критерием Кохрана, который мы рассматривать не будем, так как всегда имеется возможность попарного сравнения дисперсий по критерию Фишера и, кроме того, потребность в критерии Кохрана в наших задачах не встретится.

Сравнение средних значений двух выборок \bar{y}_1 и \bar{y}_2 правомерно тогда, когда дисперсии этих выборок в статистическом смысле равны, то есть критерий Фишера меньше табличного значения. Сравнением средних можно доказать, что обе выборки принадлежат к одной группе исследований, эти выборки полезно соединить в одну выборку, так как статистические характеристики единой выборки будут более достоверными, чем двух отдельных, или при той же достоверности окажутся в более узких доверительных интервалах. Иногда, наоборот, сравнивая средние, доказывают их различие, чтобы доказать эффективность каких-то мероприятий и т. п.

Средние следует признать одинаковыми, если они укладываются в единый доверительный интервал, то есть выполняется условие:

$$|\bar{y}_1 - \bar{y}_2| \leq t_p \cdot S_n \cdot \sqrt{(n_1 + n_2) / (n_1 \cdot n_2)},$$

где S_n^2 – средневзвешенная дисперсия, вычисляемая по формуле:

$$S_n^2 = \left[(n_1 - 1) \cdot S_{y_1}^2 + (n_2 - 1) \cdot S_{y_2}^2 \right] / f,$$

n_1 и n_2 – число опытов, $S_{y_1}^2$ и $S_{y_2}^2$ – дисперсии первой и второй выборок, $f = n_1 + n_2 - 2$ – общее число степеней свободы выборок. Коэффициент Стьюдента t_p выбирается по табл. П2 приложения при заданной надежности p и числе степеней свободы f .

Если дисперсии выборок различны, то это свидетельствует о том, что различно их происхождение, и сравнивать средние следует осторожно. Может быть, в каких-то случаях сравнение не правомерно.

Пример. Требуется установить, различаются ли производительности в дневную и ночную смены? Наблюдения в течение месяца ($n_1 = n_2 = 30$) показали средние значения производительности днем $\bar{y}_1 = 62,9$ и ночью $\bar{y}_2 = 62,4$ (в каких-то единицах продукции) при соответствующих дисперсиях $S_{y_1}^2 = 0,8$ и $S_{y_2}^2 = 0,6$. Кажется, производительности одинаковы.

Сравниваем дисперсии по критерию Фишера:

$$F = S_{y_1}^2 / S_{y_2}^2 = 0,8 / 0,6 = 1,33,$$

что меньше табличного значения $F(0,95; 29; 29) = 1,8$, следовательно, дисперсии неразличимы, поэтому средние сравнивать можно.

Средневзвешенная дисперсия

$$S_n^2 = (29 \cdot 0,8 + 29 \cdot 0,6) / (30 + 30 - 2) = 0,7 \text{ и } S_n = 0,83.$$

По табл. П2 $t_p = 1,99$ и $q = 1,99 \cdot 0,83 \cdot \sqrt{(30 + 30) / 30 \cdot 30} = 0,42$.

Разность $(62,9 - 62,4) > q$, поэтому признаем, что производительности в дневную и ночную смены различны, что, на первый взгляд, было не очевидно.

4.6. Построение экспериментальных кривых

При обработке экспериментального материала часто результаты опытов представляются в виде графиков. Многие исследователи не проводят статистическую обработку этих графиков, считая такую работу не обязательной, отнимающей время ради некоторого «бантика», приводящей к очевидным, на их взгляд, выводам. Однако пренебрежение этим важным этапом исследования иногда приводит к ошибкам и даже неверным выводам по работе в целом. Рассмотрим наиболее характерные ошибки на простых примерах.

Пусть исследуется зависимость y (выходной параметр) от x (входной параметр). В каждой точке эксперимента (то есть при одном значении x) проведено по $n > 1$ опытов. Количество уровней изменения фактора x равно m . В данном случае (при однофакторном эксперименте) количество уровней совпадает с количеством экспериментальных точек. Для определенности, пусть $n = 3$, $m = 5$ (рис. 20). Студент (исследователь) соединил средние значения в

точках прямыми линиями, как показано на рис.20, и сделал следующие выводы:

- величина y при увеличении x сначала растет, а затем падает;
- максимальное значение y при $x = 3$ обеспечивает наиболее благоприятные условия протекания исследуемого процесса .

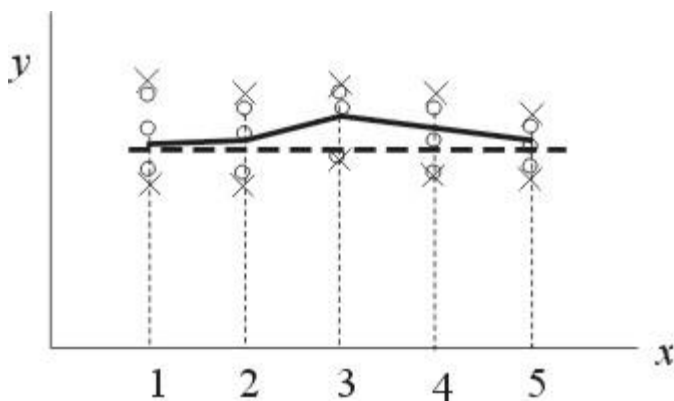


Рис. 20. Обработка однофакторного эксперимента.

Но если бы по полученным данным был получен доверительный интервал, который откладывается по вертикали вверх и вниз в каждой точке эксперимента (границы доверительного интервала обозначены крестиками), то не исключено, что выводы будут другими. В области, ограниченной доверительными интервалами, правомерна любая кривая. Проведенная по средним значениям ломаная кривая не имеет никакого преимущества перед любой другой кривой или прямой в этом интервале. Как видно из рисунка, допустима также прямая (штриховая), параллельная оси x . А последняя дает основания сделать вывод о том, что величина y не зависит от величины x . По крайней мере, из проведенной серии опытов такая зависимость не выявлена. Именно такой вывод и следует признать предпочтительным по результатам данного эксперимента.

Другой пример. Пусть при тех же входном и выходном параметрах, и также при $n = 3$ и $m = 5$ студентом проведена некоторая кривая второго порядка, проходящая близко к средним значениям $y_{\text{ср}}$ (рис. 21). Ему

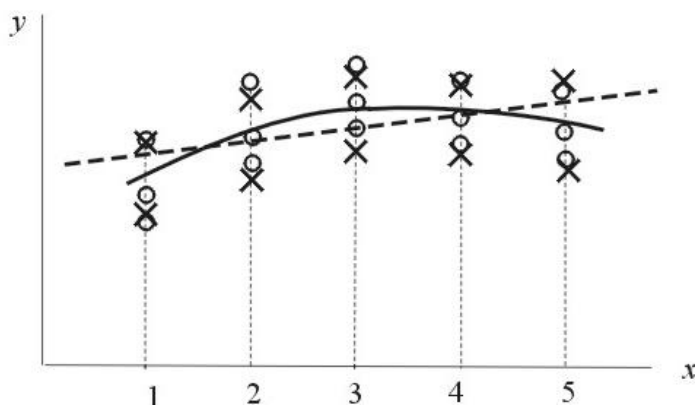


Рис. 21. Обработка экспериментальных данных

показалось, что такой характер кривой наилучшим образом описывает результаты исследования и лучше соответствует сути исследуемого процесса. Сделан вывод о том, что y и x связаны квадратичной зависимостью. Но если в каждой точке эксперимента от среднего значения отложить доверительный

интервал, то оказывается, что в этот интервал вполне вписывается прямая линия (показана штрихами). В этом случае простая кривая (у нас прямая) более предпочтительна, чем сложная. Здесь грамотным выводом будет утверждение, что указанные величины связаны линейной зависимостью. По крайней мере, приведенных данных не достаточно, чтобы достоверно подтвердить наличие квадратичной зависимости.

Из этих примеров следует, что сложившийся стереотип о том, что статистика не вносит никакой дополнительной информации при исследовании и является только украшением эксперимента, не верен. Она позволяет более грамотно и строго формулировать выводы по работе, заставляет исследователя придерживаться определенной дисциплины при обработке результатов опытов.

Как же построить доверительный интервал? Существует две методики, которые рассмотрим на тех же примерах.

Так как в каждой точке эксперимента мы имеем $n > 1$ опытов, то для каждой из них будем строить доверительный интервал $\pm \Delta y_u$, где u – номер точки эксперимента (уровня фактора x), $u = 1, 2, \dots, m$. Порядок построения доверительного интервала в одной точке приведен в разделе 4.5. Естественно, доверительный интервал откладывается в каждой точке от каждого среднего \bar{y}_u вверх и вниз.

Такая методика имеет следующий недостаток. В расчете доверительных интервалов участвуют только n опытов. Как правило, это число мало, в нашем примере $n = 3$. При малом количестве опытов, во-первых, доверительные интервалы получаются большими, так как коэффициент Стьюдента t_p и стандарт среднего S_y велики, и, во-вторых, достоверность любого вывода при малом количестве опытов невелика. При $n = 2 - 5$ даже тот вывод, который следует из статистической обработки эксперимента, также мало надежен, как противоположный.

Чтобы исправить этот недостаток, полезно все mn опытов объединить в единую выборку. Точнее, можно объединять в одну выборку не сами значения y , а отклонения $(y - \bar{y}_u)$. В основе такой возможности лежит предположение, что отклонения в каждой точке эксперимента (при любом значении u) будут примерно одинаковыми, с одинаковыми характеристиками нормального закона распределения относительно среднего отклонения, равного 0 в каждой точке. Такое предположение оправдано, если в ходе эксперимента не меняется точность измерений, остаются постоянными условия проведения эксперимента, ничто не ломается и т. д.

Таким образом, по данным эксперимента строим новую таблицу наблюдений, в которой содержится $N = mn$ опытов с новым выходным параметром $(y - \bar{y}_u)$. Статистические характеристики новой выборки определяются по стандартной методике. Доверительный интервал будет одинаковым для всех экспериментальных точек.

Что же мы выгадали? Во-первых, эти характеристики получены по mt опытом, следовательно, доверительный интервал будет уже, чем в первом случае. Достоверность всех выводов возрастает. Во-вторых, появляется возможность строить доверительный интервал и тогда, когда в каждой точке эксперимента проведено только по одному опыту ($n=1$). По первой методике обработка такого эксперимента становится невозможной. Рассмотрим такой случай на примере.

Пусть, имея только по одному опыту в каждой i -той точке эксперимента, мы провели «на глаз» некоторую кривую, проходящую близко к

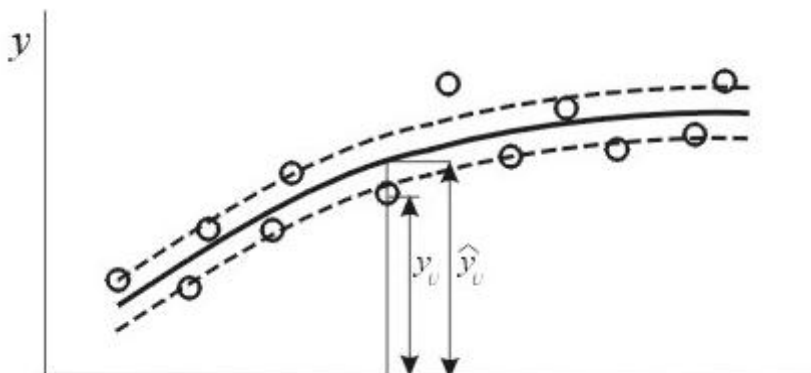


Рис. 22. Однофакторный эксперимент с одним опытом в точке.

экспериментальным данным (рис. 22). В каждой точке (у нас $m = 12$) мы имеем одно экспериментальное значение y_i и одно значение \check{y}_i , лежащее на кривой. Можно все m отклонений ($y_i - \check{y}_i$) объединить в единую выборку и для них провести стандартную статистическую обработку. Доверительный интервал будет общим для всех точек, и его необходимо отложить вверх и вниз относительно кривой. Получим область, ограниченную двумя параллельными кривыми, расположенными на расстоянии Δy от основной кривой. Теоретически в этом интервале должны лежать 95 % экспериментальных точек.

Появляется возможность, например, попытаться уложить в заданный интервал более простую кривую (прямую). И если прямая укладывается, то построенную ранее кривую следует заменить прямой. В нашем примере, прямая линия не укладывается в область доверительного интервала, следовательно, оставляем вывод о параболической связи параметров.

4.7. Упражнения

1. Построить график $y = f(x)$ по средним значениям в виде кривой Безье и вычислить сначала отдельно для каждой точки, а затем общий доверительный интервал (с вероятностью 95 %) по эксперименту, представленному в табл. 3 (5 точек по 4 опыта в точке). Сделать выводы:

- как изменяется доверительный интервал при построении его в каждой точке и по всей выборке в целом;

- какая кривая наилучшим образом укладывается в один и другой доверительный интервал.

Таблица 3

x	y_1	y_2	y_3	y_4
1	2,0	2,3	2,6	2,7
2	2,7	2,9	2,8	2,4
3	2,3	3,8	3,2	3,1
4	3,1	3,5	2,8	3,4
5	2,3	2,4	2,7	3,0

2. Построить график $y = f(x)$ и вычислить общий доверительный интервал (с вероятностью 95 %) по эксперименту, представленному в табл. 4;

Таблица 4.

x	y_1	y_2	y_3
1,2	2,1	2,3	2,4
2,4	3,7	3,9	3,8
3,6	4,3	4,8	4,2
4,4	5,1	5,5	5,8

Какая кривая должна быть проведена по этим данным?

3. Проверить, можно ли провести прямую линию по экспериментальным данным, приведенным в табл. 5:

Таблица 5.

x	y_1	y_2	y_3	y_4
1,3	0,7	1,0	1,2	1,5
3,9	1,4	1,6	2,0	2,2
6,6	3,1	3,4	3,7	4,0
10,3	4,5	5,0	5,5	5,8

4. Дана кривая второго порядка, которая точно проходит через точки, координаты которых отмечены в табл. 6. . В таблице приведены также экспериментальные данные \hat{y} в тех же точках (по три опыта в каждой точке). Найти общий доверительный интервал и попытаться уложить него прямую линию?

Таблица 6

x	\hat{y}	y_1	y_2	y_3
1	4,0	4,6	3,7	4,4
2	3,0	2,3	2,8	3,0
3	2,6	2,0	2,6	3,2
4	1,7	2,0	1,5	1,4

Контрольные вопросы

1. Какими характеристиками определяется нормальный закон распределения?
2. Как вычисляется дисперсия при конечном и бесконечном числе опытов?
3. Чем отличаются дисперсии единичного измерения и среднего?
4. Как производится отсеивание промахов?
5. Как определить доверительный интервал при измерении постоянной величины y_0 ?
6. Какие идеи лежат в основе доказательства нормальности закона распределения случайной величины y ?
7. Каков порядок статистической обработки данных по измерению постоянной величины?
8. Как строится доверительный интервал для кривой, которая проведена между экспериментальными точками?
9. Как построить доверительный интервал для кривой, построенной по средним значениям измерений в точке?
10. Какие задачи можно решать, когда построен доверительный интервал для кривой?

5. Интерполяция

5.1. Постановка задачи

Приведенные выше задачи дают возможность выбрать вид кривой, которая соответствует приведенным экспериментальным данным. Обычно требуется не только начертить такую кривую на графике, но и записать ее уравнение. Ранее указывалось, что любую кривую можно описать полиномом, представляющим собой разложение функции в ряд Тейлора:

$$\begin{aligned} y_p = & a_0 + a_1 x_1 + a_2 x_2 + \dots + a_k x_k + \\ & + a_{12} x_1 x_2 + a_{13} x_1 x_3 + \dots + a_{k-1,k} x_{k-1} x_k + \\ & + a_{11} x_1^2 + a_{22} x_2^2 + \dots + a_{kk} x_k^2 + \dots \end{aligned} \quad (32)$$

В первой строке уравнения записаны линейные члены, во второй – парные произведения первых степеней переменных, которые называются парными взаимодействиями и отражают эффекты совместного влияния факторов на величину y , в третьей – квадратичные члены и т. д. Обычно в реальной задаче стараются написать возможно меньшее количество членов уравнения, но достаточное для правильного описания процесса. Если уравнение получается сложным, многочленным, то выгоднее разбить область изменения входных параметров на ряд подобластей, в которых станут применимыми более простые модели. Чтобы построить такое уравнение, необходимо провести актив-

ный или пассивный эксперимент и построить таблицу наблюдений в некоторых узловых точках факторного пространства (области или подобласти).

Возможны две постановки задачи.

1. Построить такое уравнение, которое во всех или указанных узловых точках с номером u ($u = 1, 2, \dots, N$, где N – количество таких точек) обеспечивает расчетные значения \hat{y}_u , точно совпадающие с экспериментальными значениями y_u . В неузловых точках расхождения значений могут быть любыми и они никак не оцениваются. Операция построения такого уравнения называется *интерполяцией*, а само уравнение – *интерполяционным* полиномом.

2. Необходимо построить простое уравнение (32), содержащее небольшое количество членов и плавно проходящее вблизи экспериментальных значений y_u . В узловых точках в общем случае расчетные значения \hat{y}_u не совпадают с экспериментальными (типа кривой, представленной на рис 22). Сумма квадратов отклонений

$$SS_{ad} = \sum_u^N (\hat{y}_u - y_u)^2 \quad (33)$$

служит мерой их несовпадения или *мерой неадекватности*.

При построении интерполяционного уравнения мера неадекватности всегда равна нулю. Во второй постановке необходимо уметь оценивать меру неадекватности и минимизировать ее. Видимо, при значениях SS_{ad} меньше некоторого допустимого порога модель считается *адекватной*, то есть достаточно хорошо совпадающей с экспериментальными данными. Адекватная модель, построенная по требованиям второй постановки, называется *регрессией*. (Слова “уравнение” и “регрессия” в регрессионном анализе являются синонимами). Регрессионный анализ занимается построением именно таких моделей, а процесс построения интерполяционных уравнений в математике называется интерполяцией.

5.2. Теорема о количестве членов в уравнении.

Как всегда, мы не будем заниматься доказательством теоремы, но всесторонне обсудим суть, а также практические возможности ее использования. Сначала попытаемся на некоторых примерах продемонстрировать проблему, а потом обобщить результаты и сформулировать содержание теоремы.

Рассмотрим однофакторный эксперимент $y = f(x_1)$ – рис. 23.. Пусть проведены опыты пока только в двух точках x_{11} и x_{12} (первый индекс соответствует номеру фактора, второй – номеру точки; так как мы рассматриваем однофакторный эксперимент, первый индекс можно было бы опустить).

По двум значениям y_1 и y_2 проводим прямую 1:

$$y_p = a_0 + a_1 x_1 . \quad (34)$$

Данная модель адекватна, так как точно проходит через экспериментальные точки, она является полиномом $(N - 1)$ степени и содержит два члена ($m = 2$), что равно количеству экспериментальных точек ($N = 2$).

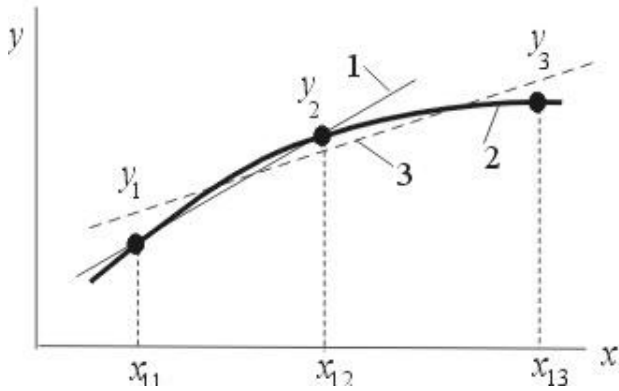


Рис. 23. Вид кривой, проходящей через две и три точки

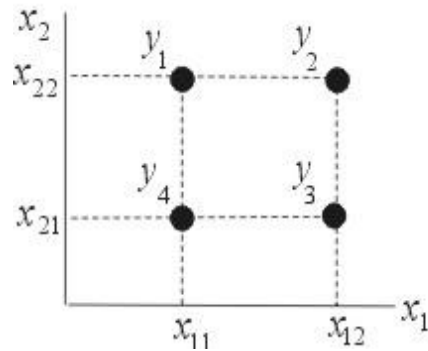


Рис. 24. Эксперимент в трех и четырех точках двухфакторного эксперимента.

Добавим третью экспериментальную точку x_{13} . Через три точки y_1, y_2 и y_3 можно провести кривую второго порядка 2, имеющую уравнение

$$y_p = a_0 + a_1 x_1 + a_2 x_1^2.$$

По-прежнему $m = N = 3$: количество экспериментальных точек и число членов уравнения совпадают, модель при этом адекватна и представляет собой полином $(N - 1)$ степени. Но если провести между тремя точками прямую линию 3 с уравнением (34), то мера неадекватности (33) станет больше нуля.

Итак, при опытах в N точках полином с уравнением:

$$y_p = a_0 + a_1 x_1 + a_2 x_1^2 + \dots + a_{N-1} x_1^{(N-1)} \quad (35)$$

точно проходит через все экспериментальные точки. Это интерполяционный полином. Мера неадекватности при этом равна нулю. Для рассмотренного однофакторного эксперимента интерполяционный полином содержит член с наивысшей степенью переменного $(N - 1)$ – полином $(N - 1)$ – й степени. Если постепенно уменьшать количество членов в этом полиноме, то мера неадекватности начнет постепенно возрастать. При каком-то количестве оставшихся членов мера неадекватности станет предельной, и дальнейшее отбрасывание членов в полиноме становится нецелесообразным.

Рассмотрим двухфакторный эксперимент $y = f(x_1, x_2)$. Если в факторном пространстве, представленном на рис.24, выбрать три узловые точки с координатами (x_{12}, x_{21}) , (x_{12}, x_{22}) и (x_{11}, x_{22}) , то через три экспериментальные точки y_1, y_2 и y_3 можно провести плоскость с уравнением

$$y_p = a_0 + a_1 x_1 + a_2 x_2.$$

Как видно, количество членов уравнения равно числу узловых точек :

$m = N = 3$. Мера неадекватности (33) равна нулю. Если добавим четвертую узловую точку с координатами (x_{11}, x_{21}) , то соответствующее ей экспериментальное значение y_4 не обязано лежать на той же плоскости. Приведенное уравнение, в котором число членов $m < N$, становится неадекватным. Чтобы получить адекватное уравнение, необходимо плоскость “закрутить” так, чтобы она захватила 4-ю точку. Эту операцию осуществляет взаимодействующий член уравнения. Уравнение

$$y_p = a_0 + a_1 x_1 + a_2 x_2 + a_{12} x_1 x_2$$

становится адекватным, оно также удовлетворяет условию $m = N$.

Можно обобщить результат и сформулировать следующую теорему о количестве членов в уравнении:

построенное по N экспериментальным узловым точкам алгебраическое уравнение (32), адекватное экспериментальным данным, содержит N членов. При удалении каких-то членов из уравнения мера неадекватности SS_{ad} возрастает с увеличением количества удаленных членов.

По терминологии дисперсионного анализа та же мысль может быть представлена следующим образом. Каждый член уравнения вносит свой определенный вклад в общую дисперсию эксперимента. Когда записаны все N членов уравнения, то общий вклад их соответствует экспериментальным данным, и мера неадекватности равна нулю. Если из уравнения исключить какие-то члены, то вместе с ними уйдет их вклад в общую дисперсию, и мера неадекватности возрастет.

Адекватное уравнение в однофакторном эксперименте является полиномом $(N - 1)$ степени, в многомерном пространстве полином содержит N членов низших степеней факторов и взаимодействий низших уровней.

5.3. Построение интерполяционного уравнения методом Ньютона.

Существует несколько способов построения интерполяционных уравнений, мы остановимся на трех, с нашей точки зрения, наиболее применяемых.

Метод Ньютона, или метод решения системы линейных уравнений, интересен тем, что пригоден как для однофакторных, так и многофакторных экспериментов. Рассмотрим многофакторную модель (32), в которой содержится m членов. Введем следующие обозначения:

$$\begin{aligned} a_{12} &= a_{k+1}, & x_1 x_2 &= x_{k+1}; \\ a_{13} &= a_{k+2}, & x_1 x_3 &= x_{k+2}; \\ & \dots & & \dots \\ a_{pp} &= a_{m-1}, & x_p &= x_{m-1}. \end{aligned}$$

Тогда m -членное уравнение (32) будет иметь вид:

$$y_p = a_0 + a_1 x_1 + a_2 x_2 + \dots + a_k x_k + a_{k+1} x_{k+1} + \dots + a_{m-1} x_{m-1} \quad (36)$$

В этом уравнении счетчик номера членов i изменяется от 0 до $(m - 1)$:

$$i = 0, 1, 2, \dots, (m - 1).$$

Для однофакторной модели (35) можно также произвести замену переменных: $x_1^2 = x_2$, $x_1^3 = x_3$ и т. д. Тогда она записывается также в виде уравнения (36). В обоих случаях $N = m$.

Для того, чтобы найти m неизвестных коэффициентов уравнения (36), необходимо провести $N = m$ опытов, по которым составляется система уравнений (каждое уравнение строится по одному опыту):

$$\begin{aligned} y_1 &= a_0 + a_1 x_{11} + a_2 x_{21} + \dots + a_{m-1} x_{m-1,1} \\ y_2 &= a_0 + a_1 x_{12} + a_2 x_{22} + \dots + a_{m-1} x_{m-1,2} \end{aligned}$$

$$\begin{aligned}
 & \dots\dots\dots \\
 & y_u = a_0 + a_1 x_{1u} + a_2 x_{2u} + \dots + a_{m-1} x_{m-1,u} \quad (37) \\
 & \dots\dots\dots \\
 & y_N = a_0 + a_1 x_{1,N} + a_2 x_{2,N} + \dots + a_{m-1} x_{m-1,N}
 \end{aligned}$$

Здесь счетчик номера опытов (второй индекс при x) $u = 1, 2, \dots, N$. Система будет иметь единственное решение тогда, когда число неповторяющихся опытов и число неизвестных совпадает. При проведении опытов очень важно проследить за тем, чтобы проведенные опыты обеспечивали независимость уравнений (невозможность получения их друг из друга.) При проведении опытов в N разных узловых точках это условие выполняется автоматически.

В результате решения системы линейных уравнений (37) с m неизвестными получаем искомые значения коэффициентов a_i .

Приведенный метод интерполяции широко применяется для написания приближенных упрощенных уравнений при обработке экспериментальных графиков. Возьмем график, представленный на рис. 25 а, который

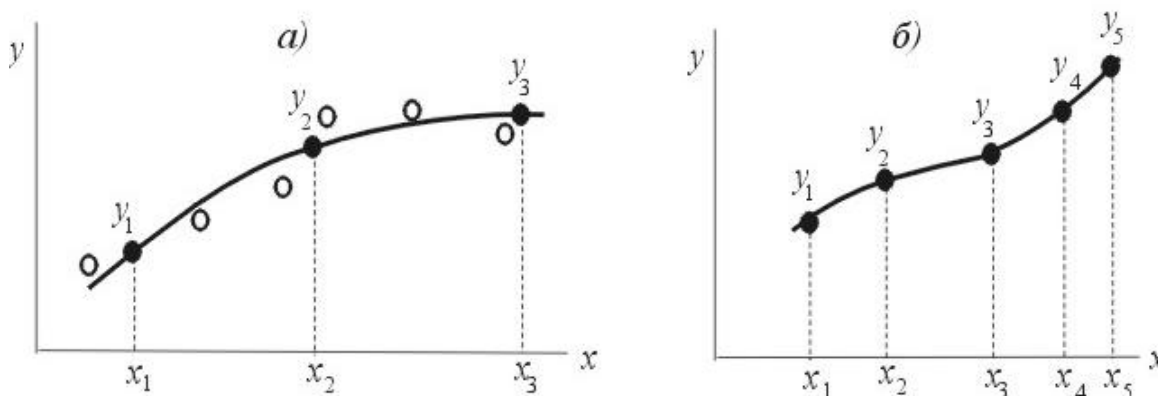


Рис.25. Построение уравнения по трем (а) и пяти (б) точкам

получен для однофакторного эксперимента. В нашем случае проведено 6 опытов, отмеченных светлыми кружочками. По ним можно было построить адекватный полином 5-й степени (с 6-ю коэффициентами уравнения (35)). Но нас вполне может устроить более простая усредненная кривая, которую мы сами провели «на глаз» между экспериментальными точками. Необходимо написать уравнение этой кривой. Вид этой кривой должен подсказать, полиномом какой степени она является. На рис. 25, а кривая визуально близка к параболе, то есть кривой второго порядка, поэтому для построения ее уравнения на кривой необходимо выбрать три точки. В общем случае на кривой должно быть выбрано $N - 1$ точек. Через эти точки обязательно пройдет искомое уравнение.

Допустим, выбрали три точки (на рис.25, а зачернены) со значениями по кривой y_1, y_2 и y_3 , по которым методом Ньютона построим полином второй степени:

$$y_1 = a_0 + a_1 x_1 + a_2 x_1^2$$

$$y_2 = a_0 + a_1 x_2 + a_2 x_2^2$$

$$y_3 = a_0 + a_1 x_3 + a_2 x_3^2$$

Полученное уравнение точно пройдет только через черные точки, но даже при небольшом представлении о кривых 2-го порядка можно достоверно утверждать, что и в остальных точках графика совпадение будет приличным.

Обычно точки, через которые мы обязуемся провести искомым полином, и количество таких точек, определяются перегибами на экспериментальной кривой. На рис. 25 б для описания кривой, видимо, потребуется 5 точек. Получим полином 4-й степени.

Упражнения

1. По экспериментальным данным, приведенным в табл.3 и 4, построить методом Ньютона полиномы, точно соответствующие средним значениям в каждой экспериментальной точке.

2. По экспериментальным данным, приведенным в табл.5 и 6, построить наиболее простые кривые, укладывающиеся в доверительный интервал, и записать для них уравнения методом Ньютона..

5. 4. Интерполяционные полиномы Лагранжа.

Метод Лагранжа применим только для одномерной модели $y = f(x_1)$, поэтому индекс «1» при x опустим. Искомым является полином $(N - 1)$ -й степени (35), если проведены опыты в N точках.

Предварительно рассмотрим следующую задачу. Пусть задан полином N -й степени $\omega(x)$, который разлагается на простые неповторяющиеся множители:

$$\omega(x) = (x - x_1)(x - x_2) \cdots (x - x_N).$$

Запишем простую дробь $f(x) / \omega(x)$ (напомним, что простой называется такая дробь, у которой наибольшая степень числителя ниже, чем степень знаменателя) и разложим ее на простые слагаемые:

$$\frac{f(x)}{\omega(x)} = \frac{A_1}{x - x_1} + \frac{A_2}{x - x_2} + \cdots + \frac{A_N}{x - x_N}. \quad (38)$$

Приведем суммы к общему знаменателю:

$$f(x) = A_1(x - x_2)(x - x_3) \cdots (x - x_N) +$$

$$+ A_2(x - x_1)(x - x_3) \cdots (x - x_N) +$$

$$\dots$$

$$+ A_N(x - x_1)(x - x_2) \cdots (x - x_{N-1}).$$

Значения функции $f(x)$ в точках x_1, x_2, \dots, x_N :

$$f(x_1) = A_1(x_1 - x_2)(x_1 - x_3) \cdots (x_1 - x_N);$$

$$f(x_2) = A_2(x_2 - x_1)(x_2 - x_3) \cdots (x_2 - x_N);$$

$$\dots$$

$$f(x_N) = A_N(x_N - x_1)(x_N - x_2) \cdots (x_N - x_{N-1}). \quad (39)$$

Вычислим производную от полинома $\omega(x)$:

$$\begin{aligned} \omega'(x) &= (x-x_2)(x-x_3)\cdots(x-x_N) + \\ &+ (x-x_1)(x-x_3)\cdots(x-x_N) + \\ &\dots\dots\dots \\ &+ (x-x_1)(x-x_2)\cdots(x-x_{N-1}). \end{aligned}$$

Значения производной $\omega'(x)$ в точках x_1, x_2, \dots, x_N :

$$\begin{aligned} \omega'(x_1) &= (x_1-x_2)(x_1-x_3)\cdots(x_1-x_N); \\ \omega'(x_2) &= (x_2-x_1)(x_2-x_3)\cdots(x_2-x_N); \\ &\dots\dots\dots \\ \omega'(x_N) &= (x_N-x_1)(x_N-x_2)\cdots(x_N-x_{N-1}). \end{aligned} \tag{40}$$

Сравнивая (39) и (40), получаем :

$$A_1 = \frac{f(x_1)}{\omega'(x_1)}; \quad A_2 = \frac{f(x_2)}{\omega'(x_2)}; \quad \dots; \quad A_N = \frac{f(x_N)}{\omega'(x_N)}.$$

Уравнение (38) приобретает вид:

$$\frac{f(x)}{\omega(x)} = \frac{f(x_1)}{\omega'(x_1) \cdot (x-x_1)} + \frac{f(x_2)}{\omega'(x_2) \cdot (x-x_2)} + \dots + \frac{f(x_N)}{\omega'(x_N) \cdot (x-x_N)}.$$

Или

$$\frac{f(x)}{\omega(x)} = \sum_{u=1}^N \frac{f(x_u)}{\omega'(x_u) \cdot (x-x_u)}.$$

Используя полученную формулу, запишем интерполяционный полином Лагранжа с учетом того, что в узловых точках $x = x_u$ функция приобретает экспериментальные значения $f(x_u) = y_u$:

$$\dots \tag{41}$$

Интерполяционный полином Лагранжа степени m обозначают $L_m(x)$.

Пример. Имеем таблицу 7, включающую 8 измерений (4 опыта) в 4 узлах. Интерполяционный полином Лагранжа имеет третью степень: $L_3(x)$.

Таблица 7

<i>u</i>	1	2	3	4
<i>x</i>	1	5	4	2
<i>y</i>	3	2	1	6

$$L_3(x) = 3 \cdot \frac{(x-5) \cdot (x-4) \cdot (x-2)}{(1-5) \cdot (1-4) \cdot (1-2)} + 2 \cdot \frac{(x-1) \cdot (x-4) \cdot (x-2)}{(5-1) \cdot (5-4) \cdot (5-2)} +$$

$$+1 \cdot \frac{(x-1) \cdot (x-5) \cdot (x-2)}{(4-1) \cdot (4-5) \cdot (4-2)} + 6 \cdot \frac{(x-1) \cdot (x-5) \cdot (x-4)}{(2-1) \cdot (2-5) \cdot (2-4)}.$$

Здесь в каждой из дробей в числителе и знаменателе по формуле (41) должно присутствовать еще по одному множителю, но они сокращаются, поэтому не записаны. После приведения подобных членов получается требуемый интерполяционный полином.

Интерполяционный полином Лагранжа, как и полиномы, получаемые некоторыми другими методами, имеют два недостатка. Во-первых, практически невозможно построить такой полином для многофакторного эксперимента, даже для двух факторов это сделать весьма трудно. Во-вторых, такие полиномы “осциллируют” и “раскачиваются”, то есть в промежутках между узловыми точками ведут себя произвольно и могут далеко отходить от каких-то плавных линий. Неоднозначность поведения полинома между узлами во многих задачах неприемлема. Такого недостатка лишен следующий метод.

5. 5. Кубические сплайны.

Кубический сплайн – это функция, составленная из отдельных кусочков (функций). Это не отдельный полином $N - 1$ степени, а кусочная функция, равнозначная искомому полиному. Такая функция – сплайн – также предназначена для описания однофакторного эксперимента и также обязана в узловых точках совпадать с экспериментальными значениями y_u , но, в отличие от полинома Лагранжа, нормирована в пространстве между узлами. Сплайн обеспечивает минимальную суммарную кривизну линии, проходящей через экспериментальные точки. Если тонкую стальную линейку изогнуть так, чтобы она прошла через точки y_u , то это и будет сплайн, так как энергия изогнутой линейки будет минимальной при минимальной суммарной кривизне ее.

Кривизну функции $y = f(x)$ определяет скорость изменения угла α между касательной к кривой и осью x при движении вдоль кривой, то есть вторая производная $f''(x)$ должна быть минимальной, часто равной нулю.

Составим функцию $f(x)$ из кусочков $\varphi_i(x)$. Нумерацию узловых точек проведем, начиная с нуля: $i = 0, 1, \dots, n$. Тогда счетчик i , приведенный на рис 26, а и б, соответствует номеру участка в границах между двумя соседними узлами с номерами $i - 1$ и i . Логично потребовать, чтобы вторые производные этих кусочков изменялись прямолинейно от узла к узлу и принимали равные значения в самих узлах, как показано на рис. 26 а. Первые производные и сами кусочные функции $\varphi_i(x)$ в узлах также должны быть равны. В такой постановке вторые производные кусочных функций представляют собой полиномы первого порядка, их первые производные будут кривыми второго, а сами кусочные функции (рис.26 б) – кривыми третьего порядка. В связи с этим описываемый сплайн называется кубическим.

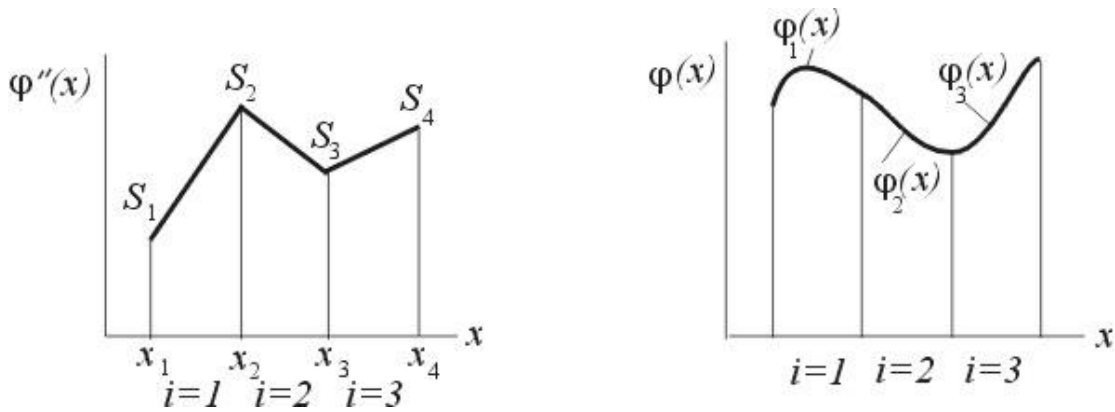


Рис. 26. Вторая производная (а) и кусочная функция (б) – сплайн.

На i -м промежутке будем иметь:

$$\begin{aligned}\varphi_{i-1}(x_i) &= \varphi_i(x_i), \\ \varphi_{i-1}'(x_i) &= \varphi_i'(x_i), \\ \varphi_{i-1}''(x_i) &= \varphi_i''(x_i).\end{aligned}$$

Пусть протяженность i -го промежутка равна $h_i = x_i - x_{i-1}$: $x_{i-1} \leq x \leq x_i$. В узловых точках промежутка нам известны значения функций $\varphi_i(x_{i-1}) = y_{i-1}$ и $\varphi_i(x_i) = y_i$. Обозначим в этих точках вторые производные (пока неизвестные) соответственно через $\varphi_i''(x_{i-1}) = S_{i-1}$ и $\varphi_i''(x_i) = S_i$. Используя интерполяционный полином Лагранжа, построим уравнение прямой для вторых производных:

$$\begin{aligned}\varphi_i''(x) = L_1(x) &= \sum_{u=i-1}^i S_u \cdot \frac{\omega(u)}{\omega'(x_u) \cdot (x - x_u)} = \\ \frac{S_{i-1}(x - x_i)}{x_{i-1} - x_i} + \frac{S_i(x - x_{i-1})}{x_i - x_{i-1}} &= \frac{1}{h_i} \cdot [S_{i-1}(x_i - x) + S_i(x - x_{i-1})].\end{aligned}$$

Проинтегрируем дважды:

$$\varphi_i'(x) = \frac{-S_{i-1}(x_i - x)^2}{2h_i} + \frac{S_i(x - x_{i-1})^2}{2h_i} + C_1,$$

$$\varphi_i(x) = \frac{S_{i-1}(x_i - x)^3}{6h_i} + \frac{S_i(x - x_{i-1})^3}{6h_i} + C_1(x - x_i) + C_2.$$

Константы интегрирования C_1 и C_2 находим из граничных значений функций на концах участка:

$$\varphi_i(x_{i-1}) = y_{i-1} = \frac{S_{i-1}h_i^2}{6} - C_1h_i + C_2,$$

$$\varphi_i(x_i) = y_i = \frac{S_ih_i^2}{6} + C_2.$$

$$\text{Отсюда } C_2 = y_i - \frac{S_i h_i^2}{6} \quad \text{и} \quad C_1 = \frac{1}{h_i} \cdot \left(\frac{S_{i-1} h_i^2}{6} - y_{i-1} + y_i - \frac{S_i h_i}{6} \right).$$

Подставим константы в исходное уравнение:

$$\begin{aligned} \varphi_i(x) = & \frac{S_{i-1}(x_i - x)^3}{6h_i} + \frac{S_i(x - x_{i-1})^3}{6h_i} + \frac{S_{i-1}h_i}{6}(x - x_i) - \\ & - \frac{y_{i-1}}{h_i}(x - x_i) + \frac{y_i}{h_i}(x - x_i) + y_i - \frac{S_i h_i}{6}(x - x_i) - \frac{S_i h_i^2}{6}. \end{aligned}$$

Так как

$$\frac{y_i}{h_i}(x - x_i) + y_i = \frac{y_i}{h_i} [(x - x_i) + (x_i - x_{i-1})] = \frac{y_i}{h_i}(x - x_{i-1})$$

и по аналогии

$$\frac{S_i h_i}{6}(x - x_i) + \frac{S_i h_i^2}{6} = \frac{S_i h_i}{6}(x - x_{i-1}),$$

то функция и производная соответственно равны:

$$\begin{aligned} \varphi_i(x) = & \frac{S_{i-1}(x_i - x)^3}{6h_i} + \frac{S_i(x - x_{i-1})^2}{6h_i} + \left(\frac{y_{i-1}}{h_i} - \frac{S_{i-1}h_i}{6} \right) (x_i - x) + \\ & \boxed{\phantom{\frac{S_{i-1}(x_i - x)^3}{6h_i} + \frac{S_i(x - x_{i-1})^2}{6h_i} + \left(\frac{y_{i-1}}{h_i} - \frac{S_{i-1}h_i}{6} \right) (x_i - x) +}}, \end{aligned} \quad (42)$$

$$\varphi_i'(x) = \frac{-S_{i-1}(x_i - x)^2}{2h_i} + \frac{S_i(x - x_{i-1})^2}{2h_i} + \frac{S_{i-1}h_i}{6} - \frac{y_{i-1}}{h_i} + \frac{y_i}{h_i} - \frac{S_i h_i}{6}.$$

Значения вторых производных S_{i-1} и S_i находим из условия равенства первых производных в узловых точках x_i на границе двух соседних участков $h_i = x_i - x_{i-1}$ и $h_{i+1} = x_{i+1} - x_i$. В точке x_i слева

$$\varphi_i'(x_i) = \frac{y_i - y_{i-1}}{h_i} + \frac{S_i h_i}{3} + \frac{S_{i-1} h_i}{6}$$

и в этой же точке справа

$$\varphi_i'(x_i) = \frac{y_{i+1} - y_i}{h_{i+1}} - \frac{S_i h_{i+1}}{3} - \frac{S_{i+1} h_{i+1}}{6}.$$

Приравнявая значения производных слева и справа, получаем окончательное уравнение:

$$\frac{y_i - y_{i-1}}{h_i} + \frac{S_i h_i}{3} + \frac{S_{i-1} h_i}{6} = \frac{y_{i+1} - y_i}{h_{i+1}} - \frac{S_i h_{i+1}}{3} - \frac{S_{i+1} h_{i+1}}{6}. \quad (43)$$

Количество таких уравнений равно $(n - 2)$ по количеству стыков, имеющих промежутки слева и справа. Для первой и последней узловой точек (с номерами $i = 1$ и $i = n$) нет парных участков. Таким образом, для нахождения n неизвестных S_i имеем только $(n - 2)$ уравнений. Два недостающих уравнения необходимо получить из граничных условий на этих концевых участках. Они должны быть либо заданы, либо определены логикой задачи. По

аналогии с изогнутой линейкой, должны быть заданы условия заделки ее концов. Если в концевых точках линейка никак не закреплена, то ее кривизна на этих участках должна быть равна нулю, потому что иное значение кривизны потребует дополнительную энергию изгиба. В этом случае получаем два недостающих уравнения: $S_0 = S_n = 0$.

Пример. Задана таблица с 4 экспериментальными данными, по которым требуется построить кубический сплайн (табл.8).

Таблица 8

i	0	1	2	3	4
x	1	2	3	4	5
y	2	4	3	5	6
S	0	-6,21	6,86	-3,21	0

Имеем $n = 4$, то есть 5 узловых точек и 4 интервала ($i = 0, 1, 2, 3, 4$), протяженность всех интервалов одинакова и равна $h_i = 1$.

Положим вторые производные на концевых участках не заданными и потому $S_0 = S_n = 0$.

Для нахождения трех неизвестных значений вторых производных имеем три уравнения

$$\begin{aligned} (4 - 2) + S_1/3 &= (3 - 4) - S_1/3 - S_2/6 \\ (3 - 4) + S_2/3 + S_1/6 &= (5 - 3) - S_2/3 - S_3/6 \\ (5 - 3) + S_3/3 + S_2/6 &= (6 - 5) - S_3/3 \end{aligned}$$

Решение системы уравнений дает искомые значения, приведенные в последней строке табл. 13. Подставив их в (42), получим четыре кубических полинома для каждого из участков, которые вместе образуют сплайн.

Первый из них имеет вид:

$$\varphi_1(x) = [-6,21(x-1)^3]/6 - 2(2-x) + (4 - 6,21/6)(x-1) .$$

В качестве упражнения студентам полезно построить по приведенным данным интерполяционные полиномы остальными двумя методами и сравнить все три кривые между собой.

Контрольные вопросы по главе 5.

1. Чем отличается интерполяционный полином от регрессии?
2. Что такое мера неадекватности и как она изменяется при изменении числа членов регрессии?
3. Как устроен метод Ньютона для построения интерполяционного полинома?
4. Выведите формулу Лагранжа для интерполяционного полинома.
5. Каковы свойства сплайна и чем он отличается от интерполяционного полинома?
6. Придумайте сами таблицу экспериментальных данных, состоящую из пяти опытов и составьте для нее интерполяционные полиномы методами Ньютона и Лагранжа. Для той же таблицы составьте сплайн.

6. Дисперсионный анализ

6.1. Задачи дисперсионного анализа

Дисперсионный анализ предназначен для качественного исследования модели процесса (1), которая представляет собой общую связь между входными и выходными параметрами:

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

Причем, как отмечалось, для теории дисперсионного анализа безразлично, в какой форме записаны входные и выходные параметры - в виде размерных, безразмерных параметров или критериев, хотя для практического проведения эксперимента желательно иметь небольшое число высокоинформативных факторов, добытых методами теории подобия. Среди входных параметров (факторов) есть сильные, изменение которых сильно отражается на значении величины y , есть послабее и совсем слабые. Последние желательно исключить из эксперимента, чтобы упростить и удешевить его. В производственных и экономических задачах выделение группы сильных факторов определяет стратегию воздействия на производство с тем, чтобы улучшить характеристики продукции, улучшить технико-экономические показатели и т. п. По теореме Парето [1], в таких задачах сильных факторов всегда мало. Например, при исследовании причин брака какой-то детали удается всегда выделить не более трех-четырех факторов, изменение которых решает проблемы.

С другой стороны, согласно теореме о существенных параметрах (см. 2. 2) в уравнение (1) необходимо включить все существенные параметры, пропуск хотя бы одного из них ведет к неверному и даже противоположному выводу. В связи с этим в начальный период исследования или при изучении сложных процессов, когда еще не ясно, какие факторы относятся к сильным, логично как можно большее количество параметров включить в уравнение (1), а затем ранжировать их по степени воздействия на y , то есть на первое место поставить самый сильный, затем послабее и т. д. Если по вертикальной оси отложить “силу” (значимость) факторов, а по горизонтальной - номер фактора по рангу, то получим кривую, похожую на представленную на рис. 27. В теории эксперимента показано, что характер этой кривой близок к экспоненте, то есть значимость факторов быстро ослабевает, особенно в технико-экономических задачах, и достойных исследования факторов сравнительно мало. В конце ряда слабые факторы, значимость которых несущественна, обязательно исключаются из эксперимента.

Математический аппарат, который занимается исследованием значимости входных параметров, называется дисперсионным анализом, так как в основе его лежит анализ вкладов каждого фактора в общую дисперсию эксперимента; тот фактор сильнее, который вносит больше в общую “копилку”.

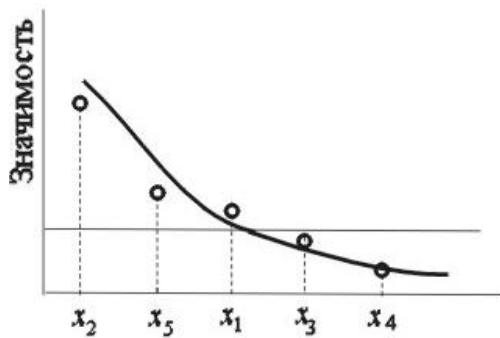


Рис. 27. Характер изменения значимости входных параметров.

Рассмотрим однофакторный эксперимент: $y = f(x_1)$. Дисперсионную модель этого эксперимента можно представить в виде:

$$y = A + \varepsilon, \quad (44)$$

где y - общий вклад в общую дисперсию, который вносят все факторы или, как говорят, общий эффект модели. Если угодно, его можно принять за 100 %. Соответственно, здесь A - эффект фактора x_1 , ε - эффект ошибки воспроизводимости.

На последнем следует остановиться подробнее. На рис. 28 приведено графическое представление рассматриваемого эксперимента. Если при каждом значении фактора x_1 - то есть в точках $x_{11}, x_{12}, \dots, x_{1a}$ (количество точек равно a и счетчик точек $i = 1, 2, \dots, a$) - проведено по одному измерению y - получено, соответственно, y_1, y_2, \dots, y_a -, то ни о какой воспроизводимости опытов говорить не приходится, они не воспроизводятся, в этом случае $\varepsilon = 0$. Но если хотя бы в одной точке эксперимента, например, при x_{11} ,

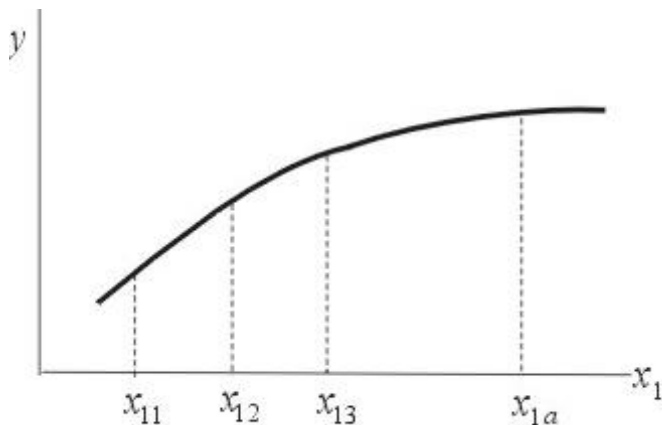


Рис. 28. Анализ однофакторного эксперимента.

проведено n измерений ($n > 1$), то y_1 как случайная величина будет отклоняться от среднего значения \bar{y}_1 в пределах доверительного интервала $\pm t_p \cdot S$, принимая значения $y_{11}, y_{12}, \dots, y_{1n}$. (Можно ввести счетчик q числа опытов, проведенных в одной точке эксперимента, $q = 1, 2, \dots, n$). Дисперсионный анализ молчаливо предполагает, что такие отклонения в точке x_{11} , а также во всех других точках эксперимента, подчиняются нормальному закону распределения. Дисперсию воспроизводимости в одной точке x_{11} эксперимента мы можем вычислить, например, по формуле (28). При числе степеней свободы

$f = (n - 1)$ сумма квадратов отклонений

$$SS = \sum_{q=1}^n y_{1q}^2 - \frac{1}{n} \cdot \left(\sum_{q=1}^n y_{1q} \right)^2.$$

(Последнюю сумму в скобках можно представлять в виде $y_{1\bullet}$).

Если во всех a точках проведено по n опытов, то все отклонения от средних \bar{y}_i можно объединить в единую выборку только в том случае, когда их дисперсии во всех точках будут статистически равнозначными (по критерию Фишера или Кохрана). Однако, дисперсионный анализ предполагает, что условие статистической равнозначности дисперсий выполняется автоматически без доказательств.

Итак, при n опытах в a точках общее число степеней свободы ошибки воспроизводимости $f = a(n - 1)$, а сумма квадратов отклонений ошибки воспроизводимости

$$SS = \sum_{i=1}^a \sum_{q=1}^n y_{iq}^2 - \sum_{i=1}^a (y_{i\bullet})^2. \quad (45)$$

Представленные суммы квадратов отклонений и рассчитанные дисперсии по формуле (26) вносят определенный вклад в общую дисперсию эксперимента, поэтому эффект ошибки воспроизводимости ε в модели (44) не равен нулю.

Аналогично, дисперсионная модель двухфакторного эксперимента $y = f(x_1, x_2)$ будет иметь вид

$$y = A + B + AB + \varepsilon, \quad (46)$$

где A и B – эффекты факторов x_1 и x_2 , а AB – эффект их совместного влияния, взаимодействия. По-прежнему, ε – эффект ошибки воспроизводимости, не равный нулю только в том случае, когда хотя бы в одной точке эксперимента было проведено больше одного опыта.

Дисперсионная модель трехфакторного эксперимента, видимо, будет содержать не только эффекты парных, но и тройного взаимодействия:

$$y = A + B + C + AB + AC + BC + ABC + \varepsilon. \quad (47)$$

6.2. Основные уравнения дисперсионного анализа.

Основные уравнения дисперсионного анализа выведем для двухфакторного эксперимента, а затем распространим результат на другие эксперименты.

На рис. 29 изображена схема проведения двухфакторного эксперимента. По одной оси откладываются значения фактора x_1 , по второй – фактора x_2 , а значения y – перпендикулярно обеим координатам.

Пусть фактор x_1 изменяется на a равноотстоящих уровнях, i – счетчик уровней, $i = 1, 2, \dots, a$, фактор x_2 – на b уровнях, также следующие через равные промежутки ($j = 1, 2, \dots, b$). В каждой точке эксперимента, находящейся в узле образовавшейся сетки, пусть проводится по n опытов ($k = 1, 2, \dots, n$). Таким образом, эксперимент содержит ab узловых точек и $N = abn$ опытов. Эксперимент, который проводится во всех узловых точках, называется пол-

ным факторным (ПФЭ), если в каких-то точках опыты не проводятся, то эксперимент называется *дробным факторным* (ДФЭ).

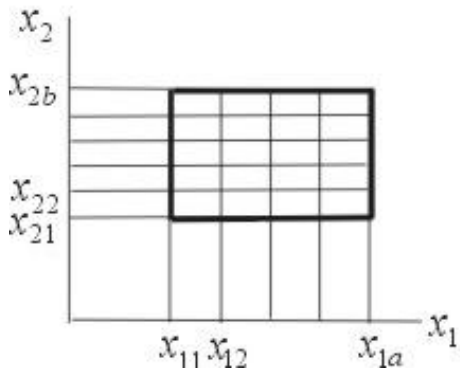


Рис. 29. Схема проведения двухфакторного эксперимента.

Следует заметить, что дисперсионный анализ чувствителен к порядку выбора опытов в эксперименте. Рекомендуется проводить эксперимент в *рандомизированном*, то есть случайном, порядке. Допустим, $a = b = n = 6$. Возьмем игральный кубик и подбросим его. Та грань, которая выпадет, определит номер i уровня первого фактора. Точно так же вторым подбрасыванием выберем номер j уровня второго фактора, а затем с третьего подбрасывания и номер опыта k в точке с координатами ij . Это и будет опыт, выбранный в рандомизированном порядке. Для следующего опыта все повторяется сначала. Никто, конечно, так эксперимент проводить не будет. Гораздо удобнее и дешевле сначала выбрать точку с координатами $i = j = 1$ и провести в ней все n опытов, затем в точке с координатами $i = 1$ и $j = 2$ – следующие n опытов и т. д. В эксперимент по всем параметрам вносится порядок. Например, требуется определить зависимость какого-то параметра качества металла в зависимости от температуры нагрева и времени выдержки металла в печи. Никто не будет подбрасывать игральный кубик, а назначит одну из температур, и в нагретую печь загрузит все образцы, предназначенные для испытания при этой температуре. Затем по прошествии требуемого времени выдержки будет доставать образцы. Тем самым вносится порядок в эксперимент, нарушая требования рандомизации. Это чревато тем, что в процессе проведения эксперимента обнаружится некий существенный фактор x_3 (в нашем случае, температура холодного спая термопары), который зависит от порядка проведения опытов. При случайном порядке проведения опытов этот фактор превратился бы в случайный и отразился бы в общей ошибке, но при упорядочении опытов этот фактор даст тренд. Как известно, пропуск существенного фактора часто ведет к неверному результату. Это всегда следует иметь в виду.

Во многих случаях, проводя упорядоченные эксперименты, исследователи вводят в эксперимент сознательно дополнительные факторы – факторы порядка, которые затем либо отсеиваются как незначимые, либо оказываются значимыми, они не будут пропущенными. Вскроется и причина появления новых существенных факторов.

Вернемся к нашему двухфакторному ПФЭ с количеством опытов равным $N = abn$. Результаты эксперимента представлены в табл. 9. Наблюдаемое

Таблица 9

	$j = 1$	$j = 2$...	$j = b$	Суммы $y_{i \bullet n}$
$i = 1$	y_{111}	y_{121}		y_{1b1}	$y_{1 \bullet 1}$
	y_{112}	y_{122}		y_{1b2}	$y_{1 \bullet 2}$

	y_{11n}	y_{12n}		y_{1bn}	$y_{1 \bullet n}$
Суммы $y_{1j \bullet}$	$y_{11 \bullet}$	$y_{12 \bullet}$		$y_{1b \bullet}$	Сумма $y_{1 \bullet \bullet}$
$i = 2$	y_{211}	y_{221}		y_{2b1}	$y_{2 \bullet 1}$
	y_{212}	y_{222}		y_{2b2}	$y_{2 \bullet 2}$

	y_{21n}	y_{22n}		y_{2bn}	$y_{2 \bullet n}$
Суммы $y_{2j \bullet}$	$y_{21 \bullet}$	$y_{22 \bullet}$		$y_{2b \bullet}$	Сумма $y_{2 \bullet \bullet}$
...
$i = a$	y_{a11}	y_{a21}		y_{ab1}	$y_{a \bullet 1}$
	y_{a12}	y_{a22}		y_{ab2}	$y_{a \bullet 2}$

	y_{a1n}	y_{a2n}		y_{abn}	$y_{a \bullet n}$
Суммы $y_{aj \bullet}$	$y_{a1 \bullet}$	$y_{a2 \bullet}$		$y_{ab \bullet}$	Сумма $y_{a \bullet \bullet}$
Суммы $y_{\bullet j \bullet}$	$y_{\bullet 1 \bullet}$	$y_{\bullet 2 \bullet}$		$y_{\bullet b \bullet}$	Сумма $y_{\bullet \bullet \bullet}$

значение y снабжены символами (y_{ijk}) , которые определяют уровни факторов и номер опыта в узловой точке с координатами ij . В таблице приведены также некоторые суммы, в которых знак суммирования заменен точкой.

Видно, что

$$\sum_k^n y_{ijk} = y_{ij \bullet} \quad ; \quad \sum_i^a y_{ij \bullet} = \sum_i^a \sum_k^n y_{ijk} = y_{\bullet j \bullet} \quad ; \quad \sum_j^b y_{ij \bullet} = \sum_j^b \sum_k^n y_{ijk} = y_{i \bullet \bullet} \quad ;$$

$$\sum_i^a y_{i \bullet \bullet} = \sum_j^b y_{\bullet j \bullet} = \sum_i^a \sum_j^b y_{ij \bullet} = \sum_i^a \sum_j^b \sum_k^n y_{ijk} = y_{\bullet \bullet \bullet}$$

Для расчетов будут необходимыми также средние значения, определенные как сумма, деленная на количество опытов, образующее эту сумму:

$$\begin{aligned} \bar{y}_{ij \bullet} &= y_{ij \bullet} / n \quad ; \quad \bar{y}_{i \bullet \bullet} = y_{i \bullet \bullet} / bn \quad ; \\ \bar{y}_{\bullet j \bullet} &= y_{\bullet j \bullet} / an \quad ; \quad \bar{y}_{\bullet \bullet \bullet} = y_{\bullet \bullet \bullet} / abn \quad . \end{aligned} \quad (48)$$

Общее число степеней свободы $f_{\text{общ}}$ эксперимента на 1 меньше общего числа опытов:

$$f_{\text{общ}} = abn - 1.$$

Число степеней свободы каждого из факторов на 1 меньше числа уровней его изменения:

$$f_{x1} = a - 1 \quad ; \quad f_{x2} = b - 1 .$$

(Далее для экономии места будем писать f_1 и f_2 вместо f_{x_1} и f_{x_2} .)

Число степеней свободы взаимодействия равно произведению чисел степеней свободы взаимодействующих факторов:

$$f_{12} = (a - 1)(b - 1).$$

Число степеней свободы ошибки воспроизводимости по ab точкам, как известно:

$$f_o = ab(n - 1).$$

Напишем тождество, в чем можно убедиться, раскрыв скобки:

$$abn - 1 = (a - 1) + (b - 1) + (a - 1)(b - 1) + ab(n - 1).$$

Отсюда получаем первое основное уравнение дисперсионного анализа:

$$f_{\text{общ}} = f_1 + f_2 + f_{12} + f_o. \quad (49)$$

Общее число степеней свободы состоит из суммы чисел степеней свободы всех эффектов дисперсионной модели (46).

По аналогии, для трехфакторной модели (47) можно записать:

$$f_{\text{общ}} = f_1 + f_2 + f_3 + f_{12} + f_{13} + f_{23} + f_{123} + f_o.$$

Здесь число уровней фактора x_3 равно c (счетчик $s = 1, 2, \dots, c$).

Недостающие числа степеней свободы равны:

$$f_3 = c - 1, \quad f_{123} = (a - 1)(b - 1)(c - 1); \quad f_o = abc(n - 1).$$

Рассмотрим суммы квадратов отклонений двухфакторного эксперимента. Общая сумма, очевидно, складывается из квадратов отклонений каждого измерения y_{ijk} от общего среднего \bar{y}_{\dots} :

$$SS_{\text{общ}} = \sum_i^a \sum_j^b \sum_k^n (y_{ijk} - \bar{y}_{\dots})^2.$$

Запишем тождество, в чем можно убедиться, раскрыв круглые скобки:

$$SS_{\text{общ}} = \left[\sum_i^a \sum_j^b \sum_k^n (y_{ijk} - \bar{y}_{i..})^2 + \sum_i^a \sum_j^b \sum_k^n (y_{ijk} - \bar{y}_{.j.})^2 + \sum_i^a \sum_j^b \sum_k^n (y_{ijk} - \bar{y}_{i..})^2 + \sum_i^a \sum_j^b \sum_k^n (y_{ijk} - \bar{y}_{.j.})^2 + \sum_i^a \sum_j^b \sum_k^n (y_{ijk} - \bar{y}_{i..})^2 + \sum_i^a \sum_j^b \sum_k^n (y_{ijk} - \bar{y}_{.j.})^2 \right].$$

Если раскрыть квадратные скобки, то получим четыре квадратичных члена плюс шесть удвоенных произведения. Доказано, что шесть удвоенных произведения попарно уничтожают друг друга, то есть в сумме равны нулю. (Можно поупражняться и доказать это самому). В результате получим:

$$SS_{\text{общ}} = \sum_i^a \sum_j^b \sum_k^n \left[(\bar{y}_{i..} - \bar{y}_{\dots})^2 + (\bar{y}_{.j.} - \bar{y}_{\dots})^2 + (\bar{y}_{ij.} - y_{i..} - \bar{y}_{.j.} + \bar{y}_{\dots})^2 + (y_{ijk} - \bar{y}_{ij.})^2 \right].$$

Каждое из оставшихся четырех слагаемых можно трактовать как сумму квадратов отклонений соответствующего эффекта модели (46):

$$SS_{\text{общ}} = SS_1 + SS_2 + SS_{12} + SS_o. \quad (50)$$

Получено второе основное уравнение дисперсионного анализа:

общая сумма квадратов отклонений состоит из суммы сумм квадратов отклонений всех эффектов дисперсионной модели (46).

По аналогии для трехфакторного эксперимента:

$$SS_{\text{общ}} = SS_1 + SS_2 + SS_3 + SS_{12} + SS_{13} + SS_{23} + SS_{123} + SS_o. \quad (51)$$

Преобразуем суммы квадратов отклонений двухфакторного эксперимента к виду, удобному для вычислений. Для первой суммы получим:

$$SS_1 = \sum_i^a \sum_j^b \sum_k^n (\bar{y}_{i..} - \bar{y}_{...})^2 = bn \cdot \frac{\sum_i^a y_{i..}^2}{b^2 n^2} + abn \cdot \frac{y_{...}^2}{a^2 b^2 n^2} - 2 \cdot \sum_i^a \sum_j^b \sum_k^n \bar{y}_{i..} \cdot \bar{y}_{...} .$$

Последнее слагаемое преобразуем отдельно, вынеся за скобки число $\bar{y}_{...}$:

$$\frac{y_{...}}{abn} \cdot \sum_i^a \sum_j^b \sum_k^n \frac{y_{i..}}{bn} = \frac{y_{...}}{abn} \cdot \frac{bn}{bn} \cdot \sum_i^a y_{i..} = \frac{y_{...}^2}{abn} .$$

Соберем все слагаемые этой суммы и по аналогии преобразовать остальные суммы квадратов отклонений. Получим формулы, приведенные в табл. 10.

Таблица 10

Эффект модели	Число степеней свободы f	Сумма квадратов отклонений SS
A (фактор x_1)	$a - 1$	$\sum_i^a \frac{y_{i..}^2}{bn} - \frac{y_{...}^2}{abn}$
B (фактор x_2)	$b - 1$	$\sum_j^b \frac{y_{.j.}^2}{an} - \frac{y_{...}^2}{abn}$
AB (взаимодействие $x_1 x_2$)	$(a - 1)(b - 1) = ab - a - b + 1$	$\sum_i^a \sum_j^b \frac{y_{ij.}^2}{n} - \sum_i^a \frac{y_{i..}^2}{bn} - \sum_j^b \frac{y_{.j.}^2}{an} + \frac{y_{...}^2}{abn}$
Ошибка воспроизводимости ε	$ab(n - 1) = abn - ab$	$\sum_i^a \sum_j^b \sum_k^n y_{ijk}^2 - \sum_i^a \sum_j^b \frac{y_{ij.}^2}{n}$
Общий эффект	$abn - 1$	$\sum_i^a \sum_j^b \sum_k^n y_{ijk}^2 - \frac{y_{...}^2}{abn}$

Можно проследить некоторую аналогию в записи выражения для числа степеней свободы после раскрытия скобок и выражения для суммы квадратов отклонений:

- 1) количество слагаемых и знаки перед ними в выражениях для числа степеней свободы и соответствующей суммы квадратов отклонений совпадают;
- 2) в каждом слагаемом для SS знаки Σ содержат индексы, аналогичные индексам при f ;

- 3) эти же индексы присутствуют в числителе при y^2 , а недостающие индексы числителя заменены точками;
- 4) в знаменателе дроби оказывается число опытов, образующих одну соответствующую сумму числителя; для ПФЭ знаменатель можно записать по недостающим индексам числителя, а при проведении ДФЭ знаменатель соответственно уменьшается

Эту аналогию используем в качестве правила для формального написания суммы квадратов отклонений. Для этого сначала необходимо написать выражение для числа степеней свободы и раскрыть в нем скобки. Затем, придерживаясь п.1 – 4, написать соответствующие члены искомым сумм.

Например, для тройного взаимодействия трехфакторного эксперимента будем иметь число степеней свободы:

$$f_{123} = (a - 1)(b - 1)(c - 1) = abc - ab - ac - bc + a + b + c - 1.$$

По приведенному правилу чисто формально записываем все 8 членов суммы квадратов отклонений:

$$\begin{aligned}
 SS_{123} = & \sum_i^a \sum_j^b \sum_s^c \frac{y_{ijs\bullet}^2}{n} - \sum_i^a \sum_j^b \frac{y_{ij\bullet\bullet}^2}{cn} - \sum_i^a \sum_s^c \frac{y_{i\bullet j\bullet}^2}{bn} - \sum_j^b \sum_s^c \frac{y_{\bullet js\bullet}^2}{an} + \\
 & + \sum_i^a \frac{y_{i\bullet\bullet\bullet}^2}{bcn} + \sum_j^b \frac{y_{\bullet j\bullet\bullet}^2}{acn} + \sum_s^c \frac{y_{\bullet\bullet s\bullet}^2}{abn} - \frac{y_{\bullet\bullet\bullet\bullet}^2}{abcn}.
 \end{aligned} \tag{52}$$

6.3. Общий анализ дисперсионной модели.

Итак, каждый из эффектов модели вносит свой вклад в общую сумму квадратов отклонений при соответствующем числе степеней свободы. Вклад ошибки воспроизводимости, вносимый разбросом данных только за счет случайных факторов, по определению должен быть небольшим (незначимым). Точнее, этот разброс органически присущ самому процессу измерения, это естественный “шум”, на фоне которого могут “прослушиваться” другие эффекты. Таким образом, числитель критерия Фишера, по которому сравниваются две дисперсии, должен содержать дисперсию оцениваемого эффекта, а знаменатель – дисперсию ошибки воспроизводимости. Если критерий Фишера некоторого i -того эффекта модели меньше табличного значения

$$F_i = \frac{SS_i / f_i}{SS_o / f_o} \leq F_T(p, f_i, f_o), \tag{53}$$

то этот эффект следует признать незначимым (он будет находиться на уровне “шума”). При обратном неравенстве вклад данного эффекта значим, так как различим на фоне ошибок воспроизводимости. И, конечно, тот эффект, который покажет наибольшее значение F -отношения, является наиболее значимым. На этом построен весь дисперсионный анализ, в результате которого все эффекты модели расставляются по рангу.

Необходимо еще раз подчеркнуть, что дисперсионный анализ возможен только тогда, когда вычисляется ошибка воспроизводимости эксперимента, а последнюю можно определить только в том случае, когда хотя бы в одной точке эксперимента проведено больше одного опыта. Чем больше число степеней свободы ошибки воспроизводимости, тем с большей достоверностью делаются выводы о значимости эффектов.

Существует теорема, доказывающая, что в дисперсионной модели все тройные и более высокие взаимодействия всегда незначимы, а значимость парных взаимодействий при увеличении количества факторов снижается, и где-то при 9-10 факторах их также можно считать мало значимыми. Таким образом, для четырехфакторного эксперимента дисперсионную модель можно записать в виде:

$$y = A + B + C + D + AB + AC + AD + BC + BD + CD + OT + \varepsilon,$$

а для девятифакторного – справедлива линейная модель:

$$y = A_1 + A_2 + \dots + A_9 + OT + \varepsilon,$$

в которых член OT представляет сумму эффектов отброшенных членов. В первой модели отброшены четыре тройных и одно четверное взаимодействие, а во второй - все парные, тройные и т. д. до 9 -го порядка взаимодействия.

В общем случае любую модель можно представить следующим образом:

$$y = OCT + OT + \varepsilon,$$

где OCT - сумма оставленных в модели членов.

В последней модели мы легко можем вычислить число степеней свободы и суммы квадратов отклонений каждого из оставленных членов, пользуясь формальными правилами предыдущего параграфа, поэтому для отброшенных членов имеем:

$$f_{от} = f_{общ} - f_{ост} - f_0 \quad \text{и} \\ SS_{от} = SS_{общ} - SS_{ост} - SS_0.$$

Значимость отброшенных членов можно проверить по критерию Фишера. Если

$$F_{от} = \frac{SS_{от} / f_{от}}{SS_0 / f_0} \leq F_T(p; f_{от}; f_0),$$

то все отброшенные эффекты суммарно незначимы. По приведенной выше теореме это неравенство должно выполняться автоматически, однако нелишнее это проверить. Кроме того, в группу отброшенных мы можем перенести незначимые линейные и парные эффекты, отсеянные по критерию Фишера (53), и, может быть, эффекты, значимость которых по условию задачи нас не интересует, при этом суммарная значимость всех отброшенных эффектов может оказаться существенной. Технолог должен решить, как поступить в том случае, когда суммарная значимость малозначимых эффектов существенна. Вполне возможно, что он поднимет “планку” шумового уровня и все-таки зачислит все указанные эффекты в число незначимых.

6. 4. Дисперсионный анализ количественных и качественных факторов.

В дисперсионном анализе количественными называются такие факторы, которые можно измерить в каких-то физических единицах измерения (m, c, kg и т. д.), а также в относительных или комбинированных единицах, составленных из физических единиц (m/c^2 , проценты, безразмерные отношения типа коэффициента трения и т. п.). Это рассматриваемые в теории подобия размерные, безразмерные параметры и критерии. Качественные факторы не измеряются, а нумеруются. К ним относятся, например, тип станка, день недели, марка автомобиля, номер очереди и т. д. В отличие от многих других видов анализа, в том числе рассматриваемого в следующей главе регрессионного, дисперсионный анализ успешно работает как с количественными, так и с качественными факторами, в этом его одно из главных достоинств, определяющих область применения.

Пусть в нашем двухфакторном эксперименте фактор x_1 количественный. Он изменяется на a уровнях и по y принимает средние значения, равные

$$\bar{y}_{1..}, \bar{y}_{2..}, \dots, \bar{y}_{a..} .$$

Для такого фактора можно не только установить его значимость, но и примерно оценить вид зависимости $\bar{y}_{i..}$ от x_1 , то есть установить, будет ли эта зависимость близка к прямолинейной (будут сильны линейные эффекты) или она больше походит на кривую второго порядка типа параболы или гиперболы (сильны квадратичные эффекты). Можно также оценить кубические и более высокие составляющие в том уравнении, которое больше всего подошло бы для описания экспериментальных данных. Заметим, что само уравнение мы не получаем. Этим занимается другой вид анализа - регрессионный, которым мы будем заниматься в следующей главе. Дисперсионный анализ позволяет оценить лишь значимость указанных эффектов. На начальных стадиях исследования или при изучении сложных процессов могут представлять большую ценность даже те оценочные характеристики, которые получены дисперсионным анализом.

Пусть сначала фактор x_1 изменяется всего на двух уровнях x_{11} и x_{12} (рис 23). Через две точки $\bar{y}_{1..}$ и $\bar{y}_{2..}$ (на рисунке y_1 и y_2) можно провести прямую линию. Чем больше тангенс угла наклона прямой к горизонтальной оси ($\operatorname{tg} \alpha = \Delta y / \Delta x$), тем выше значимость линейного эффекта. Если интервал $\Delta x = x_{11} - x_{12}$ обозначить за 1, то количественной характеристикой значимости линейного эффекта, которую обозначим через A_1 , может служить только разность $(\bar{y}_{2..} - \bar{y}_{1..})$. Можно записать ее в следующем виде:

$$A_1 = \bar{y}_{2..} - \bar{y}_{1..} = -1 \cdot \bar{y}_{1..} + 1 \cdot \bar{y}_{2..} .$$

Добавим в эксперимент третий опыт в точке x_{13} , отстоящей от предыдущей через тот же интервал Δx . Общий линейный эффект складывается из двух разностей:

$$A_1 = (\bar{y}_{2..} - \bar{y}_{1..}) + (\bar{y}_{3..} - \bar{y}_{2..}) = -1 \cdot \bar{y}_{1..} + 0 \cdot \bar{y}_{2..} + 1 \cdot \bar{y}_{3..} .$$

Если бы все три точки $\bar{y}_{1..}$, $\bar{y}_{2..}$ и $\bar{y}_{3..}$ лежали на одной прямой, то, видимо, квадратичный эффект был бы равен нулю. В этом случае равна нулю разность двух разностей:

$$A_2 = (\bar{y}_{3..} - \bar{y}_{2..}) - (\bar{y}_{2..} - \bar{y}_{1..}) = +1 \cdot \bar{y}_{1..} - 2 \cdot \bar{y}_{2..} + 1 \cdot \bar{y}_{3..} .$$

Чем дальше линия, соединяющая эти точки, отклоняется от прямой, тем больше разность A_2 , тем круче будет та парабола, которая проходит через три экспериментальные точки. Таким образом, разность A_2 является оценкой квадратичного эффекта.

Если к эксперименту добавить четвертую точку, то таким же образом можно оценить линейный, квадратичный и кубический эффект. Очевидно, при a уровнях изменения фактора x_1 оцениваются все эффекты до $(a - 1)$ -го порядка. Введем счетчик эффектов $t = 1, 2, \dots, (a - 1)$. Видно, что оценкой любого t -го эффекта является сумма всех экспериментальных значений $\bar{y}_{i..}$, которые запишем без точек \bar{y}_i , умноженных на некоторые коэффициенты c_{it} , где i - номер экспериментальной точки, $i = 1, 2, \dots, a$, и t - номер эффекта:

$$A_t = \sum_i^a (c_{it} \cdot \bar{y}_i) .$$

Коэффициенты c_{it} называются ортогональными контрастами, и их значения приведены в табл. П8 приложения.

Фактор x_1 имеет $(a - 1)$ степеней свободы, ровно столько же будет эффектов. Следовательно, каждому эффекту принадлежит по одной степени свободы: $f_t = 1$.

Сумма квадратов отклонений каждого из эффектов вычисляется по формуле, которую запишем без вывода:

$$SS_t = \frac{\left[\sum_i^a (c_{it} \cdot \bar{y}_i) \right]^2}{n_i \cdot \sum_i^a (c_{it})^2} . \quad (54)$$

Здесь n_i - количество опытов по всему эксперименту, находящихся на i -том уровне, в нашем случае $n_i = bn$.

Заметим, что для количественного фактора в целом выполняются основные уравнения дисперсионного анализа, а именно: сумма чисел степеней свободы всех эффектов равно числу степеней свободы фактора $(a - 1)$, а сумма сумм квадратов отклонения всех эффектов равна сумме квадратов отклонений по фактору в целом. Таким образом, составные части дисперсии фактора раскладываются по линейным, квадратичным и прочим эффектам, Возможен дисперсионный анализ каждого из эффектов. Значимость каждого эффекта оценивается традиционным методом по критерию Фишера. Если F - отношение меньше табличного значения:

$$F = \frac{SS_t / 1}{SS_o / f_o} \leq F_T(p; 1; f_o) ,$$

то указанный эффект незначим.

Итак, дисперсионный анализ количественных факторов осуществляется в два этапа. На первом производится общий анализ значимости фактора. Если он значим, то можно вторым этапом произвести анализ линейных, квадратичных и прочих эффектов, оценивая характер зависимости величины y от фактора x_1 .

Пример. Пусть произведен однофакторный эксперимент по исследованию дисперсионной модели

$$y = A + \varepsilon ,$$

где A - эффект количественного фактора x_1 . Пусть количество уровней изменения фактора $a = 4$ ($i = 1, 2, 3, 4$) и на каждом уровне произведено по 8 измерений ($k = 1, 2, \dots, 8; n = 8$). Данные эксперимент проведен фактически. В нем исследовали зависимость освещенности экрана телевизионной трубки (точнее, отклонения освещенности от стандартного значения) от толщины люминофорного покрытия экрана. Результаты приведены в табл. 11.

Таблица 11

k	$i = 1$	$i = 2$	$i = 3$	$i = 4$
1	-7	-5	+5	+5
2	-3	-3	+5	-1
3	-5	-5	+5	+5
4	-5	-1	+5	+5
5	-5	-7	+5	+5
6	+5	-5	-3	+5
7	-3	+5	+5	+5
8	-5	-1	+1	+5
$y_{i\bullet}$	-28	-22	+28	+34

Последней строкой в ней представлены суммы значений по столбцам $y_{i\bullet}$. Вычислим недостающие суммы:

Общий дисперсионный анализ:

$$f_{\text{общ}} = an - 1 = 31;$$

$$f_1 = a - 1 = 3;$$

$$f_o = a(n - 1) = 28 . \text{ Проверка : } f_{\text{общ}} = f_1 + f_o .$$

$$SS_{\text{общ}} = \sum_i \sum_k y_{ik}^2 - \frac{y_{\bullet\bullet}^2}{an} = 688 - 144 / 32 = 683,5$$

$$SS_1 = \sum_i^4 y_{i\bullet}^2 / n - \frac{y_{\bullet\bullet}^2}{an} = \frac{1}{8} (28^2 + 22^2 + 28^2 + 34^2) - 144 / 32 = 396,5$$

$$SS_o = SS_{\text{общ}} - SS_1 = 683,5 - 396,5 = 287,0.$$

Критерий Фишера

$$F = \frac{SS_1 / f_1}{SS_o / f_o} = \frac{296,5 / 3}{287,0 / 28} = 12,96.$$

Табличное значение $F_T(0,95; 3; 28) = 2,95$, поэтому делаем вывод о том, что фактор x_1 значим. Так как он изменяется на 4-х уровнях, есть возможность оценить линейный, квадратичный и кубический эффекты этого фактора. Из табл. П8 приложения имеем контрасты c_{it} :

линейный эффект, $t=1$:	-3	-1	+1	+3;
квадратичный эффект, $t=2$:	+1	-1	-1	+1;
кубический эффект, $t=3$:	-1	+3	-3	+1.

По формуле (54) получаем:

$$SS_{t=1} = [-3 \cdot (-28) - 1 \cdot (-22) + 1 \cdot 28 + 3 \cdot 34]^2 / [8(9 + 1 + 1 + 9)] = 348,1$$

$$SS_{t=2} = [+1 \cdot (28) - 1 \cdot (-22) - 1 \cdot 28 + 1 \cdot 34]^2 / [8(1 + 1 + 1 + 1)] = 0$$

$$SS_{t=3} = [-1 \cdot (-28) + 3 \cdot (-22) - 3 \cdot 28 + 1 \cdot 34]^2 / [8(1 + 9 + 9 + 1)] = 48,4.$$

Проверка: $SS_1 = SS_{t=1} + SS_{t=2} + SS_{t=3}$.

Видно, что основную долю в общем эффекте фактора x_1 составляет линейный эффект, значимость его можно даже не проверять. Значимость квадратичного эффекта нулевая, а кубического - находится на предельно низком значимом уровне, так как по критерию Фишера

$$F = (48,4 / 1) / (287,0 / 28) = 4,75 < F_T(0,95; 1; 28) = 4,2.$$

Зависимость y от фактора x_1 очень слабо отклоняется от прямолинейной.

Для качественного фактора x_2 также сначала определяют его общую значимость и если он значим, то анализируют различие значений выходного параметра y по уровням изменения фактора x_2 . Например, пусть в рассмотренном выше двухфакторном эксперименте исследуют зависимость производительности (y) от количественного фактора x_1 - скорости резания - и качественного фактора x_2 - типа станка. Произведены измерения с повторением опытов в одной точке эксперимента n раз. Определяем средние значения по каждому типу станка $\bar{y}_{\bullet j}$.

Последовательность дальнейших операций дисперсионного анализа качественного фактора x_2 следующая.

1. Все значения $\bar{y}_{\bullet j}$ по всем типам станка (по $j = 1, 2, \dots, b$ уровням) ранжируем в порядке убывания значений.

2. Вычисляем нормированную ошибку среднего $S_y = \sqrt{SS_o^2 / (b \cdot f_o)}$,

где SS_o^2 и f_o - вычисленные ранее сумма квадратов отклонений и число степеней свободы ошибки воспроизводимости.

3. Выписываем из табл. П7 приложения значения рангов Дункана (при заданной вероятности p), количество которых равно $(b-1) - R_1, R_2, \dots, R_{b-1}$.

3. Умножаем ранги Дункана на нормированную ошибку, получая ряд:

$$R_1 S_y, R_2 S_y, \dots, R_{b-1} S_y.$$

5. Если для двух соседних, например, третьего и второго, значений $\bar{y}_{\cdot j}$ в ранжированном ряду разность

$$|\bar{y}_{\cdot 3} - \bar{y}_{\cdot 2}| \leq R_1 S_y$$

то такие значения статистически неразличимы. Следовательно, несмотря на кажущееся различие двух типов станков нельзя с достоверностью p утверждать, что они различны по производительности.

Аналогично, если сравниваются два значения $\bar{y}_{\cdot j}$, стоящие в ряду через один интервал, например, 3-й и 5-й, то при

$$|\bar{y}_{\cdot 3} - \bar{y}_{\cdot 5}| \leq R_2 S_y$$

устанавливается их статистическая неразличимость. Если сравниваются средние значения $\bar{y}_{\cdot j}$, отстоящие друг от друга на два и большее количество интервалов, то применяются следующие ранги Дункана.

6.5. Задача.

Методами дисперсионного анализа исследовать давление металла на валки при холодной прокатке ленты в зависимости от обжатия, толщины ленты и типа применяемой смазки валков. Пусть предварительными действиями все параметры приведены к безразмерному виду. Имеем трехфакторный эксперимент $y = f(x_1, x_2, x_3)$, причем два первых фактора – обжатие и толщина ленты – количественные, а третий – тип смазки – качественный. Пусть $a = 4$ ($i = 1, 2, 3, 4$), $b = 3$ ($j = 1, 2, 3$), $c = 3$ ($s = 1, 2, 3$). Итого количество экспериментальных узловых точек равно $abc = 4 \cdot 3 \cdot 3 = 36$.

Чтобы найти ошибку воспроизводимости, необходимо либо в каждом узле провести хотя бы по 2 опыта, тогда общее количество опытов возрастет до 72, что уже дорого, либо хотя бы в одной узловой точке провести несколько опытов, что также связано с некоторыми дополнительными затратами. С другой стороны известно, что в трехфакторном эксперименте дисперсия тройного взаимодействия находится где-то на уровне дисперсии ошибки воспроизводимости, и потому это взаимодействие незначимо. Если дисперсию тройного взаимодействия использовать в качестве знаменателя критерия Фишера, то можно установить, насколько другие эффекты более значимы, чем тройной эффект. Это позволит ранжировать все эффекты. При этом мы можем обойтись одним опытом в точке, хотя останется некоторое сомнение в том, что дисперсии тройного взаимодействия и настоящей ошибки воспроизводимости действительно близки. Не исключен, например, случай, когда дисперсия тройного взаимодействия будет близка к нулю, тогда сравнение с ней обеспечит высокую (почти бесконечную) значимость всех эффектов, в том числе тех, которые были бы незначимыми на фоне действительной

ошибки воспроизводимости.. Но, как известно, опаснее исключить из рассмотрения значимый фактор, чем сохранить незначимый.

Ради экономии мы проведем полный факторный эксперимент с одним опытом в узловой точке, будем иметь полное количество опытов $N = abc = 36$. Модель трехфакторного эксперимента (47) не содержит члена ε , так как ошибку воспроизводимости мы не имеем. В роли ошибки эксперимента будет выступать тройной эффект ABC .

В табл. 12 приведены результаты наблюдений y_{ijs} (по k суммировать нет необходимости, так как $n = 1$). В табл. 13 подсчитаны суммы, участвующие в расчетах. Там же в последнем столбце приведено число опытов, участвующих в образовании одной суммы, что позволит проследить технику исполнения дисперсионного анализа. Как видно, это число соответствует количеству членов, находящихся внутри одной скобки в предыдущем столбце.

Таблица 12

	$i = 1$	$i = 2$	$i = 3$	$i = 4$
$s = 1$	$j = 1:$ 25	31	38	43
	$j = 2:$ 30	36	44	48
	$j = 3:$ 37	44	50	54
$s = 2$	$j = 1:$ 35	42	48	54
	$j = 2:$ 37	45	50	56
	$j = 3:$ 41	48	54	60
$s = 3$	$j = 1:$ 29	35	43	46
	$j = 2:$ 30	38	45	48
	$j = 3:$ 35	41	49	51

Вычисляем числа степеней свободы и суммы квадратов отклонений по всем эффектам:

$$f_1 = a - 1 = 3; \quad SS_1 = \sum_i^a \frac{y_{i\bullet\bullet}^2}{bc} - \frac{y_{\bullet\bullet\bullet}^2}{abc} = \frac{607842}{9} - \frac{2371600}{36} = 1660,2$$

$$f_2 = b - 1 = 2; \quad SS_2 = \sum_j^b \frac{y_{\bullet j \bullet}^2}{ac} - \frac{y_{\bullet\bullet\bullet}^2}{abc} = 381,0 ;$$

$$f_3 = c - 1 = 2; \quad SS_3 = \sum_s^c \frac{y_{\bullet\bullet s}^2}{ab} - \frac{y_{\bullet\bullet\bullet}^2}{abc} = 405,6 ;$$

$$f_{12} = (a - 1)(b - 1) = ab - a - b + 1 = 6;$$

$$SS_{12} = \sum_i^a \sum_j^b \frac{y_{ij\bullet}^2}{c} - \sum_i^a \frac{y_{i\bullet\bullet}^2}{bc} - \sum_j^b \frac{y_{\bullet j\bullet}^2}{ac} + \frac{y_{\bullet\bullet\bullet}^2}{abc} = 1,6 ;$$

$$f_{13} = (a - 1)(c - 1) = ac - a - c + 1 = 6 ;$$

$$SS_{13} = \sum_i^a \sum_s^c \frac{y_{i\bullet s}^2}{b} - \sum_i^a \frac{y_{i\bullet\bullet}^2}{bc} - \sum_s^c \frac{y_{\bullet\bullet s}^2}{ab} + \frac{y_{\bullet\bullet\bullet}^2}{abc} = 9,1;$$

$$f_{23} = (b - 1)(c - 1) = bc - b - c + 1 = 4;$$

$$SS_{23} = \sum_j^b \sum_s^c \frac{y_{\bullet js}^2}{a} - \sum_s^c \frac{y_{\bullet\bullet s}^2}{ab} - \sum_j^b \frac{y_{\bullet j\bullet}^2}{ac} + \frac{y_{\bullet\bullet\bullet}^2}{abc} = 50,3 ;$$

Таблица 13

Сумма	Значение	Вычисление суммы квадратов отклонений	Число опытов
$y_{\bullet\bullet\bullet}^2$	2371600	$(25 + 30 + \dots + 48 + 51)^2$	$abc=36$
$\sum_i^a \sum_j^b \sum_s^c y_{ijs}^2$	68388	$25^2 + 30^2 + \dots + 48^2 + 51^2$	1
$\sum_j^b \sum_s^c y_{\bullet js}^2$	266859	$(25+31+38+43)^2 + (30+36+44+48)^2 + \dots + (35+41+49+51)^2$	$a = 4$
$\sum_i^a \sum_s^c y_{i\bullet s}^2$	203858	$(25+30+37)^2 + (31+36+44)^2 + \dots + (43+45+49)^2 + (46+48+51)^2$	$b = 3$
$\sum_i^a \sum_j^b y_{ij\bullet}^2$	203762	$(25+35+29)^2 + (30+37+30)^2 + \dots + (48+56+48)^2 + (54+60+51)^2$	$c = 3$
$\sum_s^c y_{\bullet\bullet s}^2$	795400	$(25+30+37+31+\dots+43+48+54)^2 + \dots + (29+30+\dots+48+51)^2$	$ab = 12$
$\sum_i^a y_{i\bullet\bullet}^2$	607842	$(25+30+37+35+\dots+30+35)^2 + \dots + (43+48+\dots+48+51)^2$	$bc = 9$
$\sum_j^b y_{\bullet j\bullet}^2$	795106	$(25+35+29+31+42+\dots+54+46)^2 + \dots + (37+41+35+\dots+54+60+51)^2$	$ac = 12$

$$f_{123} = (a - 1)(b - 1)(c - 1) = abc - ab - ac - bc + a + b + c - 1 = 12;$$

$$SS_{123} = \sum_i^a \sum_j^b \sum_s^c y_{ijs}^2 - \sum_i^a \sum_j^b \frac{y_{ij\bullet}^2}{c} - \sum_i^a \sum_s^c \frac{y_{i\bullet s}^2}{b} - \sum_j^b \sum_s^c \frac{y_{\bullet js}^2}{a} +$$

$$+ \sum_i^a \frac{y_{i\bullet\bullet}^2}{bc} + \sum_s^c \frac{y_{\bullet\bullet s}^2}{ab} + \sum_j^b \frac{y_{\bullet j\bullet}^2}{ac} - \frac{y_{\bullet\bullet\bullet}^2}{abc} = 2,3 .$$

Общее число степеней свободы и общая сумма квадратов отклонений :

$$f_{\text{общ}} = abc - 1 = 35;$$

$$SS_{\text{общ}} = \sum_i^a \sum_j^b \sum_s^c y_{ijs}^2 - \frac{y_{\dots}^2}{abc} = 2510,3 .$$

Проверка: $2510,3 = 1660,2 + 381,0 + 405,6 + 1,6 + 9,1 + 50,3 + 2,3 ,$
 $35 = 3 + 2 + 2 + 6 + 6 + 4 + 12 .$

Видно, что основной вклад в общую дисперсию вносят линейные эффекты, все они высоко значимы (можно не проверять, так как $F_T(0,95; 2; 12) = 3,9$). Но значимость первого фактора (обжатие) выше, чем остальных, у которых значимость примерна одинакова. Значимость парных взаимодействий сомнительна, поэтому их необходимо проверить по критерию Фишера:

$$F_{12} = (1,6 / 6) / (2,3 / 12) = 1,39 < F_T(6; 12) = 3,0 ;$$

взаимодействие x_1x_2 незначимо;

$$F_{13} = (9,1 / 4) / (2,3 / 12) = 11,9 > F_T(4; 12) = 3,4 ;$$

взаимодействие x_1x_3 значимо;

$$F_{23} = (50,3 / 4) / (3,3 / 12) = 65,8 > 3,4;$$

взаимодействие x_2x_3 значимо.

Два последних взаимодействия, возможно, значимы только потому, что сравнение их эффектов произведено по эффекту тройного взаимодействия. Если бы была определена ошибка воспроизводимости, то не исключено, что они также стали бы незначимыми. Эта неопределенность – плата за экономию в эксперименте.

В результате дисперсионного анализа установлено, что парные эффекты на 2-3 порядка слабее, чем линейные, поэтому из технологических соображений вполне возможно их причислить к незначимым. С большой долей уверенности можно считать, что модель нашего эксперимента могла бы быть записана в виде

$$y = A + B + C + OT , \tag{55}$$

при невысокой значимости отброшенных членов OT.

Встает вопрос, а правильно ли была выбрана стратегия эксперимента? Мы построили ПФЭ для исследования сложной модели (47), но она оказалась значительно проще. Логичнее было сначала предположить, что модель может иметь вид (55), то есть парные взаимодействия могут быть незначимыми. Для исследования более простой модели можно провести более дешевый ДФЭ и сэкономить таким образом значительное количество опытов. Если при исследовании простой модели окажется, что отброшенные члены значимы, тогда только следует усложнять эксперимент и постепенно двигаться в сторону ПФЭ. Такая логика композиционного построения эксперимента требует планирования его перед началом экспериментирования.

Прежде чем перейти к планируемым экспериментам, проведем дисперсионный анализ качественного фактора x_3 в нашей задаче и выясним, какая из смазок лучшая (способствует снижению давления на валки в наибольшей степени). Вычисляем средние значения по уровням фактора:

$$\bar{y}_{\bullet\bullet 1} = \sum_i^a \sum_j^b \frac{y_{ij1}}{ab} = (25 + 30 + \dots + 48 + 54) / 12 = 40,0$$

Аналогично $\bar{y}_{\bullet\bullet 2} = 47,5$ и $\bar{y}_{\bullet\bullet 3} = 40,8$.

По табл. П7 приложения ранги Дункана $R_1 = R_2 = 4,5$. Нормированная ошибка эксперимента (по тройному взаимодействию за неимением настоящей):

$$S_y = \sqrt{SS_o / bf_o} = \sqrt{2,3 / (3 \cdot 12)} = 0,25.$$

Получаем $R_1 S_y = R_2 S_y = 4,5 \cdot 0,25 = 1,13$.

Разность $47,5 - 40,0 > 1,13$, а разность $40,8 - 40,0 < 1,13$.

Отсюда делаем вывод, что смазки № 1 и № 3 статистически не различимы, то есть примерно одинаково воздействуют на давление металла на валки при прокатке, а смазка № 2 отличается от них в худшую сторону.

6.6. Латинские квадраты.

Для трехфакторного эксперимента самым дешевым является план, который называется латинским квадратом. Он предназначен для анализа дисперсионной модели, содержащей только линейные эффекты факторов:

$$y = A + B + C + OT + \varepsilon \quad (56)$$

Как видно, такой эксперимент служит начальной стадией исследования, когда заранее предполагается, что парные взаимодействия факторов не значимы. Если затем окажется, что это не так, то необходимо будет наметить стратегию последовательного достраивания такого эксперимента до полного факторного.

Латинским квадратом в математике (в комбинаторике) называется квадратная матрица, состоящая из a элементов, которая в каждом столбце и в каждой строке содержит каждый элемент только один раз. Например, ниже приведено два латинских квадрата при $a = 3$ и $a = 4$:

$M \ N \ L$	$A \ B \ C \ D$
$N \ L \ M$	$B \ C \ D \ A$
$L \ M \ N$	$C \ D \ A \ B$
	$D \ A \ B \ C$

Элементами латинского квадрата могут быть любые события, процессы, понятия, числа и т. п. Они обычно обозначаются латинскими буквами, поэтому и матрица получила такое название.

Применительно к планированию эксперимента элементами будут уровни одного из факторов трехфакторного эксперимента. Латинский квадрат в этом случае – это план эксперимента. Для построения такого эксперимента требуется, чтобы по всем трем факторам содержалось одинаковое количество уровней: $a = b = c$. Заметим, что ПФЭ будет содержать a^3 экспериментальных узловых точек. Латинский квадрат строится так, что ПФЭ выполняется только по двум из трех факторов, и следовательно содержит a^2 экспериментальных узлов. Уровни третьего фактора добавляются в этих точках как элементы математического латинского квадрата. Следовательно, латинский квадрат

как ДФЭ в a раз дешевле, чем ПФЭ. Дисперсионный анализ такого эксперимента осуществляется точно так же, как для полного факторного. Лишь при формальной записи сумм квадратов эффектов необходимо следить за количеством измерений, участвующих в формировании каждой соответствующей суммы числителя, которое также в a раз меньше, чем в ПФЭ; количество опытов в знаменателе также в a раз меньше, чем при ПФЭ.

В качестве примера вновь рассмотрим приведенную выше задачу о холодной прокатке ленты. В отличие от полного факторного эксперимента теперь мы обязаны установить одинаковое количество уровней по всем факторам. Если возьмем $a = b = c = 3$, то экспериментальных точек будет $a^2 = 9$. Можно позволить себе в каждой точке провести по 2 опыта ($k = 1, 2$; $n = 2$). В этом случае общее количество опытов составит $N = a^2 \cdot n = 18$, что ровно в два раза меньше, чем в ранее проведенном эксперименте. Кроме того, мы сможем определить истинную ошибку воспроизводимости, которой раньше не было. Эксперимент дешевле, но богаче по информативности. Правда, за это заплачено риском пропустить значимые парные эффекты.

Кстати, даже если бы число уровней увеличилось до 4, то при $n = 2$ эксперимент, спланированный по латинскому квадрату, все равно оказался бы дешевле ранее проведенного: $N = a^2 \cdot n = 16 \times 2 = 32$.

В табл. 14 приведены экспериментальные данные эксперимента y_{ijsk} (жирным шрифтом), спланированного по латинскому квадрату. Видно, что

Таблица 14

	$i = 1$	$i = 2$	$i = 3$	$y_{\dots s \bullet}$
$s = 1$! 26 $j=1$! 24 ! 50	! 38 $j=2$! 34 ! 72	! 51 $j=3$! 49 ! 100	222
$s = 2$! 36 $j=2$! 38 ! 74	! 50 $j=3$! 46 ! 96	! 49 $j=1$! 47 ! 96	266
$s = 3$! 36 $j=3$! 34 ! 70	! 37 $j=1$! 33 ! 70	! 46 $j=2$! 44 ! 90	230
$y_{i \dots}$	194	238	286	718

$a = b = c = 3$ и $n = 2$, всего $N = 18$ опытов. Уровни факторов x_1 и x_3 перебирают все возможные их сочетания (ПФЭ 3×3), они образуют узловые точки эксперимента. Уровни второго фактора x_2 приписываются в каждом узле в соответствии с буквами M , N и L латинского квадрата. В таблице приведены также некоторые суммы (нежирным шрифтом). Имеется 9 сумм $y_{ijs \bullet}$ внутри больших клеточек (в узловых точках), по три суммы $y_{i \dots}$ (в

нижней строке) и $y_{\dots s}$ (в правом столбце) и одна общая сумма $y_{\dots} = 718$.
Недостает лишь сумм $y_{\cdot j \dots}$, вычислим их отдельно:

$$y_{\cdot 1 \dots} = 50 + 96 + 70 = 216;$$

$$y_{\cdot 2 \dots} = 72 + 74 + 90 = 236;$$

$$y_{\cdot 3 \dots} = 100 + 96 + 70 = 266.$$

Для вычислений потребуется также общая сумма квадратов:

$$\sum_i^a \sum_j^a \sum_s^a \sum_k^n y_{ijsk}^2 = 26^2 + 24^2 + 36^2 + \dots + 46^2 + 44^2 = 29782.$$

Проводим дисперсионный анализ модели (56).

$$f_1 = f_2 = f_3 = (a - 1) = 2; \quad f_{\text{общ}} = N - 1 = 17; \quad f_o = a^2 \cdot (n - 1) = 9.$$

Отброшенные члены имеют число степеней свободы

$$f_{\text{от}} = 17 - 2 - 2 - 2 - 9 = 2.$$

Суммы квадратов:

$$SS_{\text{общ}} = \sum_i^3 \sum_j^3 \sum_s^3 \sum_k^2 y_{ijsk}^2 - \frac{y_{\dots}^2}{a^2 n} = 1141,8;$$

$$SS_1 = \sum_i^3 \frac{y_{i \dots}^2}{an} - \frac{y_{\dots}^2}{a^2 n} = 705,7;$$

$$SS_2 = \sum_j^3 \frac{y_{\cdot j \dots}^2}{an} - \frac{y_{\dots}^2}{a^2 n} = 211,1;$$

$$SS_3 = \sum_s^3 \frac{y_{\dots s}^2}{an} - \frac{y_{\dots}^2}{a^2 n} = 183,1;$$

$$SS_o = \sum_i^3 \sum_j^3 \sum_s^3 \sum_k^2 y_{ijsk}^2 - \sum_i^3 \sum_j^3 \sum_s^3 \frac{y_{ijs}^2}{n} = 36;$$

$$SS_{\text{от}} = SS_{\text{общ}} - SS_1 - SS_2 - SS_3 - SS_o = 1141,8 - 705,7 - 211,1 - 183,1 - 36 = 5,9$$

Следует обратить внимание на то, как записываются знаменатели дробей, входящих в выражения для сумм квадратов: по-прежнему значение знаменателя равно количеству слагаемых, образующих одну сумму, входящую в числитель. Оно везде в a раз меньше, чем при ПФЭ.

Проверка по критерию Фишера при $F_T(2; 9) = 4,3$:

$$F_1 = (705,7 / 2) / (36 / 9) = 88,2;$$

$$F_2 = (211,1 / 2) / (36 / 9) = 26,4;$$

$$F_3 = (183,1 / 2) / (36 / 9) = 22,9;$$

$$F_{\text{от}} = (5,9 / 2) / (36 / 9) = 0,73.$$

Видно, что все линейные эффекты сильно значимы. Отброшенные члены все в сумме незначимы, то есть начальное предположение об их незначимости оправдалось. Мы провели дешевый эксперимент по плану латинского квадрата и сэкономили значительное количество опытов. Эксперимент получился более информативным, чем ранее проведенный ПФЭ, поскольку появи-

лась уверенность в том, что парные эффекты незначимы. Остальные выводы точно такие же, что и в том эксперименте. Из линейных эффектов наиболее значимым является фактор x_1 (обжатие), а два других – толщина x_2 и тип смазки x_3 – по значимости примерно одинаковы. Так же по рангам Дункана можно было бы провести анализ качественного фактора с теми же выводами об эффективности применяемых смазок.

6.7 Латинские кубы, греко-латинские квадраты, кубы и гиперкубы.

Идея латинского квадрата, примененная для трехфакторного, может быть распространена на четырехфакторный эксперимент. По трем факторам x_1 , x_2 и x_3 следует спланировать ПФЭ, который при a уровнях изменения каждого из них потребует a^3 узловых точек. Уровни четвертого фактора x_4 приписываются к опытам в этих точках как элементы латинского куба. В математике латинским кубом называется трехмерная матрица, которая по всем трем осям имеет по a элементов, причем в каждом направлении каждый элемент встречается только один раз. Например, при $a = 3$ получим трехмерную матрицу,

Таблица 15

1-й слой			2-й слой			3-й слой		
1 1 1	1 1 2	1 1 3	2 1 1	2 1 2	2 1 3	3 1 1	3 1 2	3 1 3
1	2	3	2	3	1	3	1	2
1 2 1	1 2 2	1 2 3	2 2 1	2 2 2	2 2 3	3 2 1	3 2 2	3 2 3
2	3	1	3	1	2	1	2	3
1 3 1	1 3 2	1 3 3	2 3 1	2 3 2	2 3 3	3 3 1	3 3 2	3 3 3
3	1	2	1	2	3	2	3	1

состоящую из трех слоев, наложенных друг на друга. План такого эксперимента представлен тремя плоскими таблицами 15. В ней верхней строчкой обозначены номера уровней первых трех факторов, образующих полный их перебор, а второй строчкой – номера уровней 4 го фактора, образующих латинский куб.

Такой эксперимент в a раз дешевле ПФЭ, содержащего a^4 опытов. Он предназначен для анализа дисперсионной модели

$$y = A + B + C + D + OT + \varepsilon, \quad (57)$$

в которой представлены только линейные эффекты, а все парные и более высокие взаимодействия причислены в группу отброшенных эффектов. Дисперсионный анализ модели ничем не отличается от приведенного выше.

Латинский квадрат как план 4-х факторного эксперимента уже становится дорогатым. Так при $a = 3$ и при количестве опытов в одной точке $n = 2$ будем иметь $N = a^3 \cdot n = 54$. Это много, хотя ПФЭ потребовал бы 162 опыта. В связи с этим идея латинских кубов не применяется для экспериментов с ко-

личеством факторов больше 4-х, не говоря уж о том, что сложно написать многомерный латинский гиперкуб в многомерном пространстве.

Для 4-хфакторного эксперимента существуют более дешевые планы. Один из них построен на использовании греко-латинского квадрата. Греко-латинским квадратом в комбинаторике называется квадратная матрица, которая состоит из двух групп по a элементов, обозначаемых соответственно греческими и латинскими буквами, причем в каждой строке и в каждом столбце элементы из каждой группы встречаются только по одному разу. Например, при $a = 5$ и $a = 3$:

$A\alpha$	$B\beta$	$C\gamma$	$D\delta$	$E\epsilon$		$A\alpha$	$B\beta$	$C\gamma$
$C\delta$	$D\epsilon$	$E\alpha$	$A\beta$	$B\gamma$		$B\gamma$	$C\alpha$	$A\beta$
$E\beta$	$A\gamma$	$B\delta$	$C\epsilon$	$D\alpha$		$C\beta$	$A\gamma$	$B\alpha$
$B\epsilon$	$C\alpha$	$D\beta$	$E\gamma$	$A\delta$				
$D\gamma$	$E\delta$	$A\epsilon$	$B\alpha$	$C\beta$				

Четырехфакторный эксперимент, план которого называется греко-латинским квадратом, содержит a^2 опытных узловых точек, которые образуются путем перебора всех уровней только двух факторов (ПФЭ по двум факторам), а уровни следующих двух факторов представляют собой элементы греко-латинского квадрата. Вместо a^4 опытов, необходимых при ПФЭ, имеем только a^2 опытов, экономия в a^2 раз ! Например, при $a = 3$ план эксперимента содержит всего 9 точек, представленных в табл. 16. Для лучшего сравнения этого плана с предыдущим (латинским кубом, табл. 16) здесь так же первой строкой показаны номера уровней первых двух факторов, выбранных по принципу полного перебора, и второй строкой - номера уровней других двух факторов, образующих греко-латинский квадрат. При двух опытах в каждой точке общее количество опытов в эксперименте $N = a^2 \cdot n = 18$, что вполне приемлемо.

Таблица 16

1 1	2 1	3 1
1 1	2 2	3 3
1 2	2 2	3 2
2 3	3 1	1 2
1 3	2 3	3 3
3 2	1 3	2 1

Такой эксперимент позволяет произвести дисперсионный анализ четырехфакторной модели (57), который проводится совершенно по тем же правилам, которые рассмотрены и продемонстрированы на примерах выше. Только количество членов в суммах, и значение знаменателей при расчетах сумм квадратов будет в a^2 раз меньше, чем при ПФЭ.

Для 5-ифакторного эксперимента, видимо, неплохими свойствами будет обладать план, построенный по принципу греко-латинского куба, в котором

только три фактора выбираются путем полного перебора, а уровни двух оставшихся факторов выбираются в этих опытах как элементы греко-латинского куба. Для такого 5-факторного эксперимента потребуется a^3 опытов. По-прежнему, такой эксперимент даст возможность проанализировать значимость только пяти линейных эффектов и оценить суммарную значимость всех остальных (отброшенных) эффектов.

Практически идеи латинских и греко-латинских квадратов и кубов далее пятифакторных экспериментов продвинуть трудно. Существуют другие также экономные и даже более дешевые планы. Свойства каждого из них, а также виды обслуживаемых ими моделей, можно найти в специальной литературе, например, в справочнике [8]. Среди таких планов существуют сбалансированные и несбалансированные, полноблочные и неполноблочные. Сбалансированными называются такие планы, в которых в каждой строке и каждом столбце одинаковые уровни факторов встречаются одинаковое количество раз, иначе они становятся несбалансированными. Полноблочные планы предусматривают проведение опытов во всех точках эксперимента, а в неполноблочных некоторые плановые точки пропускаются. Например, в табл.17 приведен неполноблочный сбалансированный план Юдена,

Таблица 17.

Уровни фактора x_1	Уровни фактора x_2			
	1	2	3	4
1	+	-	+	+
2	-	+	+	+
3	+	+	+	-
4	+	+	-	+

применяемый для двухфакторного эксперимента, в котором каждый фактор изменяется на 4-х уровнях, а знаком “ - ” обозначены пропускаемые опыты. Этот план дешевле, чем ПФЭ 4×4 , он содержит 12 опытов вместо 16, но также позволяет провести полный дисперсионный анализ всех эффектов двухфакторной модели (46).

Контрольные вопросы по главе 6.

1. Какие задачи решает дисперсионный анализ?
2. Как выглядит дисперсионная модель процесса для двух- и трехфакторного эксперимента?
3. Напишите дисперсионную модель четырехфакторного эксперимента.
4. Как выглядят основные уравнения дисперсионного анализа, и как проводится дисперсионный анализ?
5. Напишите, пользуясь правилом, сумму квадратов отклонений для парного взаимодействия факторов x_1 и x_3 трехфакторного эксперимента.
6. Как вычисляется ошибка воспроизводимости эксперимента?

7. Какие достоинства и недостатки имеет эксперимент, построенный по плану «латинский квадрат»?
8. Как и для каких целей проводится эксперимент латинский квадрат?
9. Как обрабатывается эксперимент, построенный по плану латинского квадрата?
10. Как и для каких экспериментов строится план «латинский куб» и «греко-латинский квадрат»?

7. РЕГРЕССИОННЫЙ АНАЛИЗ.

7.1. Задачи регрессионного анализа

Регрессионный анализ, в отличие от дисперсионного, может работать только с количественными факторами. Он предназначен для построения количественной модели процесса, в которой общее уравнение (1), связывающее k входных параметров (x_1, x_2, \dots, x_k) с выходным параметром y , представляет собой некоторую формулу, расчеты по которой хорошо совпадают с экспериментальными данными. Такая формула или математическая модель процесса в общем случае может иметь любой вид, но ее всегда можно разложить в ряд Тейлора (32). Мы уже отмечали, что сложные процессы выгодно разбить на несколько более простых с тем, чтобы каждый из них можно было бы описать уравнением, в котором переменные содержат степени не выше второй:

$$\begin{aligned}
 y_p = & a_0 + a_1 x_1 + a_2 x_2 + \dots + a_k x_k + \\
 & + a_{12} x_1 x_2 + a_{13} x_1 x_3 + \dots + a_{k-1,k} x_{k-1} x_k + \\
 & + a_{11} x_1^2 + a_{22} x_2^2 + \dots + a_{kk} x_k^2
 \end{aligned} \tag{58}$$

Каждая строка в этом уравнении носит отдельное название. В первой строке записаны свободный и линейные члены, во второй – парные произведения первых степеней переменных, которые называются парными взаимодействиями и отражают эффекты совместного влияния факторов на величину y , в третьей – квадратичные члены. Более сложные модели в регрессионном анализе не рассматриваются. Обычно в реальной задаче стараются оставить возможно меньшее количество членов уравнения (58), но достаточное для правильного (адекватного) описания процесса. Уравнение (58) называется также регрессией. Таким образом, слова «уравнение» и «регрессия» в регрессионном анализе (и только в нем!) – синонимы. Модель процесса, содержащая только первые две строки уравнения (58), называется линейной. Она содержит свободный член, переменные в первой степени и парные взаимодействия линейных эффектов.

7.2. Построение регрессии.

Пусть в уравнении (58) оставлено $(m + 1)$ членов и их число меньше, чем количество N экспериментальных точек (узлов). В соответствии с тео-

ремой 5.2, такая модель в общем случае неадекватна, и мера неадекватности (33) больше нуля. Введем следующие обозначения:

$$\begin{aligned} a_{12} &= a_{k+1}, & x_1 x_2 &= x_{k+1}; \\ a_{13} &= a_{k+2}, & x_1 x_3 &= x_{k+2}; \\ & \dots\dots\dots \\ a_p &= a_m, & x_p &= x_m. \end{aligned}$$

Здесь p – индекс при последнем сохраненном члене.

Введем также фиктивное переменное x_0 , которое всегда равно 1, тогда уравнение (58) будет представлено в виде:

$$y_p = a_0 x_0 + a_1 x_1 + a_2 x_2 + a_k x_k + a_{k+1} x_{k+1} + \dots + a_m x_m = \sum_{i=0}^m a_i x_i. \quad (59)$$

В этом уравнении счетчик номера членов i изменяется от 0 до m :

$$i = 0, 1, 2, \dots, m.$$

Нулевой номер присвоен свободному члену $a_0 x_0$, который в регрессионном анализе играет особую роль.

Для однофакторной модели, как отмечалось в 5.3, можно также ввести фиктивное переменное x_0 и произвести замену переменных: $x_1^2 = x_2$, $x_1^3 = x_3$ и т. д. Тогда она записывается также в виде уравнения (59). В обоих случаях $(m + 1) < N$ и уравнение (59) становится линейным относительно неизвестных коэффициентов a_0, a_1, \dots, a_m , входящих в него.

Необходимо так построить данное уравнение, чтобы мера неадекватности была минимальной. Такая постановка задачи приводит к известному методу наименьших квадратов. Мы повторим некоторые основные положения этого метода.

Итак, пусть в каждой узловой точке проведено по одному опыту. Получим таблицу 18 экспериментальных данных, содержащую N строк.

Таблица 18

№ опыта u	Уровни факторов				y_u
	x_{1u}	x_{2u}	\dots	x_{ku}	
1	x_{11}	x_{21}		x_{k1}	y_1
2	x_{12}	x_{22}		x_{k2}	y_2
N	x_{1N}	x_{2N}		x_{kN}	y_N

Требуется еще раз разобраться со счетчиками, которые мы зафиксируем на все дальнейшее изложение:

k - количество входных параметров (факторов),

u - счетчик числа узловых точек эксперимента, $u = 1, 2, \dots, N$,

N - число таких точек,

i - счетчик числа членов регрессии (59), $i = 0, 1, 2, \dots, m$, для этих же целей потребуется еще один счетчик $j = 0, 1, 2, \dots, m$.

Минимизация меры неадекватности (33) равнозначна решению системы $(m + 1)$ уравнений с $(m + 1)$ неизвестными коэффициентами a_0, a_1, \dots, a_m :

$$\frac{\partial SS_{ad}}{\partial a_0} = 0; \quad \frac{\partial SS_{ad}}{\partial a_1} = 0; \quad \dots, \quad \frac{\partial SS_{ad}}{\partial a_m} = 0 \quad (60)$$

Запишем SS_{ad} подробнее, подставив в (59) опытные данные из табл. 18:

$$SS_{ad} = [(y_1 - a_0 x_{01} - a_1 x_{11} - \dots - a_m x_{m1})^2 + \quad (\text{по 1-й строке табл.18}) \\ + (y_2 - a_0 x_{02} - a_1 x_{12} - \dots - a_m x_{m2})^2 + \quad (\text{по 2-й строке}) \\ \dots \dots \dots \\ + (y_N - a_0 x_{0N} - a_1 x_{1N} - \dots - a_m x_{mN})^2] \quad (\text{по } N\text{-й строке}).$$

Для примера возьмем частную производную по a_0 только от первой строки выражения. Получим:

$$2 a_0 x_{01} x_{01} + 2 a_1 x_{01} x_{11} + 2 a_2 x_{01} x_{21} + \dots + 2 a_m x_{01} x_{m1} - 2 x_{01} y_1 .$$

Просуммируем по всем строкам. Затем выполним подобные действия с производными по другим неизвестным. Получим систему уравнений, которая называется *системой нормальных уравнений*:

$$a_0 \sum_u^N x_{0u} x_{0u} + a_1 \sum_u^N x_{0u} x_{1u} + \dots + a_m \sum_u^N x_{0u} x_{mu} = \sum_u^N x_{0u} y_u \\ a_0 \sum_u^N x_{1u} x_{0u} + a_1 \sum_u^N x_{1u} x_{1u} + \dots + a_m \sum_u^N x_{1u} x_{mu} = \sum_u^N x_{1u} y_u \quad (61) \\ \dots \dots \dots \\ a_0 \sum_u^N x_{mu} x_{0u} + a_1 \sum_u^N x_{mu} x_{1u} + \dots + a_m \sum_u^N x_{mu} x_{mu} = \sum_u^N x_{mu} y_u .$$

Для лучшего запоминания систему нормальных уравнений удобно записать в виде двух матриц, введя обозначения

$$\sum_u^N x_{iu} x_{ju} = ij \quad \text{и} \quad \sum_u^N x_{ju} y_u = jy ,$$

то есть, записав только индексы при факторах:

$$(ij) = \begin{vmatrix} 00 & 01 & \dots & 0m \\ 10 & 11 & \dots & 1m \\ \dots & \dots & \dots & \dots \\ m0 & m1 & \dots & mm \end{vmatrix}; \quad (jy) = \begin{vmatrix} 0y \\ 1y \\ \dots \\ my \end{vmatrix} .$$

Первая из этих матриц называется нормальной или информационной. Таким образом, исходная система нормальных уравнений (61) в матричной форме записывается в виде:

$$(ij) = (jy) .$$

Как видно, в матрицах и в системе уравнений (61) счетчик i применяется для нумерации столбцов или слагаемых в уравнении, а счетчик j - для но-

мера строк. Счетчик i также совпадает с индексом при неизвестном коэффициенте a .

7. 3. Обращенная матрица. Планирование эксперимента.

Один из методов решения системы нормальных уравнений (61) состоит в обращении нормальной матрицы, то есть построении обращенной матрицы (C_{ij}) той же размерности, что и нормальная:

$$C_{ij} = \begin{vmatrix} C_{00} & C_{01} & \dots & C_{0m} \\ C_{10} & C_{11} & \dots & C_{1m} \\ \dots & \dots & \dots & \dots \\ C_{m0} & C_{m1} & \dots & C_{mm} \end{vmatrix} . \quad (62)$$

Известно, что элементы обращенной матрицы вычисляются по формуле:

$$C_{ij} = (ij)' / \Delta ,$$

где Δ - определитель матрицы $\{ij\}$,

а $(ij)'$ - алгебраическое дополнение, то есть умноженный на $(-1)^{(i+j)}$ определитель матрицы, которая отличается от нормальной тем, что в ней отсутствует j -я строка и i -й столбец. Обращение матриц осуществляется на ЭВМ по стандартной программе.

Обращенная матрица имеет большое значение в планировании эксперимента и в регрессионном анализе. Прежде всего, она позволяет вычислять коэффициенты регрессии по формуле:

$$a_i = \sum_{j=0}^m [C_{ij} \cdot (jy)] . \quad (63)$$

Для нахождения коэффициента a_i необходимо все элементы i - го столбца перемножить на элементы матрицы $\{jy\}$. Например,

$$\begin{aligned} a_1 &= C_{01} \cdot (0y) + C_{11} \cdot (1y) + \dots + C_{m1} \cdot (my) = \\ &= C_{01} \cdot \sum_{u=1}^N x_{0u} y_u + C_{11} \cdot \sum_{u=1}^N x_{1u} y_u + \dots + C_{m1} \cdot \sum_{u=1}^N x_{mu} y_u . \end{aligned}$$

Коэффициенты регрессии a_i являются случайными величинами, зависящими от серии проведенных опытов (в том смысле, что при другом эксперименте также из N опытов получатся несколько отличные коэффициенты). Точные значения коэффициентов находятся в доверительном интервале $a_i \pm \Delta a_i$, который определяется дисперсией этого коэффициента $S^2(a_i)$ (точнее, стандартом $S(a_i)$):

$$\Delta a_i = t_p \cdot S(a_i),$$

где t_p - коэффициент Стьюдента.

Оценка дисперсии $S^2(a_i)$ осуществляется с помощью диагональных членов обращенной матрицы:

$$S^2(a_i) = C_{ii} \cdot S_0^2 ,$$

где S_0^2 - известная из дисперсионного анализа ошибка воспроизводимости эксперимента (позже мы к ней вернемся).

Недиагональные члены обращенной матрицы также характеризуют своеобразную ошибку вычисления коэффициентов регрессии, которая называется *ковариацией*. В связи с этим обращенную матрицу называют также *матрицей ошибок* или *ковариационной матрицей*. Чтобы понять, что такое ковариация, необходимо проследить алгоритм решения системы нормальных уравнений методом Гаусса (в основе метода обращения нормальной матрицы по существу лежат те же преобразования). По методу Гаусса решение системы осуществляется в два этапа. На первом – сначала из первого уравнения первое неизвестное выражается через все остальные и подставляется в другие уравнения, в которых число неизвестных уменьшается на одно. (Напомним, что неизвестными здесь являются коэффициенты регрессии a_i). Затем то же самое проделывается со вторым неизвестным второго уравнения. И так далее. На каждом шаге во всех последующих уравнениях количества переменных сокращается. В последнем уравнении останется лишь одно неизвестное, которое из него определяется. Вторым этапом, двигаясь в обратном направлении, находят все остальные неизвестные. Такой алгоритм реализуется на ЭВМ по стандартной программе. Важно отметить два существенных недостатка этого алгоритма. Первый связан с методикой вычислений. Результат расчетов может зависеть от того, в какой последовательности записаны уравнения в исходной системе нормальных уравнений. К примеру, если эту систему переписать в обратной последовательности, то коэффициенты регрессии могут получиться иными. Это связано с тем, что при численном решении системы результаты обязательно округляются, и при переходе от одного уравнения к другому ошибка округления постепенно накапливается. Последний коэффициент может содержать уже существенную ошибку. Чтобы уменьшить (но не устранить полностью) ошибку вычислений, в стандартную программу на ЭВМ включают подпрограмму сортировки уравнений исходной системы нормальных уравнений, по которой на первое место всегда выводится уравнение, которое имеет наибольший коэффициент при том неизвестном, которое выражается через остальные неизвестные.

Второй недостаток вызван тем, что сами коэффициенты регрессии a_i являются величинами случайными и содержат ошибку, связанную со случайным характером серии проделанных опытов. При описанных преобразованиях ошибка одного коэффициента переносится на все следующие коэффициенты. Влияние ошибок коэффициента a_i на ошибки коэффициента a_j , связанные со случайным характером этих величин, называется ковариацией. Она оценивается недиагональными членами обращенной матрицы:

$$\text{cov}(a_i, a_j) = C_{ij} \cdot S_0^2.$$

Количественной мерой ковариации служит коэффициент ковариации

$$\rho(a_i, a_j) = \frac{C_{ij}}{\sqrt{C_{ii} \cdot C_{jj}}},$$

который изменяется от 0 до 1. Он близок к 1, когда влияние по ошибкам одного коэффициента на другой значительно. При $\rho(a_i, a_j) = 0$ случайная

ошибка коэффициента a_i никак не отражается на ошибке коэффициента a_j , в этом смысле эти коэффициенты независимы.

Поскольку коэффициенты регрессии являются случайными величинами, то случайным будет также расчетное значение y_p , вычисленное по формуле (59). По закону сложения ошибок дисперсия значения y_p складывается из дисперсий членов регрессии. Для u -той точки эксперимента эта сумма будет иметь вид:

$$S^2(y_{pu}) = \left(\frac{\partial y_p}{\partial a_0} \right)_u^2 \cdot S^2(a_0) + \left(\frac{\partial y_p}{\partial a_1} \right)_u^2 \cdot S^2(a_1) + \dots + \left(\frac{\partial y_p}{\partial a_m} \right)_u^2 \cdot S^2(a_m)$$

Так как $\left(\frac{\partial y_p}{\partial a_i} \right)_u^2 = x_{iu}^2$, то

$$S^2(y_{pu}) = \sum_{i=0}^m x_{iu}^2 \cdot S^2(a_i). \quad (64)$$

Выражение (64) представляет собой уравнение эллипсоида в m -мерном пространстве и называется эллипсоидом рассеяния. Чем меньше эллипсоид рассеяния, тем с большей точностью расчетное значение y_{pu} совпадает с экспериментальным y_u .

Очевидно, выгодно заранее спланировать эксперимент так, чтобы коэффициенты ковариации по всем коэффициентам регрессии равнялись нулю, и каждый из членов регрессии содержал бы только собственную ошибку, не зависящую от ошибок других коэффициентов. Желательно также, чтобы эллипсоид рассеяния был минимальным и не очень вытянутым по осям. Но, к сожалению, не всегда удается удовлетворить эти требования в одном плане. Все существующие планы экспериментов отличаются между собой теми или иными требованиями к матрице ошибок или к эллипсоиду рассеяния.

Очень эффективны так называемые *ортогональные* планы, которые строятся так, чтобы в матрице ошибок все элементы, не лежащие на главной диагонали, обратились в нуль: $C_{ij} = 0$ при $i \neq j$. Это произойдет тогда, когда в системе нормальных уравнений (61) все недиагональные члены будут равны нулю:

$$\sum_u^N x_{iu} x_{ju} = 0 \text{ при всех значениях } i \neq j. \quad (65)$$

Выражение (65) является математической формулировкой ортогонального плана эксперимента. При этом система нормальных уравнений (61) распадается на N независимых друг от друга уравнений (строк), и каждый коэффициент регрессии вычисляется непосредственно из одной строки:

$$a_i = \frac{\sum_{u=1}^N x_{iu} y_u}{\sum_{u=1}^N x_{iu} x_{iu}}. \quad (66)$$

Дисперсии и доверительные интервалы каждого из коэффициентов регрессии будут также независимыми друг от друга. Диагональные члены обращенной матрицы будут определяться выражением:

$$C_{ij} = \frac{1}{\sum_{u=1}^N x_{iu} x_{ju}}$$

Но ортогональные планы не всегда обеспечивают минимальное значение эллипсоида рассеяния. Кроме того, вытянутость эллипсоида в направлении каждой из осей может быть сильно различной, а это значит, что точность вычисления y_{pi} будет разной в различных направлениях. Планы, которые требуют, чтобы рассеяние по всем осям было одинаковым, называется *рототабельными*. Они также достигаются при определенных соотношениях элементов в матрице ошибок. Но рототабельные планы не всегда ортогональны. И, в общем случае, планов, оптимальных по всем критериям, не существует.

D-оптимальным называется такой план, которому соответствует минимальный определитель матрицы ошибок (или максимальный определитель нормальной матрицы). Так как определитель пропорционален объему эллипсоида рассеяния, то такие планы обеспечивают минимум обобщенной дисперсии всех оценок коэффициентов регрессии.

A-оптимальный план строится так, чтобы в матрице ошибок была минимальной сумма элементов, лежащих на главной диагонали. Такие планы минимизируют сумму дисперсий только линейных, то есть наиболее важных коэффициентов регрессии.

E-оптимальному плану соответствует минимальное значение максимального из диагональных членов матрицы ошибок, при этом минимизируется дисперсия (и ошибка) того члена регрессии, для которого она максимальна. Эллипсоид рассеяния становится менее вытянутым, ковариации таким образом также снижаются.

Выбор того или иного плана определяется характером задачи.

7. 4. Ортогональные планы первого порядка.

Линейной регрессией в *k*-факторном эксперименте называется уравнение вида:

$$\begin{aligned} y_p = & a_0 + a_1 x_1 + a_2 x_2 + \dots + a_k x_k + \\ & + a_{12} x_1 x_2 + a_{13} x_1 x_3 + \dots + a_{k-1,k} x_{k-1} x_k + \\ & + a_{123} x_1 x_2 x_3 + \dots \\ & + a_{1234} x_1 x_2 x_3 x_4 + \dots + a_{12\dots k} x_1 x_2 \dots x_k . \end{aligned}$$

В нем содержится свободный член, члены с первыми степенями факторов, а также парные, тройные и т. д. до *k*-го порядка взаимодействия, представляющие собой произведения первых степеней факторов.

Планы экспериментов, которые предназначены для построения линейной регрессии, называются планами первого порядка. Доказано, что ортогональные планы первого порядка одновременно являются рототабельными и *D* -

оптимальными, поэтому ортогональность таких экспериментов является универсальным критерием их оптимальности.

Чтобы план первого порядка стал ортогональным, необходимо выполнить три условия:

- эксперимент должен быть полным факторным (или специальным образом спланированным), и в каждой узловой точке такого эксперимента должно быть проведено по одному опыту; если в некоторых точках проведено несколько опытов, то в расчетах должны использоваться средние значения;
- по каждому фактору x_1, x_2, \dots, x_k уровни изменения факторов должны быть равноотстоящими, то есть расстояния между уровнями $\Delta x_i = const$;
- оси координат факторов должны быть перенесены в центр эксперимента путем замены переменных по формуле:

$$x'_i = \frac{x_i - (x_{i \max} + x_{i \min})}{0,5 \cdot \Delta x_i}, \quad (67)$$

где x'_i и x_i - новые и старые координаты, $x_{i \max}$ и $x_{i \min}$ - максимальный и минимальный уровни фактора x_i .

Пример.

На рис. 30 графически представлен полный двухфакторный эксперимент с равноотстоящими уровнями. Фактор x_1 изменяется на трех уровнях, принимая значения 3,0; 6,0 и 9,0. По фактору x_2 четыре уровня - 2,0; 4,0; 6,0 и 8,0. В каждой точке проведено по три опыта, опытные данные представлены в табл. 19. Итого имеем 12 экспериментальных точек. Средние значения по трем опытам в одной точке, участвующие в построении регрессии, также приведены в таблице.

Таблица 19

x_1	x_2							
	2,0		4,0		6,0		8,0	
3,0	15,1	15,3	17,3	17,5	19,6	19,8	22,0	22,0
	15,3		17,8		19,8		21,8	
	15,4		17,4		20,0		22,2	
6,0	14,2	14,4	16,9	17,1	20,0	20,0	22,6	22,8
	14,7		17,3		20,1		22,8	
	14,4		17,1		19,8		23,0	
8,0	13,3	13,3	16,6	16,6	19,9	19,9	23,5	23,5
	13,2		16,8		20,0		23,6	
	13,4		16,4		19,8		23,5	

По результатам эксперимента можно построить регрессию, содержащую 12 членов, но мы ограничимся четырьмя, построив линейную модель:

$$y_p = a_0 + a_1 x_1 + a_2 x_2 + a_{12} x_1 x_2.$$

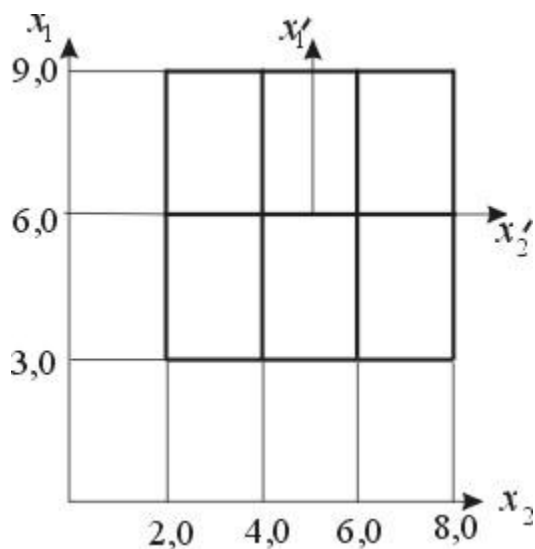


Рис. 30. План двухфакторного эксперимента

Такое уравнение не обязано быть адекватным, однако к анализу адекватности мы вернемся позже.

. Для регрессионного анализа удобно построить рабочую табл. 20 экспериментальных данных, в которой под y_u фигурируют усредненные экспериментальные значения. В нее введено фиктивное переменное x_0 , которое всегда равно +1, а также переменное $x_3 = x_1x_2$.

Таблица 20

u	x_{0u}	x_{1u}	x_{2u}	x_{3u}	y_u
1	1	3	2	6	15,3
2	1	3	4	12	17,5
3	1	3	6	18	19,8
4	1	3	8	24	22,0
5	1	6	2	12	14,4
6	1	6	4	24	17,1
7	1	6	6	36	20,0
8	1	6	8	48	22,8
9	1	9	2	18	13,3
10	1	9	4	36	16,6
11	1	9	6	54	19,9
12	1	9	8	72	23,5

Сначала сознательно не будем вводить новые оси координат с тем, чтобы увидеть недостатки неортогонального плана. Обрабатываем данные методом наименьших квадратов. Получим систему нормальных уравнений (61), содержащую 4 уравнения с 4-мя неизвестными коэффициентами a_i . Запишем ее в матричном виде, то есть в виде двух матриц:

- матрицы (ij)

$$(ij) = \begin{vmatrix} 00 & 01 & 02 & 03 \\ 10 & 11 & 12 & 13 \\ 20 & 21 & 22 & 23 \\ 30 & 31 & 32 & 33 \end{vmatrix} = \begin{vmatrix} 12 & 72 & 60 & 360 \\ 72 & 504 & 360 & 2520 \\ 60 & 360 & 360 & 2160 \\ 360 & 2520 & 2160 & 15120 \end{vmatrix} .$$

- и матрица (jy)

$$(jy) = \begin{vmatrix} 0y \\ 1y \\ 2y \\ 3y \end{vmatrix} = \begin{vmatrix} 222,2 \\ 1329,3 \\ 1194,4 \\ 7187,4 \end{vmatrix} .$$

Для примера, вычислим некоторые из числовых значений, занесенных в матрицу (данные берем из табл. 20):

$$(12) = \sum_{u=1}^N x_{1u}x_{2u} = 3 \cdot 2 + 3 \cdot 4 + \dots + 9 \cdot 8 = 360 ,$$

$$(3y) = \sum_{u=1}^N x_{3u}y_u = 6 \cdot 15,3 + 12 \cdot 17,5 + \dots + 72 \cdot 23,5 = 7187,4 .$$

Обратная матрица, которая получена на ЭВМ по стандартной программе обращения матриц, имеет значения:

$$(C_{ij}) = \begin{vmatrix} 3,5 & -0,5 & -0,583 & 0,0833 \\ -0,5 & 0,083 & 0,083 & -0,0139 \\ -0,583 & 0,083 & 0,117 & -0,0167 \\ 0,0833 & -0,0139 & 0,0167 & 0,0028 \end{vmatrix}$$

Коэффициенты регрессии находятся по формуле (63), например,

$$a_0 = 3,5 \cdot 222,2 - 0,5 \cdot 1329,3 - 0,583 \cdot 1194,4 + 0,0833 \cdot 7187,4 = + 15,26 .$$

Получаем регрессию:

$$y_p = \mathbf{15,26 - 0,617 x_1 + 0,715 x_2 + 0,112 x_1 x_2} . \quad (68)$$

Коэффициенты ковариации:

$$\rho(a_0, a_1) = -0,5 / \sqrt{3,5 \cdot 0,083} = -0,927;$$

$$\rho(a_0, a_2) = 0,083 / \sqrt{0,083 \cdot 0,117} = 0,84 \text{ и т. д.}$$

Как видно, коэффициенты ковариации близки к 1, влияние ошибок коэффициентов друг на друга очень велико.

Теперь сделаем все снова, введя новые переменные по формулам (67) - (см. рис .30):

$$x'_1 = \frac{x_1 - 0,5(9 + 3)}{0,5(6 - 3)} = \frac{x_1 - 6}{1,5} ;$$

$$x'_2 = \frac{x_2 - 0,5(8 + 2)}{0,5(4 - 2)} = x_2 - 5 .$$

В новых переменных значения уровней следующие:

$$x'_1 = -2; 0; +2 \quad \text{и} \quad x'_2 = -3; -1; +1; +3.$$

Таблица 21 экспериментальных данных теперь выглядит следующим образом.

Таблица 21

u	x'_{0u}	x'_{1u}	x'_{2u}	x'_{3u}	y_u
1	1	-2	-3	+6	15,3
2	1	-2	-1	+2	17,5
3	1	-2	+1	-2	19,8
4	1	-2	+3	-6	22,0
5	1	0	-3	0	14,4
6	1	0	-1	0	17,1
7	1	0	+1	0	20,0
8	1	0	+3	0	22,8
9	1	+2	-3	-6	13,3
10	1	+2	-1	-2	16,6
11	1	+2	+1	+2	19,9
12	1	+2	+3	+6	23,5

Новая нормальная матрица (ij) и матрица (jy) будут иметь вид:

$$(ij) = \begin{vmatrix} 12 & 0 & 0 & 0 \\ 0 & 32 & 0 & 0 \\ 0 & 0 & 60 & 0 \\ 0 & 0 & 0 & 160 \end{vmatrix}; \quad (jy) = \begin{vmatrix} 222,2 \\ -2,6 \\ 84,4 \\ 23,0 \end{vmatrix}.$$

Например, $(12) = 6 + 2 - 2 - 6 + 0 - 6 - 2 + 2 + 6 = 0$;

$$(1y) = -2 \cdot 15,3 - 2 \cdot 17,5 + \dots + 2 \cdot 23,5 = -2,6.$$

Следует обратить внимание на то, что в неортогональной матрице разница в минимальном и максимальном значениях элементов (соответственно 12 и 160) очень велика, а чем выше эта разница, тем меньше точность вычислений коэффициентов a_i . В ортогональной матрице все коэффициенты ковариации равны 0 и коэффициенты регрессии a_i вычисляются независимо друг от друга по формуле (66):

$$a_0 = 222,2 / 12 = 18,52; \quad a_1 = -2,6 / 32 = -0,081;$$

$$a_2 = 84,4 / 60 = 1,41; \quad a_3 = a_{12} = 23,0 / 160 = 0,144.$$

Получаем регрессию, которая, после подстановки старых координат, будет выглядеть следующим образом:

$$y_p = 18,52 - 0,081 \cdot \frac{x_1 - 6}{1,5} + 1,41 \cdot (x_2 - 5) + \frac{(x_1 - 6)(x_2 - 5)}{1,5}.$$

После преобразований получим:

$$y_p = 14,65 - 0,53 x_1 + 0,83 x_2 + 0,096 x_1 x_2 .$$

Как видно, это уравнение отличается от полученного ранее (68) и, надо полагать, в сторону большей точности, так как в предыдущем сказались округления чисел при вычислениях и ковариации коэффициентов.

Необходимо иметь в виду, что при обработке данных методом наименьших квадратов всегда полезно вводить новые координаты, даже если эксперимент неортогонален, так как этот прием позволяет уменьшать ковариации и с большей точностью вычислять коэффициенты регрессии.

7. 5. Планы 2^k

Особое место в теории планирования эксперимента занимают полные факторные эксперименты 2^k , в которых каждый из k факторов изменяется только на двух уровнях. Если перенести оси координат в центр эксперимента, то такой эксперимент автоматически становится ортогональным. Используя формулы (67), получим только два возможных уровня каждого из факторов в новых координатах: $x'_i = +1$ и $x'_i = -1$.

При дальнейшем изложении планов экспериментов 2^k мы будем использовать только приведенные к центру координаты, поэтому ради простоты штрихи при x_i будем опускать.

Чтобы оценить возможности экспериментов 2^k , рассмотрим несколько примеров. Пусть $k = 2$. Полный двухфакторный эксперимент содержит $N = 2^2 = 4$ узла. Координаты этих точек (в новых переменных) приведены в табл. 22 и на рис. 31 План эксперимента 2^2 приведен в таблице 23. Аналогично построен план эксперимента 2^3 , приведенный в табл. 24.

Таблица 22

x_1	x_2
+1	+1
+1	-1
-1	+1
-1	-1

Таблица 23

u	x_0	x_1	x_2	$x_1 x_2$	y_u
1	+	+	+	+	y_1
2	+	+	-	-	y_2
3	+	-	+	-	y_3
4	+	-	-	+	y_4

В этих планах содержится столбец с фиктивным переменным x_0 , которое всегда равно +1, столбцы с координатами факторов и взаимодействиями, а так же значения выходного параметра y_u . В планах обычно единицу не записывают (подразумевая ее), поскольку при расчетах важным оказывается только знак при ней.

Для эксперимента 2^2 уравнение

$$y_p = a_0 + a_1 x_1 + a_2 x_2 + a_{12} x_1 x_2 ,$$

содержащее 4 члена, оказывается адекватным ($m = N$).

Оно имеет свободный член, все линейные члены и парное взаимодействие. Поскольку план ортогонален, то коэффициенты регрессии легко вычисляются по формуле (66). Например,

$$a_0 = \frac{+y_1 + y_2 + y_3 + y_4}{4}; \quad a_{12} = \frac{+y_1 - y_2 - y_3 + y_4}{4}$$

Как видно, в числителе знаки столбца x_i приписываются к значениям y_u , а в знаменателе оказывается число N .

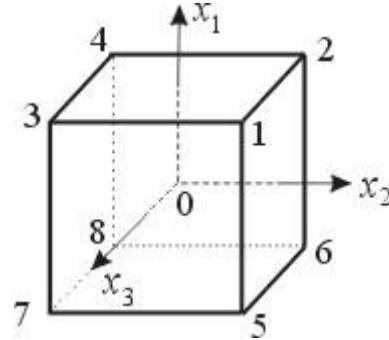
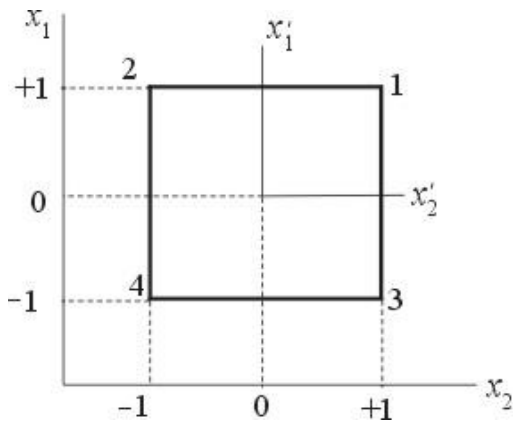


Рис. 31. План эксперимента 2^2 .

Рис. 32. План эксперимента 2^3 .

В эксперименте 2^3 опыты проводятся в вершинах куба (рис. 32), количество узловых точек $N = 2^3 = 8$. По его результатам можно построить адекватное уравнение, содержащее 8 членов:

$$y_p = a_0 + a_1 x_1 + a_2 x_2 + a_3 x_3 + a_{12} x_1 x_2 + a_{13} x_1 x_3 + a_{23} x_2 x_3 + a_{123} x_1 x_2 x_3.$$

. Эксперимент ортогонален, поэтому коэффициенты регрессии определяются по формуле (66), например,

$$a_{13} = \frac{y_1 - y_2 + y_3 - y_4 - y_5 + y_6 - y_7 + y_8}{8}$$

Таблица 24

u	x_0	x_1	x_2	x_3	$x_1 x_2$	$x_1 x_3$	$x_2 x_3$	$x_1 x_2 x_3$	y_u
1	+	+	+	+	+	+	+	+	y_1
2	+	+	+	-	+	-	-	-	y_2
3	+	+	-	+	-	+	-	-	y_3
4	+	+	-	-	-	-	+	+	y_4
5	+	-	+	+	-	-	+	-	y_5
6	+	-	+	-	-	+	-	+	y_6
7	+	-	-	+	+	-	-	+	y_7
8	+	-	-	-	+	+	+	-	y_8

В общем случае, полный факторный эксперимент 2^k содержит $N = 2^k$ экспериментальных точек. После приведения координат к центру эксперимента план становится ортогональным, что позволяет записать адекватное уравнение, содержащее $N = 2^k$ членов, в число которых входят свободный член, ли-

нейные члены и линейные взаимодействия парные, тройные, четверные и более высокого порядка; наивысший порядок взаимодействия равен k .

Таким образом, ПФЭ 2^k сравнительно дешев и очень информативен, поэтому широко применяется в экспериментальных исследованиях.

7. 6. Дробные реплики.

Доказано, что все тройные и более высокого порядка взаимодействия мало значимы, поэтому должны быть исключены из регрессии. Количество таких членов резко возрастает по мере увеличения числа факторов. Если для трехфакторного эксперимента незначимым является только один член – тройное взаимодействие, то при 5-ти факторах из $m = N = 2^5 = 32$ членов в число значимых должны быть включены только 16 членов, ровно половина от общего числа: 1 свободный, 5 линейных членов и 10 парных взаимодействий. При $k = 7$ из 128 членов лишь 29 могут быть значимыми: свободных – 1, линейных – 7 и взаимодействий – 21.

Возникает вопрос: нельзя ли так организовать (спланировать) эксперимент, чтобы определялись только значимые коэффициенты? При удалении членов из полного уравнения часть информации, как известно, неизбежно теряется. Необходимо так спланировать эксперимент, чтобы потерять информацию только о незначимых членах. В планируемом эксперименте количество опытов может быть значительно сокращено по сравнению с ПФЭ 2^k . Желательно, конечно, чтобы сокращенный эксперимент сохранял ортогональность плана.

Такие эксперименты строятся как дробная реплика от полного факторного эксперимента. Дробной репликой называется один из ортогональных блоков опытов, на которые можно разделить ПФЭ. Рассмотрим технику разделения ПФЭ на ортогональные блоки на трехфакторном эксперименте. Разобьем ПФЭ 2^3 , приведенный в табл. 24, на два блока: в первый включим опыты № 1, 4, 6 и 7. Во второй войдут все остальные. Оба блока характеризуются тем, что по первым двум факторам x_1 и x_2 осуществлен полный перебор уровней, и по ним каждый блок представляет собой ПФЭ $2^{k-1} = 2^2$. Уровни третьего фактора x_3 связаны с уровнями первых двух в первом блоке соотношением

$$x_3 = + x_1 x_2, \quad (69)$$

а во втором –

$$x_3 = - x_1 x_2. \quad (69 \text{ а})$$

Такие блоки сохраняют ортогональность и являются дробными полурепликами от полного факторного эксперимента.

Выражения (69) и (69 а) называются генерирующими соотношениями, которые и отличают один блок от другого. Генерирующее соотношение каждого блока обладают тем свойством, что из него можно получить другие генерирующие соотношения, принадлежащие этому блоку. Для этого необходимо левую и правую часть генерирующего соотношения умножить на любое переменное x_i ($i = 1, 2, 3$). При этом сомножитель x_i^2 можно не записывать, так

как $(+1)^2 = (-1)^2 = 1$. Например, умножив уравнение (69) соответственно на x_1, x_2, x_3 , получим новые генерирующие соотношения:

$$\begin{aligned}x_3 x_1 &= x_2, \\x_3 x_2 &= x_1, \\1 &= x_1 x_2 x_3.\end{aligned}$$

Последнее генерирующее соотношение имеет отдельное название - определяющий контраст, оно считается главным среди генерирующих соотношений. Если посмотреть на опыты № 1, 4, 6 и 7 табл. 24, образующие первую полуреплику от ПФЭ, то можно увидеть, что генерирующие соотношения указывают на столбцы, у которых знаки плана одинаковы. Определяющий контраст показывает, что в данном блоке знаки при x_0 и $x_1 x_2 x_3$ совпадают. (Соответственно, во втором блоке генерирующие соотношения указывают на столбцы, у которых знаки противоположны). Отсюда следует, что при вы-

числении коэффициентов регрессии по формуле (66) сумма $\sum_u^N x_{iu} y_u$ принадлежит сразу двум коэффициентам. Если построить систему нормальных уравнений и решить ее, то по 4-м опытам, входящим в блок, получим 4 уравнения:

$$\begin{aligned}a_0 + a_{123} &= (+ y_1 + y_2 + y_3 + y_4) / 4 ; \\a_1 + a_{23} &= (+ y_1 + y_2 - y_3 - y_4) / 4 ; \\a_2 + a_{13} &= (+ y_1 - y_2 + y_3 - y_4) / 4 ; \\a_3 + a_{12} &= (+ y_1 - y_2 - y_3 + y_4) / 4 .\end{aligned}$$

Только из первой суммы можно вычислить коэффициент a_0 достаточно точно, поскольку коэффициент a_{123} незначим. По другим суммам разделить линейные и парные эффекты не удастся. Полуреплика 2^{3-1} от ПФЭ 2^3 не эффективна, поскольку смешивает между собой значимые эффекты. Для их разделения необходимо проводить все 8 опытов ПФЭ. Однако, отсюда уже видно, что дробные реплики следует выбирать так, чтобы линейные и парные эффекты перемешивались с эффектами более высокого порядка. Проследить за этим можно по генерирующим соотношениям.

Рассмотрим 5-факторный эксперимент ($k = 5$). ПФЭ 2^5 содержит 32 опытные точки, а полуреплика от него 2^{5-1} в два раза дешевле. Чтобы получить ее, как и раньше, по четырем факторам сделаем полный перебор уровней, а пятый будем вычислять по одному из двух возможных генерирующих соотношений, например,

$$x_5 = + x_1 x_2 x_3 x_4 .$$

Умножая левую и правую часть на x_i ($i = 1, 2, \dots, 5$), получим другие генерирующие соотношения, которые укажут, какие пары коэффициентов регрессии вычисляются совместно:

$$\begin{aligned}1 &= + x_1 x_2 x_3 x_4 x_5 - \text{контраст} \\x_1 &= + x_2 x_3 x_4 x_5 \\x_2 &= + x_1 x_3 x_4 x_5 \text{ и т. д.}\end{aligned}$$

Видно, что свободный член регрессии вычисляется совместно с малозначимым пятерным взаимодействием, а линейные коэффициенты - с четверными. Далее

$$\begin{aligned} x_1 x_2 &= + x_3 x_4 x_5 & x_1 x_5 &= + x_2 x_3 x_4 \\ x_2 x_3 &= + x_1 x_4 x_5 & x_2 x_5 &= + x_1 x_3 x_4 \quad \text{и т. д.} \end{aligned}$$

Парные взаимодействия перемешаны с тройными и также разделяются. Таким образом, все нужные нам коэффициенты вычисляются с достаточной точностью, дробный факторный эксперимент 2^{5-1} эффективен, поэтому в любом случае нет смысла проводить ПФЭ 2^5 .

При семи факторах ($k = 7$) эффективной становится 1/4-реплика 2^{7-2} , содержащая 32 опытные точки вместо ПФЭ 2^7 , требующего 128 опытов. Такой эксперимент позволит вычислить все 29 значимых коэффициентов. По формуле (66) будем вычислять уже суммы четырех коэффициентов регрессии. Чтобы определить, какие эффекты будут перемешиваться, необходимо составить генерирующие соотношения. Пусть по первым пяти факторам x_1, x_2, x_3, x_4, x_5 осуществляется полный перебор уровней, а остальные два фактора будут вычисляться по таким генерирующим соотношениям:

$$\begin{aligned} x_7 &= + x_1 x_2 x_3 x_4 x_5 \\ x_6 &= + x_1 x_2 x_3 x_4 . \end{aligned}$$

Следует заметить, что начальные генерирующие соотношения исследователь назначает сам и в значительной мере произвольно, он может перемешивать между собой различные эффекты по собственному желанию. Видимо, всегда полезно линейные эффекты перемешивать с эффектами наивысшего порядка.

Остальные генерирующие соотношения находят по известной методике, однако рекомендуется начинать их формирование с определяющих контрастов, в нашем случае их будет два:

$$1 = + x_1 x_2 x_3 x_4 x_5 x_7 \quad \text{и} \quad 1 = + x_1 x_2 x_3 x_4 x_6 .$$

Если их перемножить, то получим третий контраст: $1 = + x_5 x_6 x_7$. Видно, что свободный член вычисляется совместно с шестерным, пятерным и тройным взаимодействиями.

Каждый из контрастов умножим на x_1 , получим:

$$x_1 = + x_2 x_3 x_4 x_5 x_7 = x_2 x_3 x_4 x_6 = x_1 x_5 x_6 x_7 .$$

Видно, что коэффициент при факторе x_1 , а по аналогии и остальные линейные коэффициенты, перемешиваются с коэффициентами при высоких взаимодействиях. Также можно показать, что парные коэффициенты вычисляются вместе с незначимыми коэффициентами.

Подводя итог, сформулируем правила проведения эксперимента для получения линейной модели

$$y = a_0 + a_1 x_1 + \dots + a_k x_k + a_{12} x_1 x_2 + a_{13} x_1 x_3 + \dots + a_{k-1, k} x_{k-1} x_k : \quad (70)$$

- при $k = 2, 3$ и 4 следует проводить ПФЭ 2^k ,
- при $k = 5$ и 6 эффективны дробные 1/2-реплики от ПФЭ 2^{k-1} ,
- при $k = 7$ и 8 эксперимент должен представлять собой дробную 1/4-реплику от ПФЭ 2^{k-2} .

В общей записи при $k = 2, \dots, 8$ проводится эксперимент 2^{k-l} , где $l = 0$ при $k \leq 4$, $l = 1$ при $k = 5$ и 6 и $l = 2$ при $k = 7$ и 8 .

7.7. Симплекс-планы. Планы Плакетта -Бермана.

Существует теорема, доказывающая, что при увеличении количества факторов k в эксперименте значимость парных взаимодействий ослабевает. При количестве факторов $k > 8$ можно считать незначимыми также парные взаимодействия. Модель эксперимента (70) упрощается и содержит только свободный и линейные члены:

$$y = a_0 + a_1 x_1 + a_2 x_2 + \dots + a_k x_k . \quad (71)$$

Количество членов регрессии сократилось до $(k + 1)$, желательно соответственно сократить число опытов для их нахождения.

Кстати, потребность в такой простой модели может возникнуть не обязательно при большом количестве факторов. Если по условию задачи точно известно, что парные взаимодействия незначимы или не интересны, то логично сразу начать изучение модели (71) и соответствующим образом под нее спланировать эксперимент, содержащий лишь требуемое количество опытов.

Самым дешевым для построения таких моделей является симплекс-план эксперимента. В математике симплексом называется геометрическая фигура в k - мерном пространстве, которая содержит $(k + 1)$ равноудаленных друг от друга вершин. Например, при двух факторах (рис. 33) симплексом является равносторонний треугольник (три вершины), в трехмерном пространстве – тетраэдр (четыре вершины). При большем количестве факторов геометрическое представление симплекса невозможно, но абстрактная конструкция симплекса в k -мерном пространстве легко представляется.

Поскольку симплекс-план эксперимента содержит ровно такое же количество экспериментальных точек, что и число членов в регрессии (71), то полученная модель будет адекватной. Симплекс-планы обладают только одним недостатком - они не ортогональны.

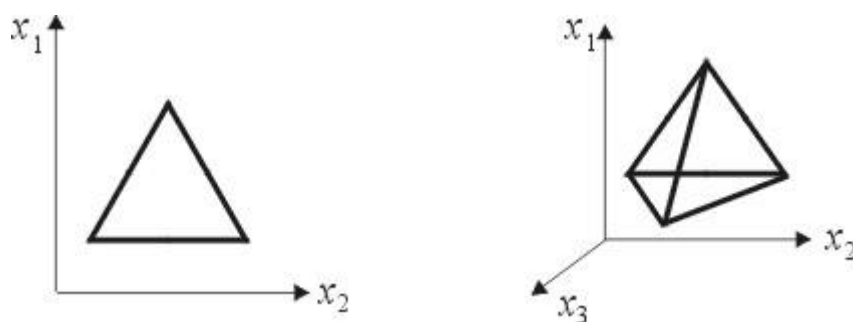


Рис. 33. Симплексы в двумерном и трехмерном пространстве.

Немного более дорогими, но ортогональными являются планы Плакетта-Бермана. Этими учеными было доказано, что k -факторные эксперименты ортогональны тогда, когда выполняются следующие условия:

- число экспериментальных точек N делится на 4,
- уровни всех факторов закодированы по формулам (67) и принимают только значения $+1$ и -1 ,

- по каждому фактору количество верхних равно количеству нижних уровней (и равно $N/2$),
- по каждой паре факторов в произведении количество плюсов равно количеству минусов.

Таблица 25.

k	N	Уровни фактора x_1 ($N - 1$ знаков)
9, 10, 11	12	++-++++--+-
16, 17, 18, 19	20	++--++++-+-+-----++-

Таблица 26.

u	x_0	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	y_u
1	+	+	-	+	-	-	-	+	+	+	-	19,15
2	+	+	+	-	+	-	-	-	+	+	+	34,44
3	+	-	+	+	-	+	-	-	-	+	+	85,08
4	+	+	-	+	+	-	+	-	-	-	+	92,88
5	+	+	+	-	+	+	-	+	-	-	-	90,91
6	+	+	+	+	-	+	+	-	+	-	-	51,76
7	+	-	+	+	+	-	+	+	-	+	-	101,33
8	+	-	-	+	+	+	-	+	+	-	+	101,34
9	+	-	-	-	+	+	+	-	+	+	-	98,62
10	+	+	-	-	-	+	+	+	-	+	+	87,85
11	+	-	+	-	-	-	+	+	+	-	+	84,49
12	+	-	-	-	-	-	-	-	-	-	-	89,89
a_i	a_0	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	-	-	-

Приведенные выше двухуровневые ПФЭ и дробные реплики от них удовлетворяют этим условиям, поэтому в этом смысле они могут рассматриваться как планы Плакетта-Бермана, точнее, при $k < 8$ эти планы совпадают. Однако при $k > 9$ помимо дробных реплик от ПФЭ существуют другие ортогональные эксперименты. Чтобы облегчить их построение, предложена вспомогательная таблица, две строки из которой приведены в табл. 26. Более полную таблицу можно найти в литературе, например, в [9, 10]. В табл. 26 находится строка с требуемым числом факторов k , по ней определяется число опытных точек N . По этой строке строится план эксперимента Плакетта-Бермана (табл. 27). Первые два столбца плана тривиальны. Третий столбец, соответствующий фактору x_1 , – это переписанная из табл.26 строка (за исключением последнего значения). Все остальные столбцы (также содержащие $N - 1$ значение) получаются из предыдущего путем переписывания знаков со сдвигом на один вниз, последнее значение предыдущего столбца становится в нем первым. В последней N -й строке всем факторам придается значение $- 1$. Для примера, табл.26 заполнена для 10-факторного экспери-

мента, для которого план Плакетта-Бермана содержит 12 экспериментальных точек. Можно убедиться в том, что заполненная таблица отвечает всем приведенным выше требованиям, и план будет ортогональным. Он применим при $k = 9, 10$ и 11 . При количестве факторов $k = 12, 13, 14$ или 15 ближайшим ортогональным планом является дробная реплика 2^4 , содержащая 16 опытов. Поэтому для такого количества факторов в табл. 25 строка отсутствует.

Коэффициенты регрессии нашего плана, естественно, также независимо друг от друга рассчитываются по формуле (66), например,

$$a_2 = (-19,15 + 34,44 + 85,08 - \dots - 89,89) / 12 = -3,47.$$

В табл.26 для коэффициентов a_i отведена отдельная строка.

7. 8. Дисперсионный анализ регрессионной модели.

Полученная регрессионная модель (70) или (71) содержит меньшее количество членов ($m + 1$), чем количество экспериментальных точек (N), следовательно, $N - (m + 1)$ член являются отброшенными. Каждый из отброшенных членов, хотя они и малозначимы, вынес из общей дисперсии эксперимента по одной степени свободы и некоторую сумму квадратов отклонений, что суммарно может представлять значительную величину. Среди оставленных в модели членов также могут встретиться такие, значимость которых невелика, и их следует исключить из модели. Если их оставить, то никаких последствий не будет, но загромождение модели свидетельствует о некоторой небрежности или неграмотности исследователя.

Напомним, что все оставшиеся члены регрессии также имеют по одной степени свободы, кроме свободного члена, который является числом. Отсюда следует, что, если регрессия содержит ($m + 1$) член, то число степеней свободы оставленных членов равно

$$f_{\text{ост}} = m, \quad (72)$$

а число степеней свободы отброшенных членов равно:

$$f_{\text{от}} = f_{\text{общ}} - f_{\text{ост}} = N - 1 - m. \quad (73)$$

Общее число степеней свободы $f_{\text{общ}} = N - 1$.

Общая сумма квадратов отклонений по эксперименту, проведенному в N точках, в регрессионном анализе вычисляется по формуле

$$SS_{\text{общ}} = \sum_u^N y_u^2.$$

Без вывода запишем сумму квадратов отклонений, принадлежащую каждому члену регрессии a_i (эту формулу надо запомнить):

$$SS_{a_i} = a_i \cdot \sum_u^N x_{iu} y_u. \quad (74)$$

Для ортогонального плана (первого порядка), с учетом формулы (66) и равенства

$$\sum_u^N x_{iu}x_{iu} = N ,$$

будем иметь:

$$SS_{a_i} = a_i \cdot \sum_u^N x_{iu}x_{iu} \cdot \frac{\sum_u^N x_{iu}y_u}{\sum_u^N x_{iu}x_{iu}} = Na_i^2 . \quad (74 a)$$

Все оставленные в регрессии члены имеют суммарную сумму квадратов отклонений равную

$$SS_{\text{ост}} = \sum_{i=0}^m a_i \cdot \sum_u^N x_{iu}y_u . \quad (75)$$

Следовательно, все отброшенные члены регрессии содержат сумму квадратов отклонений

$$SS_{\text{от}} = SS_{\text{общ}} - SS_{\text{ост}} = \sum_u^N y_u^2 - \sum_{i=0}^m a_i \cdot \sum_u^N x_{iu}y_u . \quad (76)$$

Таким образом, заключительной операцией при построении регрессионной модели является ее дисперсионный анализ. Решение о значимости или незначимости каких-то эффектов принимается на основе критерия Фишера, знаменателем которого является дисперсия ошибки воспроизводимости. Для ее нахождения требуется хотя бы в одной точке провести несколько опытов.

В планировании эксперимента принято двояким образом строить эксперимент. Во - первых, можно выполнять n_0 опытов действительно в одной точке, а именно в центре эксперимента, так как на эти опыты будут возложены дополнительные функции, на которых мы остановимся ниже. В этом случае имеем счетчик опытов в центральной точке $k = 1, 2, \dots, n_0$ и соответствующие экспериментальные значения y_{0k} . Число степеней свободы ошибки воспроизводимости

$$f_0 = n_0 - 1$$

и сумма квадратов отклонений

$$SS_0 = \sum_{k=1}^{n_0} y_{0k}^2 - \frac{y_{0\bullet}^2}{n_0} , \quad (77)$$

где сумма $y_{0\bullet} = \sum_{k=1}^{n_0} y_{0k}$.

Эксперимент, содержащий n_0 опытов центральной точке эксперимента, имеющей нулевые координаты по всем осям, называется центральным.

Вторая возможность определить дисперсию ошибки воспроизводимости - провести эксперимент, содержащий в каждой точке по n опытов. Это будет дороже, чем в первом случае, но надежнее, так как число степеней свободы возрастет. Поскольку количество экспериментальных точек равно N , то общее число степеней свободы

$$f_0 = N(n - 1)$$

и сумма квадратов отклонений

$$SS_0 = \sum_{u=1}^N \sum_{k=1}^n y_{uk}^2 - \sum_{u=1}^N \frac{y_{u\bullet}^2}{n}. \quad (78)$$

Если в каждой u -той точке проведено разное количество опытов n_u , то последние формулы будут иметь вид:

$$f_0 = \sum_u^N (n_u - 1),$$

$$SS_0 = \sum_{u=1}^N \sum_{k=1}^{n_u} y_{uk}^2 - \sum_{u=1}^N \frac{y_{u\bullet}^2}{n_u}.$$

Теперь возможен обычный дисперсионный анализ регрессионной модели по критерию Фишера. Если для некоторого коэффициента регрессии a_i выполняется неравенство

$$F_{a_i} = \frac{SS_{a_i} / 1}{SS_0 / f_0} \leq F(p; 1, f_0),$$

то с вероятностью p соответствующий член регрессии признается незначимым, и его следует исключить из регрессии. Точно также можно оценить суммарную значимость всех отброшенных членов, Если

$$F_{от} = \frac{SS_{от} / f_{от}}{SS_0 / f_0} \leq F(p; f_{от}; f_0),$$

то все отброшенные члены суммарно незначимы. При обратном неравенстве ситуация неоднозначна. С одной стороны, каждый из отброшенных членов регрессии незначим, но суммарный эффект их значим. Обычно решение принимается, исходя из конкретных особенностей задачи.

В качестве упражнения проведем дисперсионный анализ описанного выше двухфакторного эксперимента, приведенного в табл.19 и 21, который содержит $N = 12$ экспериментальных точек. В табл. 19 в каждой точке содержится по три опыта ($n = 3$), что позволяет определить ошибку воспроизводимости по формуле (78):

$$SS_0 = (15,1^2 + 15,3^2 + \dots + 23,5^2) - [(15,1 + 15,3 + 15,4)^2 + \dots + (23,5 + 23,6 + 23,5)^2] / 3 = 0,81.$$

Число степеней свободы ошибки воспроизводимости равно

$$f_0 = N(n - 1) = 12(3 - 1) = 24.$$

Оставлено в регрессии членов $(m + 1) = 4$ члена, следовательно $f_{ост} = 3$.

Общее число степеней свободы $f_{общ} = 12 - 1 = 11$

Число степеней отброшенных членов $f_{от} = 11 - 3 = 8$.

Общая сумма квадратов отклонений (по табл. 21) равна

$$SS_{общ} = 15,3^2 + 17,5^2 + \dots + 23,5^2 = 4236,7.$$

Сумму квадратов оставленных членов вычислим по формуле (75):

$$SS_{ост} = 18,52 \cdot 222,2 - 0,81 \cdot (-2,6) + 1,41 \cdot 84,4 + 0,144 \cdot 23,0 = 4233,2.$$

Сумма квадратов отброшенных членов

$$SS_{от} = 4236,7 - 4233,2 = 3,5 .$$

Критерий Фишера для отброшенных членов

$$F = \frac{3,5/8}{0,81/24} = 12,5 .$$

При табличном значении $F(0,95 ; 8 ; 24) = 3,1$ суммарно отброшенные члены значимы. Модель эксперимента

$$y_p = 14,65 - 0,53 x_1 + 0,83 x_2 + 0,096 x_1 x_2 ,$$

полученная в эксперименте (см. 7.4), не адекватна. Четырех оставленных в модели членов мало. Можно дописать еще какие-нибудь члены. Предельное количество членов равно 12, когда модель становится безусловно адекватной. Однако сравнение цифровых значений сумм квадратов отклонений оставшихся и отброшенных членов (4233,2 и 3,5) дает основание утверждать, что отброшенные члены вносят очень небольшой вклад в общую дисперсию. В большинстве практических задач можно считать полученную выше модель вполне удовлетворительной.

7.9. Центральный композиционный план.

Изложенные идеи лежат в основе центрального композиционного планирования эксперимента, то есть постепенного построения такого плана, который ставит своей целью построение адекватного уравнения при минимальном числе опытов. Первоначально предполагают, что модель процесса линейна, то есть содержит свободный и линейные члены и парные взаимодействия. Такой эксперимент содержит две серии опытов. В первой серии при небольшом количестве факторов ($k < 9$) эксперимент содержит 2^{k-1} узловых точек и описывается уравнением (70). При $k > 9$ применяются планы Плаккетта - Бермана или симплекс - планы для получения уравнения (71).

Вторую серию из n_0 опытов проводят в центре эксперимента, чтобы найти ошибку воспроизводимости. Относительно количества этих опытов можно высказать разные мнения. С одной стороны, чем их больше, тем лучше, так как с увеличением числа степеней свободы увеличивается точность оценки ошибки воспроизводимости и достоверность всего анализа. С другой - желательно получить достаточно достоверный результат при минимальном количестве опытов. Существуют таблицы для определения минимального значения n_0 в зависимости от типа эксперимента, обеспечивающего такую же достоверность при вычислении ошибки воспроизводимости, что и при вычислении коэффициентов регрессии, например,

$$\begin{aligned} \text{при } k = 2 \quad n_0 &= 4, \\ k = 3 \quad n_0 &= 6, \\ k = 4 \quad n_0 &= 6 \text{ и т. д.} \end{aligned}$$

На эксперименты в центральной точке возлагается дополнительная функция: оценка адекватности модели в этой точке. Идею такой операции проследим на однофакторном эксперименте. При $k = 1$ эксперимент содержит лишь две периферийных точки с координатами $+1$ и -1 , как показано на рис. 34.

Уравнение прямой

$$y_p = a_0 + a_1 x_1 \quad (79)$$

точно проходит через экспериментальные точки y_1 и y_2 , то есть адекватно в периферийных точках. В центральной точке с координатой $x = 0$ по уравнению имеем $y_{p0} = a_0$. Но значение y_0 , полученное как среднее по опытам, проведенным в этой точке, равно

$$y_0 = \frac{\sum_{u=1}^{n_0} y_{0u}}{n_0}.$$

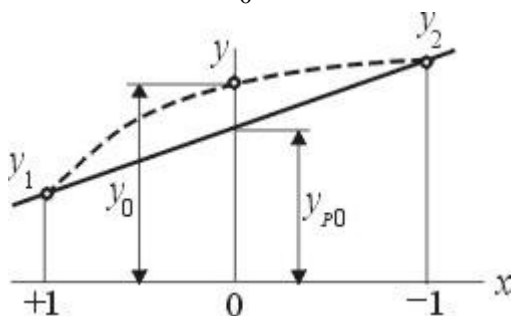


Рис. 34. Проверка адекватности модели в центре эксперимента.

Если эти значения совпадают, точнее находятся в одном доверительном интервале

$$a_0 = y_0 \pm \Delta y,$$

то уравнение прямой (79) адекватно также в центральной точке. Напомним, что доверительный интервал определяется по формуле $\Delta y = t_p \cdot S_x$, где коэффициент Стьюдента t_p находится по таблице П2 приложения при заданной вероятности p и числе опытов n_0 , а среднеквадратическое отклонение

$$S_x = \sqrt{SS_0 / f_0 / n_0}.$$

Если расчетное значение $y_{p0} = a_0$ не укладывается в заданный интервал, то с вероятностью p модель (прямая) не адекватна в центральной точке. Чтобы сделать ее адекватной и в этой точке, необходимо провести кривую второго порядка, показанную на рис. 34 штрихами, имеющую уравнение:

$$y_p = a_0 + a_1 x_1 + a_{11} x_1^2.$$

В общем случае, в k -факторном эксперименте модель (70) или (71) также необходимо проверить на адекватность в центральной точке, имеющей координаты $x_s = 0$ (счетчик числа факторов $s = 1, 2, \dots, k$). Расчетное по уравнению значение $y_{p0} = a_0$ должно укладываться в доверительный интервал Δy . Если этого не произошло, то линейная модель не адекватна в центральной точке. Необходимо перейти к модели второго порядка. Для ее получения не нужно проводить новый эксперимент, а к проведенным уже опытам следует добавить минимальное количество новых (третью серию) опытов. В целом по опытам всех трех серий необходимо записать модель, содержащую линейные, парные и квадратичные члены:

$$y = a_0 + a_1 x_1 + \dots + a_k x_k + a_{12} x_1 x_2 + a_{13} x_1 x_3 + \dots + a_{k-1,k} x_{k-1} x_k + a_{11} x_1^2 + a_{22} x_2^2 + \dots + a_{kk} x_k^2. \quad (80)$$

При этом необходимо в максимальной степени сохранить добытые результаты, то есть полученные ранее коэффициенты линейной модели (70) желательно уже не пересчитывать. Применяемые для построения квадратичной модели (80) планы эксперименты называются планами второго порядка. Они позволяют пересчитывать только один коэффициент – a_0 .

Для построения плана второго порядка применяется стратегия, которая называется центральным композиционным планированием. Опыты всех трех серий komponуются вокруг центральной точки эксперимента, поэтому план является центральным. Он выполняется по пути постепенного усложнения модели от простой линейной (70) к более сложной квадратичной (80) регрессии при постепенным увеличением количества затраченных опытов. Такая стратегия является самой экономной из всех возможных.

7. 10. Планирование второго порядка.

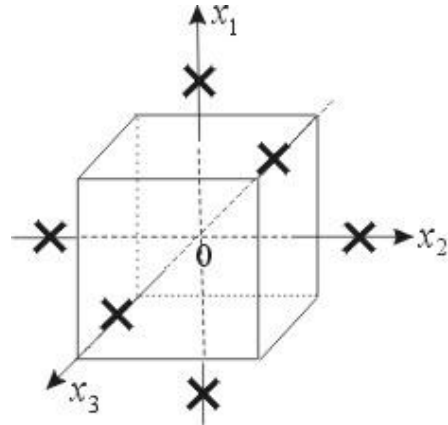
Каждый план второго порядка имеет свои достоинства и недостатки и не может быть оптимальным сразу по нескольким критериям. Только линейные планы были одновременно ортогональными, рототабельными и D - оптимальными. Особенности построения и свойства планов второго порядка можно найти в специальной литературе, например, в [8].

Особое место среди планов второго порядка занимают ортогональные и рототабельные планы, так как содержат минимальное и строго определенное количество опытов третьей серии, добавленных к опытам, затраченным при построении линейной модели. Такие эксперименты уже не D - оптимальны, рототабельные эксперименты не ортогональны, а ортогональные – не обладают рототабельностью. Напомним, что все свойства планов определяются требованиями, предъявляемыми к матрице ошибок (обращенной матрице), с помощью которой решается система линейных уравнений в методе наименьших квадратов. В ортогональных планах все члены матрицы ошибок, не лежащие на главной диагонали, равны нулю, что позволяет вычислять все коэффициенты регрессии независимо друг от друга с той ошибкой, которая свойственна каждому из них в отдельности. Но общая ошибка модели (80) при этом не минимальна. Рототабельные планы минимизируют общую ошибку, поэтому предпочтительнее ортогональных, но они не ортогональны. D - оптимальные планы дороги, не ортогональны и не рототабельны, хотя минимизируют ошибку при вычислении линейных членов модели (80) при неизвестной ошибке взаимодействий и квадратичных членов.

Опыты третьей серии ортогональных и рототабельных планов выполняются в так называемых *звездных* точках плана, расположенных на каждой оси на расстоянии звездного плеча α от центральной точки в положительном и отрицательном направлении. Таким образом, в k - факторном эксперименте на k осях расположится $2k$ звездных точек. В каждой из звездных точек необходимо провести по одному опыту, следовательно, третья серия состоит из $2k$ опытов. На рис. 35 в качестве примера приведены звездные точки трехфакторного эксперимента (6 штук), координаты их - в табл. 27.

Таблица 27

u	x_1	x_2	x_3
1	$+\alpha$	0	0
2	$-\alpha$	0	0
3	0	$+\alpha$	0
4	0	$-\alpha$	0
5	0	0	$+\alpha$
6	0	0	$-\alpha$

Рис. 35. Опыты в звездных точках эксперимента при $k = 3$.

План становится ортогональным, если так подобрано звездное плечо α , что нормальная матрица в методе наименьших квадратов вырождается в диагональную, следовательно, $\sum_u^N x_{iu}x_{ju} = 0$ при $i \neq j$. Это происходит при

следующих значениях звездного плеча α :

$$\alpha = 1,0 \text{ при } k=2,$$

$$\alpha = 1,21 \text{ при } k=3,$$

$$\alpha = 1,41 \text{ при } k=4 \text{ и т. д.}$$

Для такого эксперимента полученные ранее коэффициенты регрессии при линейных членах и парных взаимодействиях пересчитывать не надо. Все три серии опытов участвуют в расчете новых коэффициентов $a_{11}, a_{22}, \dots, a_{kk}$, которые рассчитываются по формуле (66), справедливой для ортогонального плана. Пересчитывается также один коэффициент a_0 (по формуле (66) и также по трем сериям опытов)..

Рототабельный план второго порядка более предпочтителен, чем ортогональный, так как помимо отмеченных достоинств (минимальная погрешность при вычислении y и почти равные ошибки в направлении каждой из осей) он почти ортогонален. Звездное плечо вычисляется по формуле

$$\alpha = 2^{(k-1)/4}.$$

Получаем:

$$\alpha = 1,41 \text{ при } k=2,$$

$$\alpha = 1,68 \text{ при } k=3,$$

$$\alpha = 2,0 \text{ при } k=4 \text{ и т. д.}$$

Можно сравнить с ортогональным планом.

После проведения третьей серии опытов по рототабельному плану также нет необходимости пересчитывать коэффициенты при линейных членах и парных взаимодействиях, полученные из первой серии опытов. В этом смысле рототабельный план почти ортогонален. Рассчитываются только новые коэффициенты $a_{11}, a_{22}, \dots, a_{kk}$ и пересчитывается коэффициент a_0 . По этим коэффициентам нормальная матрица не ортогональна, приходится ре-

шать систему из $(k + 1)$ уравнений с $(k + 1)$ неизвестными. Причем при составлении нормальных уравнений должны участвовать опыты всех трех серий. Серию опытов в центральной точке заменяют одним опытом со средним значением y_0 . Таким образом, в каждой строке нормального уравнения суммирование осуществляется по $N = (2^{k-l} + 1 + 2k)$ опытам.

7.11. Пример построения рототабельного плана.

Осуществим построение регрессионной модели четырехфакторного эксперимента ($k = 4$) при минимуме количества проведенных опытов. Применим стратегию центрального композиционного планирования эксперимента. В табл. 28 задано соотношение между старыми (физическими) и новыми (кодированными) переменными, рассчитанными по формуле (67).

Таблица 28.

	x_1	x_2	x_3	x_4
В центре экс-та ($x_i' = 0$)	0,87	40,0	1,0	250
Интервал Δx_i	0,15	5,0	0,25	50
Верхний ур-нь ($x_i' = +1$)	1,02	45,0	1,25	300
Нижний ур-нь ($x_i' = -1$)	0,72	35,0	0,75	200

Таблица 29

u	x_0'	x_1'	x_2'	x_3'	x_4'	y_u	y_u^2
1	+	+	+	+	+	21,5	462,2
2	+	+	-	+	-	9,0	81,0
3	+	-	-	+	+	10,2	104,0
4	+	-	+	+	-	40,2	1616,0
5	+	+	+	-	-	26,4	697,0
6	+	+	-	-	+	42,2	1780,8
7	+	-	-	-	-	13,0	169,0
8	+	-	+	-	+	13,8	190,4
9	+	-	+	+	+	17,7	313,3
10	+	+	-	+	+	29,3	858,5
11	+	+	+	-	+	16,7	278,9
12	+	+	+	+	-	32,8	1075,8
13	+	-	-	-	+	24,0	576,0
14	+	-	-	+	-	1,2	1,4
15	+	-	+	-	-	34,6	1197,2
16	+	+	-	-	-	20,2	408,0
a_i	22,05	2,71	3,41	-1,81	-0,12	352,8	9809,5

Первая серия опытов представляет собой ПФЭ 2^4 , состоящий из 16 опытов (по одному опыту в каждой точке). В табл.29 приведены план эксперимента и опытные данные. Эксперимент проведен в рандомизированном порядке, что всегда полезно, чтобы уменьшить возможность непредвиденных случайностей. Для удобства ручных расчетов в последнем столбце вычислены квадраты всех выходных значений. Последняя строка посвящена коэффициентам регрессии, вычисленным по формуле (66).

В результате такого эксперимента получаем регрессию:

$$y_p = 22,05 + 2,71 x_1 + 3,41 x_2 - 1,81 x_3 - 0,12 x_4 - 3,83 x_1 x_2 + 0,2 x_1 x_3 + 2,79 x_1 x_4 + 4,4 x_2 x_3 - 7,91 x_2 x_4 - 0,4 x_3 x_4. \quad (81)$$

Имеем 11 членов уравнения при 16 опытах, следовательно, отброшено 5 членов: четыре тройных и одно четверное взаимодействия. Они “по определению” незначимы, но в этом можно убедиться, подсчитав сумму квадратов, принадлежащую этим членам:

$$SS_{от} = SS_{общ} - SS_{ост} = \sum_{u=1}^{16} y_u^2 - N \cdot \sum_{i=0}^{10} a_i^2 =$$

$$(462,2 + 81,0 + \dots + 408,0) - 16 \cdot (22,05^2 + 2,71^2 + \dots + 0,4^2) = 0,63.$$

При 5 степенях свободы вклад всех отброшенных членов в общую дисперсию очень мал, однако для его оценки по критерию Фишера необходимо иметь ошибку воспроизводимости эксперимента. Для этого проводим вторую серию из $n_0 = 6$ опытов в центре. В табл.30 приведен план этой части эксперимента.

Таблица 30

u	x_1	x_2	x_3	x_4	y_{0u}
1	0	0	0	0	12,5
2	0	0	0	0	12,9
3	0	0	0	0	11,5
4	0	0	0	0	12,0
5	0	0	0	0	13,0
6	0	0	0	0	13,0

Вычислим суммы, среднее и доверительный интервал:

$$\sum_{u=1}^6 y_{0u} = 12,5 + \dots + 13,0 = 74,9;$$

$$y_{0\text{cp}} = y_0 = 74,9 / 6 = 12,48;$$

$$SS_0 = \sum_{u=1}^6 y_{0u}^2 - \frac{1}{n_0} \cdot \left(\sum_{u=1}^6 y_{0u} \right)^2 = (12,5^2 + \dots + 13,0^2) - 74,9^2 / 6 = 1,9;$$

$$\Delta y = t_p \cdot \sqrt{SS_0 / (n_0 - 1) / n_0} = 2,57 \cdot \sqrt{1,9 / 5 / 6} = 0,65.$$

Критерий Фишера для отброшенных членов

$$F_{от} = \frac{0,63/5}{1,9/5} = 0,4,$$

что меньше табличного значения $F_T(0,95; 5; 5) = 5,1$, поэтому все отброшенные члены не значимы.

В регрессии (81) могут быть мало значимыми члены с наименьшими значениями коэффициентов, например, a_4x_4 , $a_{13}x_1x_3$ и $a_{14}x_1x_4$. Проверим их по критерию Фишера, рассчитав по формуле (74 а) сумму квадратов отклонений $SS_{ai} = a_i^2 \cdot N$ и вклад каждого из них в общую дисперсию (при 1 степени свободы).

При табличном значении $F_T(0,95; 1; 5) = 6,6$ для указанных членов регрессии критерии Фишера будут следующими:

$$F_{a_4} = \frac{16 \cdot 0,12^2 / 1}{1,9 / 5} = 0,6 ;$$

$$F_{a_{13}} = \frac{16 \cdot 0,2^2 / 1}{1,9 / 5} = 1,68 ;$$

$$F_{a_{34}} = \frac{16 \cdot 0,4^2 / 1}{1,9 / 5} = 6,7 .$$

Члены регрессии с коэффициентами a_4 и a_{13} незначимы, все остальные будут значимыми. Член с коэффициентом a_{14} находится на пределе значимости, поэтому при желании мы могли бы его также отбросить.

В центре эксперимента значение по уравнению y_{p0} не совпадает с экспериментальным значением $y_0 = 12,48 \pm 0,65$, то есть модель в этой точке не адекватна. Необходимо добавить $2k = 8$ опытов в звездных точках и достроить модель до квадратичной. Звездное плечо $\alpha = 2^{4/4} = 2$. В таблицу 28 добавляется еще две строки :

Добавка к таблице 28.

	x_1	x_2	x_3	x_4
Уровень $x_i' = + \alpha$	1,17	50,0	1,50	350
Уровень $x_i' = - \alpha$	0,57	30,0	0,50	150

Экспериментальные данные в звездных точках приведены в табл. 31.

Рототабельный план неортогонален лишь относительно коэффициентов регрессии a_0 , a_{11} , a_{22} , a_{33} и a_{44} , следовательно в системе нормальных уравнений останутся пять уравнений с пятью неизвестными коэффициентами. Остальные коэффициенты, вычисленные ранее, пересчитывать не надо. В данной системе суммирование будет осуществляться по 16 опытам эксперимента 2^4 , 1 опыту в центре со средним значением y_0 и 8 опытам в звездных точках, итого по $N = 25$. Нормальная система уравнений будет иметь вид:

$$\begin{aligned}
a_0 \cdot \sum_u^{25} x_{0u} x_{0u} + a_{11} \cdot \sum_u^{25} x_{0u} x_{1u}^2 + \dots + a_{44} \cdot \sum_u^{25} x_{0u} x_{4u}^2 &= \sum_u^{25} x_{0u} y_u \\
a_0 \cdot \sum_u^{25} x_{1u}^2 x_{0u} + a_{11} \cdot \sum_u^{25} x_{1u}^2 x_{1u}^2 + \dots + a_{44} \cdot \sum_u^{25} x_{1u}^2 x_{4u}^2 &= \sum_u^{25} x_{1u}^2 y_u \\
a_0 \cdot \sum_u^{25} x_{2u}^2 x_{0u} + a_{11} \cdot \sum_u^{25} x_{2u}^2 x_{1u}^2 + \dots + a_{44} \cdot \sum_u^{25} x_{2u}^2 x_{4u}^2 &= \sum_u^{25} x_{2u}^2 y_u \\
a_0 \cdot \sum_u^{25} x_{31u}^2 x_{0u} + a_{11} \cdot \sum_u^{25} x_{3u}^2 x_{1u}^2 + \dots + a_{44} \cdot \sum_u^{25} x_{3u}^2 x_{4u}^2 &= \sum_u^{25} x_{3u}^2 y_u \\
a_0 \cdot \sum_u^{25} x_{41u}^2 x_{0u} + a_{11} \cdot \sum_u^{25} x_{4u}^2 x_{1u}^2 + \dots + a_{44} \cdot \sum_u^{25} x_{4u}^2 x_{4u}^2 &= \sum_u^{25} x_{4u}^2 y_u.
\end{aligned}$$

Таблица 31

u	x_0'	x_1'	x_2'	x_3'	x_4'	y_u	y_u^2
1	+	+2	0	0	0	29,4	864,4
2	+	-2	0	0	0	18,3	334,9
3	+	0	+2	0	0	19,3	372,5
4	+	0	-2	0	0	5,7	32,5
5	+	0	0	+2	0	27,7	767,3
6	+	0	0	-2	0	34,9	1218,0
7	+	0	0	0	+2	12,3	151,3
8	+	0	0	0	-2	12,7	161,3
a_{ii}	12,5	22,2	27,2	-14,4	-0,8	160,3	3902,2

После подсчета всех сумм система выглядит следующим образом:

$$\begin{aligned}
25a_0 + 24a_{11} + 24a_{22} + 24a_{33} + 24a_{44} &= 525,6 \\
24a_0 + 48a_{11} + 16a_{22} + 16a_{33} + 16a_{44} &= 543,6 \\
24a_0 + 16a_{11} + 48a_{22} + 16a_{33} + 16a_{44} &= 452,8 \\
24a_0 + 16a_{11} + 16a_{22} + 48a_{33} + 16a_{44} &= 603,2 \\
24a_0 + 16a_{11} + 16a_{22} + 16a_{33} + 48a_{44} &= 452,8.
\end{aligned}$$

Технику суммирования проследим на конкретных примерах:

$$\begin{aligned}
\sum_u^{25} x_{0u} x_{0u} &= [16 \text{ (по табл. 30)} + 1 \text{ (в центре)} + 8 \text{ (по табл. 32)}] = 25; \\
\sum_u^{25} x_{1u}^2 x_{2u}^2 &= [16 \text{ (по табл. 30)} + 0 \text{ (в центре)} + 0 \text{ (по табл. 32)}] = 16; \\
\sum_u^{25} x_{1u}^2 x_{1u}^2 &= [16 \text{ (по табл. 30)} + 0 \text{ (в центре)} + 32 \text{ (по табл. 32)}] = 48.
\end{aligned}$$

$$\sum_u^{25} x_{3u}^2 y_u = [352,8 \text{ (по табл. 30)} + 0 \text{ (в центре)} + \\ + (4 \cdot 27,7 + 4 \cdot 34,9) - \text{ по табл. 32}] = 603,2$$

Решив систему из 5 уравнений, получаем:

$$a_0 = 12,5; \quad a_{11} = 3,17; \quad a_{22} = 0,34; \quad a_{33} = 5,03; \quad a_{44} = 0,34.$$

Проверим значимость члена с коэффициентом 0,34 (таковых два):

$$SS_{a_{22}} = a_{22} \cdot \sum_u^{25} x_{2u}^2 y_u = 0,34(352,8 + 4 \cdot 19,3 + 4 \cdot 5,7) = 153,9;$$

$$F_{a_{22}} = \frac{153,9 / 1}{1,9 / 5} = 405.$$

Видно, что оба подозреваемые члена высоко значимы. Окончательное уравнение, которое будет адекватным как в периферийных, так и в центральной точках, будет следующим:

$$y_p = 12,5 + 2,71 x_1 + 3,41 x_2 - 1,81 x_3 - 3,83 x_1 x_2 + 2,79 x_1 x_4 + 4,4 x_2 x_3 - \\ - 7,91 x_2 x_4 - 0,4 x_3 x_4 + 3,17 x_1^2 + 0,34 x_2^2 + 5,03 x_3^2 + 0,34 x_4^2.$$

Оно получено самым экономным способом (30 опытов), не содержит мало-значимых членов. Проверку суммарной значимости всех отброшенных членов обычно не делают, если каждый из них в отдельности незначим. К тому же это довольно трудоемкая операция. Мы выполним ее, чтобы продемонстрировать методику практических расчетов необходимых сумм по данным приведенных экспериментальных таблиц.

Количество опытов $N = 25$, количество оставленных членов $(m + 1) = 12$, следовательно, отброшено 13 членов и $f_{от} = 13$.

Сумма квадратов отброшенных членов

$$SS_{om} = \sum_u^{25} y_u^2 - \sum_{i=0}^{11} a_i \sum_u^{25} x_{iu} y_u = \\ = (9809,5 + 12,5^2 + 3902,2) - [12,5 \cdot 525,6 + 16 \cdot (2,71^2 + 3,41^2 + 1,81^2 + 3,83^2 + \\ + 4,4^2 + 0,4^2 + 2,79^2 + 7,91^2) + 2,71 \cdot 2 \cdot (29,4 - 18,3) + 3,41 \cdot 2 \cdot (19,3 - 5,7) - \\ - 1,81 \cdot 2 \cdot (27,7 - 34,9) + 3,17 (352,8 + 4 \cdot 29,4 + 4 \cdot 18,3) + \\ + 0,34 (352,8 + 4 \cdot 19,3 + 4 \cdot 5,7) + 5,03 (352,8 + 4 \cdot 27,7 + 4 \cdot 34,9) + \\ + 0,34 (352,8 + 4 \cdot 12,3 + 4 \cdot 12,7)] = 22,3.$$

Критерий Фишера для всех отброшенных членов

$$F = \frac{22,3 / 13}{1,9 / 5} = 4,51$$

меньше табличного значения $F(0,95; 13; 5) = 4,7$, следовательно все отброшенные члены суммарно незначимы.

7.12 . Симплекс-решетчатый план второго порядка.

В качестве примера специализированного плана приведем симплекс - решетчатый план второго порядка, который применяется в металлургической технологии для изучения тройных систем сплавов и шлаков. Такими экспериментами удобно исследовать, например, жидкотекучесть (y) шлака в зависимости от содержания в нем компонентов SiO_2 , CaO и Al_2O_3 (соответственно x_1 , x_2 и x_3). Эксперимент проводится в 16 точках двухфакторного симплекса, показанного на рис. 36, причем в центральной точке № 16 проводится n_0 опытов для определения ошибки воспроизводимости, а в остальных точках - по одному опыту. Видно, что на каждой стороне симплекса находится по 5 экспериментальных точек, и кроме того имеются точки № 13,14 и 15, расположенные на расстоянии $1/3$ отрезка, проходящего от центра по направлениям к вершинам симплекса. Такой эксперимент обладает *D-оптимальными* свойствами и позволяет записать и исследовать квадратичную модель следующего вида:

$$y_p = a_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + b_{12} x_1 x_2 + b_{13} x_1 x_3 + b_{23} x_2 x_3 + c_{12} x_1 x_2 (x_1 - x_2) + c_{13} x_1 x_3 (x_1 - x_3) + c_{23} x_2 x_3 (x_2 - x_3) + d_{12} x_1 x_2 (x_1 - x_2)^2 + d_{13} x_1 x_3 (x_1 - x_3)^2 + d_{23} x_2 x_3 (x_2 - x_3)^2 + e_1 x_1^2 + e_2 x_2^2 + e_3 x_3^2 .$$

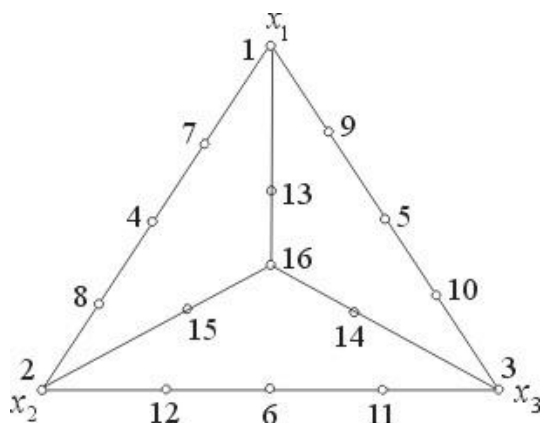


Рис.36. Симплекс-решетчатый план

План не ортогональный, поэтому для нахождения 16 неизвестных коэффициентов придется решать систему нормальных линейных уравнений (61). Модель будет адекватной, если в центральной точке расчетное значение $y_{0p} = a_0$ находится в доверительном интервале экспериментального значения $y_0 \pm t_p S$. Дисперсионный анализ значимости каждого члена регрессии осуществляется точно так же, как в приведенном рототабельном плане.

7. 13. Отсеивающие эксперименты.

Метод случайного баланса

Как известно, на начальных стадиях исследования сложных объектов важно выделить существенные параметры процесса и отсеять несущественные, и этими вопросами занимается дисперсионный анализ. Однако, когда все факторы задачи количественные более дешевым может оказаться отсеивающий

эксперимент, основанный на идеях регрессионного анализа, так как для его проведения требуется любое количество опытов N , в том числе меньше числа факторов k . Все остальные изученные нами планы экспериментов требовали, чтобы число опытов было больше числа факторов: $N \geq k$. В этом смысле отсеивающие эксперименты являются самыми дешевыми. Метод отсеивания несущественных факторов называется также методом случайного баланса. План такого эксперимента требует выполнения лишь следующего простого условия: каждый фактор обязан иметь только два уровня “+” и “-” и количество верхних уровней каждого фактора должно примерно равняться количеству нижних уровней и по произведению любых двух факторов количество “+” и “-” должно быть также примерно одинаковым. В какой-то мере эти требования перекликаются с требованиями к ортогональности эксперимента, сформулированными Плакеттом и Берманом. Однако там еще требовалось, чтобы число опытов было кратно 4 и, конечно, $N \geq k$. В общем же случае, когда эти условия не выполняются, эксперименты метода случайного баланса неортогональны, но методика их обработки близка к той, что используется при ортогональном планировании.

Рассмотрим метод случайного баланса на конкретном примере. В табл. 32 приведен отсеивающий эксперимент, содержащий 12 опытов при 8 факторах: $N = 12$, $k = 8$. Уровни каждого их факторов выбраны случайным образом, но в соответствии с изложенными требованиями.

Таблица 32.

u	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	y_u
1	+	-	+	+	+	-	+	+	100
2	-	+	+	-	+	-	-	+	35
3	+	-	-	-	-	-	-	-	90
4	-	-	+	-	-	+	+	+	65
5	-	+	+	+	-	-	-	-	45
6	-	-	-	-	+	+	+	-	40
7	-	-	+	+	+	+	-	-	65
8	+	+	+	-	-	+	+	-	75
9	+	+	-	-	+	-	-	+	100
10	-	+	-	+	-	-	+	-	50
11	+	+	-	+	+	+	+	+	105
12	+	-	-	+	-	+	-	+	90

Наиболее значимым считается тот линейный фактор x_i или парное взаимодействие $x_i x_j$ регрессии, для которого наивысшее значение принимает y_u (как характеристика эффекта)

$$\text{или } A_{ij} = \left| \frac{\sum_u^{N^+} x_{iu} x_{ju} y_u}{N^+} - \frac{\sum_u^{N^-} x_{iu} x_{ju} y_u}{N^-} \right|,$$

где N^+ и N^- – количество членов со знаками “+” и “–” соответственно. Суммирование производится отдельно по опытам на верхнем и нижнем уровнях.

Например, по данным табл. 32

$$A_1 = \frac{100 + 90 + 75 + 100 + 105 + 90}{6} - \frac{35 + 65 + 45 + 40 + 65 + 50}{6} = 43,5$$

$$A_{13} = \frac{100 + 40 + 75 + 50}{4} - \frac{35 + 90 + 65 + 45 + 65 + 100 + 105 + 90}{8} = 8,1.$$

Делаем вывод: фактор x_1 более значим, чем взаимодействие $x_1 x_3$. В методе случайного баланса ошибку воспроизводимости не определяют, а уровень значимости A_{min} назначают, исходя из каких-то технологических соображений. Все эффекты со значением A_i и A_{ij} меньше A_{min} признаются незначимыми. К примеру, если в нашем случае назначить $A_{min} = 10$, то фактор x_1 следует считать значимым, а взаимодействие $x_1 x_3$ не значимым.

Как видно, метод случайного баланса имеет элемент произвола и потому не очень точен. Это плата за дешевизну эксперимента.

7. 14. Экстремальные эксперименты. Крутое восхождение

Во многих технологических задачах требуется найти такие условия, то есть такой набор входных параметров x_i ($i = 1, 2, \dots, k$), а в терминологии планирования эксперимента – такую точку факторного пространства, в которой выходной параметр y приобретает экстремальное (максимальное или минимальное) значение. Для определенности, мы будем говорить о поиске максимума. К поиску экстремума сводится большинство задач оптимального управления, в том числе автоматического и автоматизированного. Металлургические задачи “состав – свойства” также ставятся с целью нахождения экстремальных характеристик. В них требуется определить химический состав и параметры технологической обработки сплава, при которых его эксплуатационные характеристики максимальны.

Студентам известно, что когда модель процесса (уравнение) имеется и область поиска максимума не ограничена, то искомая точка находится просто: необходимо взять частные производные от этого уравнения по всем входным параметрам и приравнять их нулю. Если область поиска максимума ограничена некоторыми плоскостями (линейными функциями), и сама модель представляет собой линейную функцию, то поиск точки максимума осуществляют методами линейного программирования. При нелинейных

границах области или нелинейной функции существует несколько методов нелинейного и динамического программирования.

В большинстве технологических, экономических и производственных задач мы имеем дело с иной ситуацией: модель процесса не известна, а технологические (входные) параметры ограничены сверху и снизу некоторыми конструктивными или технологическими значениями, и не могут изменяться в широких пределах. Технологические и экономические задачи, как правило, логически подразумевают существование и наличие одного максимума, который и требуется отыскать. Не требуется, как в математике, доказывать существование и единственность решения. Это существенное облегчение для такого класса задач.

Необходимо разработать стратегию поиска максимума экспериментальным путем, затратив на это минимальное количество опытов и времени. Причем, сам максимум в ряде случаев не так уж и важен, можно ограничиться поиском области вблизи максимума. На поиски точного максимума в этой области необходимо затратить дополнительные средства, и необходимо обосновать целесообразность таких затрат.

Математическую сущность задачи поиска максимума удобно рассмотреть на двухфакторной модели, а затем распространить на многофакторные эксперименты. На рис. 37 двумерная модель представлена графически в виде “горки” линиями равного уровня (как на географической карте) с одной вершиной в точке M . Линии равного уровня представляют некоторые замкнутые кривые в виде криволинейных эллипсов, что характерно для технико-экономических задач. В математике функции, описывающие такие поверхности, относятся к уравнениям эллиптического типа. Если в задаче модель более сложна, например, имеет несколько максимумов или максимум находится где-то в другой области, то это, скорее всего, свидетельствует о неправильной или недостаточно четкой технологической проработке и постановке задачи. Может быть, следует разбить задачу на две.

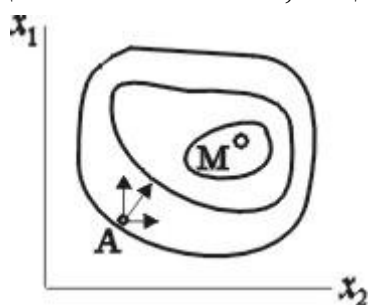


Рис. 37. Метод крутого восхождения для двухфакторной модели.

Поиска максимума в технико-экономических задачах распадается на два этапа. Сначала необходимо с помощью дешевого эксперимента наметить стратегию движения от любой точки A к точке максимума M , и осуществить движение по намеченной стратегии. При движении мы попадаем в область максимума, после чего следует решить, окупятся ли дополнительные затраты на уточнение точки максимума в этой области. И, если затраты оправданы, то в области максимума можно провести планируемый центральный композиционный эксперимент второго порядка, по типу приведенного выше

рототабельного, затем от полученного уравнения взять частные производные по всем k факторам и приравнять их нулю. Полученная система k уравнений с k неизвестными даст точные значения факторов в точке максимума. Вторая часть поиска максимума дорога и проводится редко, поэтому наибольший практический интерес представляет первая часть, выводящая в область максимума.

Известно, что из произвольной точки A путь к вершине M будет кратчайшим, если совершать движение по самому крутому направлению, поэтому первая часть стратегии называется крутым восхождением. Это направление соответствует градиенту функции, описывающей горку в окрестности точки A . Градиент функции $y = f(x_1, x_2, \dots, x_k)$ является вектором и определяется выражением:

$$\overline{grad(y)} = \overline{\Delta y} = \frac{\partial y}{\partial x_1} \cdot \bar{i}_1 + \frac{\partial y}{\partial x_2} \cdot \bar{i}_2 + \dots + \frac{\partial y}{\partial x_k} \cdot \bar{i}_k. \quad (82)$$

Единичные векторы \bar{i}_s по направлению совпадают с направлением каждой s -й оси ($s = 1, 2, \dots, k$). Модуль их, то есть величина шага Δx_s перемещения по s -й оси, назначается технологом произвольно, исходя из возможностей оборудования и технологических соображений. Чтобы найти частные производные $\partial y / \partial x_s$, необходимо в окрестности точки A провести какой-нибудь дешевый эксперимент, например, ДФЭ 2^{k-1} (или симплекс-план, а при $k > 9$ - план Плакетта-Бермана). Точка A будет центром такого эксперимента. Из этого эксперимента будет получено уравнение

$$y = a_0 + a_1 x_1 + a_2 x_2 + \dots + a_k x_k. \quad (83)$$

В нашем случае при $k = 2$ дешевым экспериментом будет план 2^2 с 4-мя точками, расположенными на расстоянии $\pm \Delta x_s$ по каждой оси (см. табл. 24). После кодировки переменных и обработки эксперимента получим коэффициенты

$$a_1 = (+y_1 + y_2 - y_3 - y_4)/4; \quad a_2 = (+y_1 - y_2 + y_3 - y_4)/4$$

и уравнение $y = a_0 + a_1 x_1 + a_2 x_2 + a_{12} x_1 x_2$.

Парные взаимодействия здесь не интересны, и обычно они мало значимыми, поскольку периферийные точки эксперимента 2^{k-1} (в нашем случае 2^2) располагаются близко от точки A .

Легко видеть, что по полученному уравнению

$$\frac{\partial y}{\partial x_1} = a_1 \quad \frac{\partial y}{\partial x_2} = a_2.$$

В k -мерном случае частные производные по осям равны линейным коэффициентам регрессии (83). Таким образом, градиент (82) в точке A , определяющий направление крутого восхождения, может быть записан следующим образом:

$$\overline{\Delta y} = a_1 \Delta x_1 + a_2 \Delta x_2 + \dots + a_k \Delta x_k = \sum_s^k a_s \Delta x_s$$

(с учетом того, что Δx_s - векторы).

Шаг в направлении s -й оси можно выбрать постоянным. Но если он выбирается большим, то, шагая к максимуму, мы можем проскочить его. При малом шаге придется долго двигаться к цели, то есть эксперимент будет дорогим. Разумно выбрать шаг пропорциональным частным производным $\partial y / \partial x_s$. В самом деле, когда точка A находится еще далеко от точки M , эти производные будут большими и выбранный шаг также будет большим. По мере приближения к точке M крутизна поверхности уменьшается, то есть производные уменьшаются, и шаг замедляется. В точке M все производные равны нулю, и шаги по всем осям также становятся равными нулю. Вблизи точки максимума шаги будут уже почти нулевыми. Таким образом, при такой стратегии мы никогда не придем точно в точку M , но сравнительно быстро окажемся в ее окрестности.

Итак, определив направление градиента, то есть величину шагов в направлении каждой оси, необходимо сделать в заданном направлении несколько шагов и перейти в другую точку A' . В этой точке необходимо повторить дешевый эксперимент с целью корректировки направления движения, и продолжить движение уже в направлении нового градиента. Через несколько таких операций мы окажемся в области максимума. Признаком того, что мы находимся в этой области будет то, что при дальнейшем движении в том же направлении значение функции y практически не будет возрастать или даже станет уменьшаться, а частные производные в этой области по всем осям будут близкими к нулю.

Рассмотрим технику проведения экстремального эксперимента на конкретном примере, описанном в [11] и ставшим классическим. Он интересен также тем, что выполнен на металлургическую тематику и относится к классу задач “состав - свойство”. Была поставлена задача получить состав сплава, включающего семь химических легирующих элемента Cr, Ni, Mo, V, Nb, Mn и C (остальное Fe), обладающий максимальными прочностными характеристиками при определенных условиях эксплуатации этого конструкционного материала. Из некоторых металлургических соображений был выбран начальный сплав, аналог точки A на рис. 37, только в семимерном пространстве. В табл. 33 приведен состав этого сплава как основной уровень факторов, соответствующий нулевым значениям закодированных переменных. В той же таблице приведены выбранные по металлургическим соображениям интервалы изменения каждого из элементов.

Таблица 33.

Уровни факторов (закодированные)	Содержание элементов в %						
	Cr	Ni	Mo	V	Nb	Mn	C
Основной ($x_s = 0$)	4	2	0,1	0,02	0,1	0,4	0,4
Интервал (Δx_s)	1	1	0,1	0,02	0,1	0,1	0,1
Верхний ур-нь ($x_s = +1$)	5	3	0,2	0,04	0,2	0,5	0,5
Нижний ур-нь ($x_s = -1$)	3	1	0	0	0	0,3	0,3

Кодирование всех семи переменных произведено по формуле (67), оно определяет содержание всех элементов на верхнем и нижнем уровне плана эксперимента. Вокруг основного уровня, принятого за центр, необходимо провести дешевый эксперимент, достаточный для получения 8 членов регрессии (1 свободный и 7 линейных членов):

$$y = a_0 + a_1 x_1 + a_2 x_2 + \dots + a_7 x_7.$$

Таким экспериментом может служить ортогональная реплика от ПФЭ 2^{7-4} , содержащая 8 опытов. Как известно, такую реплику можно получить, осуществив полный перебор уровней только по трем факторам, например, x_1 , x_2 и x_3 . Уровни остальных факторов должны вычисляться по генерирующим соотношениям, которые мы выбираем сами. В приведенной работе Налимовым В.В. выбраны генерирующие соотношения:

$$\begin{aligned} x_4 &= x_1 x_2 x_3 ; & x_6 &= - x_1 x_3 ; \\ x_5 &= - x_2 x_3 ; & x_7 &= - x_1 x_2 . \end{aligned}$$

Конечно, можно было бы выбрать иные начальные генерирующие отношения. Но когда начальный набор выбран, то остальные генерирующие соотношения вычисляются путем перемножения левой и правой части любого из них на любое переменное. Такие генерирующие соотношения помогают определить, какие эффекты перемешиваются при вычислении коэффициентов регрессии. Нас не должно волновать то, что линейные эффекты перемешиваются с парными взаимодействиями, поскольку при малых интервалах Δx_s парные эффекты практически незначимы.

Эксперимент 2^{7-4} был фактически проведен, то есть выплавлены соответствующие восемь сплавов и измерена их прочность в заданных условиях. Результаты приведены в табл. 34. Так как эксперимент ортогонален, то коэффициенты регрессии a_i можно рассчитывать по формуле (66). Результаты расчетов приведены также в таблице.

Таблица 34

Эксперимент в точке A

Номер опыта u	x_1	x_2	x_3	x_4	x_5	x_6	x_7	y_u
1	+	+	+	+	-	-	-	5,8
2	+	+	-	-	+	+	-	3,5
3	+	-	+	-	+	-	+	6,2
4	+	-	-	+	-	+	+	5,3
5	-	+	+	-	-	+	+	3,2
6	-	+	-	+	+	-	+	5,1
7	-	-	+	+	+	+	-	5,3
8	-	-	-	-	-	-	-	1,5
Коэф. a_i	0,71	-0,09	0,64	0,89	0,54	-0,16	0,46	
Шаг $a_i \cdot \Delta x_s$	0,71	-0,009	0,064	0,018	0,054	-0,016	0,046	
Шаг с округ.	0,8	-0,1	0,07	0,02	0,06	-0,02	0,05	

Таблица 35

Движение по градиенту

№ шага	Опыт	Состав сплава (элементы в %)							y_u
		Cr	Ni	Mo	V	Nb	Mn	C	
1	Мысленно	4,8	1,9	0,17	0,04	0,16	0,38	0,45	
2	Мысленно	5,6	1,8	0,24	0,06	0,22	0,36	0,50	
3	Мысленно	6,4	1,7	0,31	0,08	0,28	0,34	0,55	
4	Мысленно	7,2	1,6	0,38	0,10	0,34	0,32	0,60	
5	Реализован	8,0	1,5	0,45	0,12	0,40	0,30	0,65	10,3
6	Мысленно	8,8	1,4	0,52	0,14	0,46	0,28	0,70	
7	Реализован	9,6	1,3	0,59	0,16	0,52	0,26	0,75	11,0
8	Реализован	10,4	1,2	0,66	0,18	0,58	0,24	0,80	11,5
9	Реализован	11,2	1,1	0,73	0,20	0,64	0,22	0,85	11,2
10	Реализован	12,0	1,0	0,70	0,22	0,70	0,20	0,90	10,1

Шаги по каждой координате, обеспечивающие движение по направлению градиента полученной функции, выбраны пропорционально частным производным (то есть коэффициентам a_i). Для удобства шаги округлены так, чтобы можно было реально обеспечить требуемый химический состав расчетного сплава.

Теперь можно двигаться по градиенту, значение y на каждом шагу будет расти. В табл. 35 показано движение, демонстрирующее технику крутого восхождения к максимуму. Как видно, в заданном направлении вначале полезно сделать несколько мысленных шагов, затем только проверить правильность движения. Лишь после пяти мысленных шагов была проведена выплавка требуемого сплава и измерены его прочностные свойства y . Установлено, что по сравнению с начальными значениями y вблизи точки A свойства металла резко улучшились. Здесь технологу, как туристу, совершающему восхождение к вершине, полезно отдохнуть и полюбоваться окружающей природой, получить эстетическое наслаждение от ощущения того, что ведет нас к вершине математика, а мы послушно шагаем за ней. Это она диктует нам, в каком направлении должны изменяться элементы в сплаве!

По логике здесь следует провести новый дешевый эксперимент, то есть провести новых 8 плавов, найти новый градиент, по которому движение к вершине будет более быстрым. Однако тратить средства на новые плавки дорого. Можно пока, не меняя старого направления, двигаться дальше до перевала, хотя, возможно, перевал окажется далеко от вершины. После седьмого шага свойства продолжают расти, но медленнее, чем на предыдущих шагах. Дальше, видимо, пропускать шаги рискованно, и необходимо делать плавки металла на каждом шаге. После 9-го шага свойства начали снижаться. Для того, чтобы убедиться в том, что это не случайность и пройден пере-

вал, провели плавку 10 - го шага. Убедились в том, что на 8 - м шаге действительно был перевал. Для дальнейшего движения к максимуму уже необходимо точку этого шага сделать центром нового эксперимента 2^{7-4} , затратив на него новых 8 опытов, по результатам вычислить новый градиент, после чего осуществить несколько реализованных опытов по новому пути. Однако видно, что после 8 -го шага движения по первому градиенту уже достигнуты приличные результаты: выплавлено всего 13 плавок и первоначальные свойства (нулевой уровень) увеличены более чем в 2 раза. Вторая серия плавок для движения в новом направлении, уже не даст такого эффекта. Удастся немного улучшить результат, но, вероятнее всего, такое улучшение не окупится затратами на проведение новой серии опытов. Эксперимент было решено остановить на первой серии опытов, и оптимальным признать состав сплава 8 - го шага: 10,4 % Cr; 1,2 % Ni; 0,66 % Mo ; 0,18 % V; 0,58 % Nb ; 0,24 % Mn ; 0,8 % C . Строго говоря, максимум не достигнут, возможно, даже не достигнута область вблизи максимума. В математические аспекты поиска максимума вмешались экономические и технологические соображения. И так, видимо, должно быть всегда: при решении любых технологических задач: надо уметь правильно сочетать результат со средствами, затраченными на его достижение.

7. 15. Адаптивные модели в системах управления.

Адаптивные, то есть самообучающиеся, модели используются в системах управления промышленными объектами. Например, при производстве труб связь между выходными параметрами (например, диаметром и толщиной стенки трубы) и входными технологическими параметрами процесса непрерывно изменяется по различным причинам, и постоянной модели для управления диаметром трубы и толщиной стенки не существует. Необходимо в ходе самого процесса производства постоянно обновлять эту модель, адаптируя к конкретной обстановке. В ЭВМ поступают данные от всех технологических датчиков и измерителей размеров трубы. По последним 16 произведенным трубам методом наименьших квадратов (значительно упрощенным для конкретной задачи) определяются коэффициенты регрессии, связывающей выходной параметр с входными. Такая модель процесса будет использована для управления размерами следующей (17-й) трубы. После ее производства параметры первой трубы исключаются из системы данных, и все коэффициенты регрессии пересчитываются. Происходит непрерывная автоматическая корректировка модели. Как видно, такая адаптивная система управления будет успешно работать не только при наличии соответствующих надежных датчиков и автоматических измерителей, но и при достаточно хорошей технологической проработке задачи, позволяющей установить набор определяющих входных параметров, характер их влияния на размеры и возможности воздействия на них. При управлении любым производственным объектом надежными будут только адаптивные модели, так как постоянные модели

обязательно через какое-то время практически перестают адекватно описывать процесс.

Иные адаптивные модели используются в автоматизированных системах, осуществляющих поддержание технологического процесса в оптимальных условиях. Можно однажды методом крутого восхождения найти оптимальный набор технологических (входных) параметров, при которых выход y имеет максимальное значение. Однако в процессе производства материалы стареют, оборудование изнашивается, условия меняются. Все это ведет к тому, что точка максимума M уходит, “ дрейфует”, и требуется постоянная автоматическая настройка процесса за счет небольшого смещения входных параметров, обеспечивающая нахождение выходного параметра всегда на максимальном уровне. Предполагается, что сама модель процесса со временем может также меняться.

Работу “автомата”, поддерживающего экстремальные условия работы, удобно рассмотреть на двухфакторной модели, приведенной на рис. 38

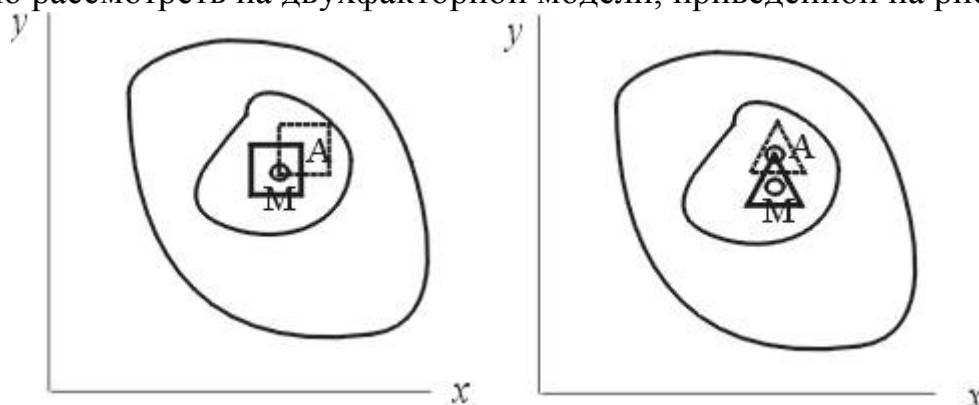


Рис. 38. Автоматическое слежение за максимумом

. Вблизи точки максимума M , приняв ее за центр, необходимо через небольшие промежутки времени проводить дешевый эксперимент ДФЭ типа 2^{k-l} . Для 2-х факторного эксперимента требуется ПФЭ 2^2 , содержащий 4 опыта, рис. 38 а). Чуть-чуть дешевле будет эксперимент на симплексе (в нашем случае, симплексом является треугольник, поэтому требуется 3 опыта, рис 38 б).

Для автоматического регулирования вблизи точки максимума M необходимо непрерывно “покачивать” процесс, то есть проводить измерения выходного параметра в упомянутых 4-х или 3-х точках. Если в центральной точке выход больше, чем в любой периферийной точке, то максимум находится где-то в области, ограниченной периферийными точками, при этом никакой настройки процесса не требуется. Если же в некоторой периферийной точке A значение выхода окажется больше, чем в точке M , то это свидетельствует о том, что максимум вышел из пределов области. В этом случае точку A следует принять за новый центр эксперимента, и вокруг нее организовать такой же непрерывный дешевый эксперимент.

7.16. Корреляционный анализ .

Корреляционный анализ предназначен для выявления наличия и силы корреляционной связи между двумя или несколькими количественными случайными величинами. Чтобы четче представлять отличие корреляционного анализа от регрессионного, вспомним однофакторную модель регрессионного анализа $y = f(x)$. Здесь выходной параметр y является нормально распределенной случайной величиной, а входной параметр x таковой не является, он принимает строго фиксированные значения на каждом уровне. Величина x , как предполагается, измеряется с точностью, на порядок превышающей точность измерения y . Связь между входным и выходным параметрами здесь причинная, то есть параметр y изменяется в опытах только потому, что изменяется фактор x .

В корреляционном анализе то же уравнение $y = f(x)$ имеет иной смысл. Здесь обе величины – x и y – случайные, распределенные по нормальному закону. В общем случае эти величины не связаны причинной зависимостью, то есть одна изменяется не обязательно потому, что изменяется другая. Может быть, они обе меняются потому, что зависят от каких-то общих причин, хотя сами по себе никакого отношения друг к другу не имеют. Корреляционные связи между величинами могут быть самыми разнообразными.

Первый крайний случай представлен на рис. 39, а. Здесь величина x не является фактором, определяющим значение величины y . Наоборот, эти величины, возможно, совершенно не имеют никакого отношения друг к другу, это некоторые равноправно существующие величины, не имеющие между собой ничего общего. Например, такая связь может наблюдаться при исследовании зависимости количества выпадаемых осадков от дня недели, хотя шутники заметили между этими параметрами тесную корреляционную связь, поскольку большинство дождливых дней приходится на воскресенье.

Другой крайний случай – зависимость между величинами x и y функциональная, описываемая каким-то уравнением. Здесь любому значению x соответствует строго или почти строго определенное значение y . При проведении эксперимента опытные точки точно или почти точно ложатся на кривую, как показано на рис. 39, б. Функциональные связи наиболее просты для экспериментального исследования, а исследуемый процесс отражает некоторый закон природы. Например, такова связь между давлением и объемом газа в сосуде при неизменной температуре (закон Менделеева-Клапейрона).

В промежутке между рассмотренными крайними случаями, находится все многообразие корреляционных связей между величинами x и y (в многофакторных моделях между x_i и x_j или между y_i и y_j) - рис 39 в.

При исследовании многофакторных моделей корреляционный анализ применяется для установления связей, прежде всего, между входными параметрами. Если, например, два фактора x_1 и x_2 имеют функциональную связь, то один из них следует исключить из рассмотрения, так как он ничего нового не вносит в модель. Таким образом, модель $y = f(x_1, x_2, \dots, x_k)$ упрощается. Наоборот, такие факторы, которые никак корреляционно не связаны, вносят

в модель свой независимый вклад, они независимы друг от друга и, как правило, характеризуют какие-то существенные стороны изучаемого процесса. Такими факторами, например, являются критерии процесса. К формированию таких входных параметров надо стремиться.

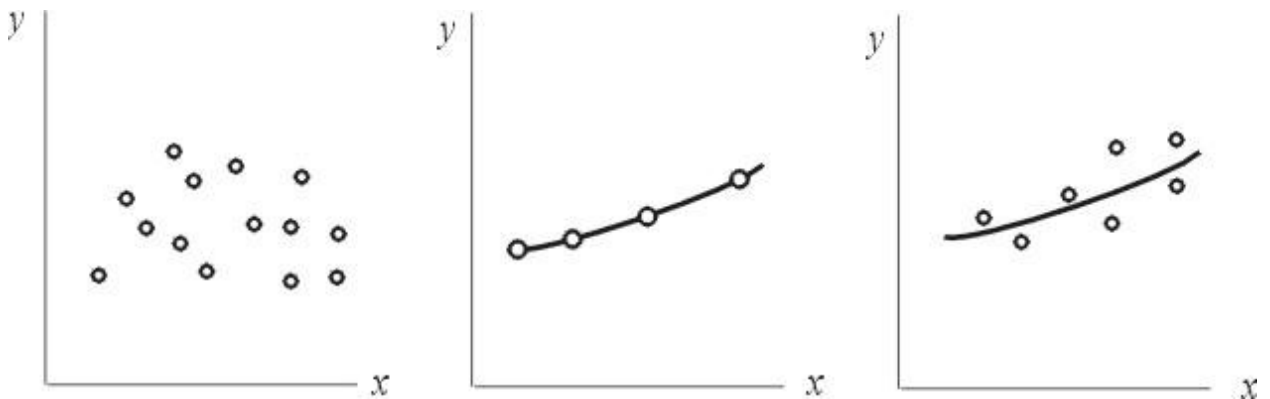


Рис. 39. Отсутствие зависимости между переменными (а), функциональная (б) и статистическая (в) связь переменных.

Корреляционный анализ применяется также для установления связей между несколькими выходными параметрами. К примеру, в [12] приведен эксперимент по изучению нескольких литейных характеристик сплава – жидкотекучести, объема и высоты усадочной раковины, усадочной микропористости – от различных технологических параметров плавки. Установлено, что между всеми перечисленными выходными характеристиками существует тесная линейная связь. Следовательно, нет необходимости изучать все эти характеристики, можно провести эксперимент только с одной из них, а остальные просто вычислить по установленным ранее линейным уравнениям. Более того, оказалось, что в данном эксперименте вообще можно не проводить измерений ни одного из выходных параметров, то есть вообще не проводить натурный эксперимент, а ограничиться только расчетным экспериментом, так как микропористость, а через нее и остальные выходные параметры, можно вычислить по технологическим (входным) параметрам. Аналогично, для многих сплавов установлена тесная корреляционная линейная связь между такими выходными параметрами как твердость и предел текучести, причем первая из этих характеристик измеряется сравнительно просто, а для определения второй требуется изготовление специальных образцов. Естественно, при планировании технологических экспериментов мы исключим из числа выходных параметров предел текучести материала, а включим параметр “твердость”. Вторую характеристику будем вычислять.

Как показано на рис. 39 в, в общем случае значение y не точно определяется величиной x . При одном и том же значении x можно получить некоторый диапазон значений y . Чем меньше связь этих величин, тем больше разброс. Однако видно, что при заданном значении x можно с определенной вероятностью прогнозировать значение величины y . Такова, например, связь

между количеством осадков в летние месяцы и урожаем, количеством затрачиваемого труда студентом на учебу и его успеваемостью и т. д.

При проведении многофакторных экспериментов наличие таких нежестких корреляционных связей между входными параметрами нежелательно, так как при осознанном изменении одного параметра неосознанно изменяется другой, да еще случайным образом. В целом, эксперимент может быть трудно воспроизводимым. Еще раз подчеркнем, что в правильно поставленном эксперименте входные параметры процесса должны быть некоррелированными и детерминированными.

Корреляционный анализ, таким образом, выявляет самые простые, но практически очень важные связи между переменными. Наибольшее практическое применение имеет раздел корреляционного анализа для установления связи между двумя величинами (парная корреляция). Этот раздел наилучшим образом разработан.

Если требуется установить связь сразу между несколькими случайными величинами, то используют методы множественного корреляционного анализа, но надежнее последовательно установить попарные связи между ними. В связи с меньшей практической значимостью множественной корреляции по сравнению с парной мы множественную корреляцию рассматривать не будем.

В парном корреляционном анализе (впрочем, как и в множественном) требуется сравнительно большое количество n измерений величин x и y . Если $n > 100$, то по каждому переменному необходимо весь диапазон изменения от минимального до максимального значения разделить на 8-20 равных интервалов. Соответственно при количестве k_i и k_j интервалов протяженность одного интервала по каждой переменной будет равна:

$$C_i = \frac{x_{max} - x_{min}}{k_i} \quad \text{и} \quad C_j = \frac{y_{max} - y_{min}}{k_j} .$$

Всем значениям, находящимся в одном интервале, присваивают среднее значение X_j и Y_j . В каждом i -м интервале по x и j -м интервале по y подсчитывают количество измерений n_{ij} и составляют таблицу наблюдений.

Для примера, пусть изучается зависимость разнотолщинности y холоднокатаного листа от разнотолщинности x исходного горячекатаного подката, являющегося заготовкой для холодной прокатки. Сначала длину исходного листа разделим на 100 равных участков и в определенной точке каждого участка проведем измерения толщины (точнее, отклонения ее от среднего значения, которое называется разнотолщинностью x). Найдем минимальное и максимальное значение x , назначим количество интервалов k_i и по приведенной выше формуле вычислим протяженность C_i каждого участка и количество n_i измерений, укладываемых в i -й интервал. Всем значениям x в этом интервале присваиваем одинаковое значение X_i . Аналогично поступим со значениями y разнотолщинности листа после прокатки его в холодном состоянии. В общем случае n_i точек i -го интервала исходного листа после прокатки распределятся на несколько интервалов n_{ij} готового листа. Конкретные данные приведены в табл. 36.

Таблица наблюдений n_{ij} .

j	Y_j	Значения X_i								n_j
		105	115	125	135	145	155	165	175	
1	70	4								4
2	80	3	6	2	1					12
3	90	5	6	7	1					19
4	100		3	10	5	2				20
5	110			3	6	8	7			24
6	120			1	1	1	7	1		11
7	130						2	1		3
8	140						1	1	5	7
n_i		12	15	23	14	11	17	3	5	100
Y_i		80,8	88,8	97,4	103,6	109,1	118,2	130,0	140,0	

Очевидно

$$n_i = \sum_{j=1}^{k_j} n_{ij}; \quad n_j = \sum_{i=1}^{k_i} n_{ij}; \quad n = \sum_{i=1}^{k_i} n_i = \sum_{j=1}^{k_j} n_j = \sum_{i=1}^{k_i} \sum_{j=1}^{k_j} n_{ij}.$$

Суммарное количество точек равно 100.

Общее среднее равно:

$$\bar{X} = \frac{\sum_{i=1}^{k_i} n_i X_i}{n} \quad \text{и} \quad \bar{Y} = \frac{\sum_{j=1}^{k_j} n_j Y_j}{n}.$$

Затем вычисляем ряд промежуточных величин:

корреляционный момент

$$C_{xy} = \frac{\sum_{i=1}^{k_i} X_i \sum_{j=1}^{k_j} n_{ij} Y_j}{n} - \bar{X} \cdot \bar{Y},$$

стандарты

$$S_x = \sqrt{\frac{\sum_{i=1}^{k_i} n_i X_i^2}{n} - \bar{X}^2} \quad \text{и} \quad S_y = \sqrt{\frac{\sum_{j=1}^{k_j} n_j Y_j^2}{n} - \bar{Y}^2},$$

а также выражение

$$S_{xy} = \sqrt{\frac{\sum_{i=1}^{k_i} n_i (Y_i - \bar{Y})^2}{n}}, \quad \text{где} \quad Y_i = \frac{1}{n_i} \cdot \sum_{j=1}^{k_j} n_{ij} Y_j.$$

Далее вычисляем коэффициент корреляции ρ_{ij} и корреляционное отношение θ_y по формулам:

$$\rho_{ij} = \frac{C_{xy}}{S_x \cdot S_y}; \quad \theta_y = \frac{S_{yx}}{S_y} . \quad (84)$$

Корреляционный анализ возможен также при сравнительно небольшом числе опытов n . Однако надежность выводов при этом быстро падает. В этом случае не выделяют участки изменения переменных, и в каждом i -м опыте ($i = 1, 2, \dots, n$) рассматривают оба измеряемых параметра – x_i и y_i . В этом случае средние, стандарты и корреляционный момент рассчитываются по формулам:

$$\begin{aligned} \bar{x} &= \frac{1}{n} \sum_{i=1}^n x_i; & \bar{y} &= \frac{1}{n} \sum_{i=1}^n y_i; \\ S_x &= \sqrt{\frac{\sum_{i=1}^n x_i^2}{n-1} - \frac{n \cdot \bar{x}^2}{n-1}}; & S_y &= \sqrt{\frac{\sum_{i=1}^n y_i^2}{n-1} - \frac{n \cdot \bar{y}^2}{n-1}}; \\ C_{xy} &= \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = \frac{1}{n-1} \sum_{i=1}^n x_i y_i - \frac{n \bar{x} \bar{y}}{n-1} . \end{aligned}$$

В обоих случаях вычисленные по формулам (84) коэффициент корреляции и корреляционное отношение позволяют установить наличие и силу связи между величинами x и y :

- а) если $\rho_{ij} = \pm 1$, то между переменными x и y существует линейная функциональная связь вида $y = a_0 + a_1 x$; если $\rho_{ij} = 0$, то эта связь либо не существует вовсе, либо имеет нелинейный вид;
- б) если $\theta_y = 0$, то между x и y нет корреляционной связи, а при $\theta_y = 1$ эта связь функциональная (не обязательно линейная);
- в) если $\rho_{ij} = \theta_y$, то между переменными существует только линейная связь, при неравенстве, наоборот, связь нелинейная.

Равенство или неравенство величин ρ_{ij} и θ_y надо понимать в статистическом смысле: если критерий Фишера меньше или равен табличному значению:

$$F = \frac{n - k_j - 2}{k_j - 2} \cdot \frac{\theta_y^2 - \rho_{ij}^2}{1 - \theta_y^2} \leq F(p, n - 2) ,$$

то указанные параметры статистически не различимы и в этом смысле равны. Точно также в статистическом смысле следует понимать отличие от нуля вычисленного значения коэффициента корреляции. В табл. П 5 приведены значения $\rho_{кр}$, применяемые при $n < 100$, зависящие от уровня надежности p ($p = 0, 95$ и $0,99$) и числа опытов n (точнее, числа степеней свободы $f = n - 2$). Если $\rho_{ij} < \rho_{кр}$, то найденное значение коэффициента корреляции статистически не отличается от нуля, в противном случае следует признать линейную связь между параметрами статистически значимой. Видно, что при

малом числе опытов критическое значение $\rho_{кр}$ очень велико, поэтому корреляционный анализ не эффективен.

При наличии линейной связи уравнение прямой имеет вид:

$$y - \bar{Y} = \rho_{ij} \frac{S_y}{S_x} (x - \bar{X}) . \quad (85)$$

Для примера, по данным табл. 32 получаем:

$$\bar{X} = (12 \cdot 105 + 15 \cdot 115 + \dots) / 100 = 133,5 ;$$

$$\bar{Y} = (4 \cdot 70 + 12 \cdot 80 + \dots) / 100 = 102,8 ;$$

$$C_{xy} = [105 \cdot (4 \cdot 70 + 3 \cdot 80 + 5 \cdot 90) + 115 \cdot (6 \cdot 80 + 6 \cdot 90 + 3 \cdot 100) + \dots + 175 \cdot 5 \cdot 140] / 100 - 133,5 \cdot 102,8 = 294,2$$

$$S_x = \sqrt{(12 \cdot 105^2 + 15 \cdot 115^2 + \dots) / 100 - 133,5^2} = 19,36 ,$$

$$S_y = \sqrt{(4 \cdot 70^2 + 12 \cdot 80^2 + \dots) / 100 - 102,8^2} = 17,38 .$$

Значения Y_i приведено в таблице, например, для $i = 1$ имеем:

$$Y_1 = (4 \cdot 70 + 3 \cdot 80 + 5 \cdot 90) / 12 = 80,8 .$$

$$S_{yx} = \sqrt{12 \cdot (80,8 - 102,8)^2 + 15 \cdot (88,8 - 102,8)^2 + \dots} = 15,2 .$$

Коэффициент корреляции

$$\rho_{ij} = 294,2 / (19,36 \cdot 17,38) = 0,87,$$

корреляционное отношение

$$\theta_y = 15,2 / 17,38 = 0,87.$$

При $n = 100$ и $p = 0,95$ $\rho_{кр} = 0,164$.

Вывод: между разнотолщинностью до и после холодной прокатки существует тесная, почти функциональная прямолинейная корреляционная связь. Ее можно представить уравнением

$$y - 102,8 = 0,87 \frac{17,38}{19,36} (x - 133,5) = 0,78(x - 133,5).$$

Заметим еще раз, что в этом уравнении обе переменные – x и y – нормально распределенные случайные величины.

Получив такой результат, технолог может сделать целый ряд практически важных выводов. Во-первых, если он хочет иметь низкую разнотолщинность холоднокатаного листа, то должен позаботиться об этом на стадии горячей прокатки, при холодной прокатке уже ничего не исправить. Во-вторых, если измерена разнотолщинность после горячей, то можно не измерять ее после холодной прокатки, поскольку ее можно вычислить с высокой вероятностью по уравнению прямой (85). В-третьих, если все-таки необходимо выпускать более высококачественный холоднокатаный лист (менее разнотолщинный), то необходимо на стане холодной прокатки поставить систему автоматического регулирования толщины листа. И так далее.

Контрольные вопросы по главе 7.

1. Какие задачи решает регрессионный анализ?
2. Каковы процедуры метода наименьших квадратов?
3. Каковы свойства обращенной матрицы?
4. Какие планы экспериментов существуют и какими свойствами они обладают?
5. Что такое ковариация?
6. Ортогональные планы и формулы для расчета коэффициентов регрессии и ошибок в его вычислении.
7. При каких условиях линейная модель становится ортогональной.
8. Какими свойствами обладает эксперимент 2^k ?
9. Как проводится эксперимент 2^k Как найти ошибку воспроизводимости в таком эксперименте?
10. Свойства генерирующих соотношений.
11. Как с помощью генерирующих соотношений построить ДФЭ 2^{k-l} ?
12. Какими свойствами обладает симплекс-план?
13. Как строятся планы Плакетта-Бермана?
14. Какими свойствами обладают планы Плакетта-Бермана?
15. Что такое центральный композиционный план?
16. Каков порядок построения центрального композиционного плана второго порядка?
17. Как вычисляются суммы квадратов отклонений и число степеней свободы в регрессионном анализе:
 - a. общие для всего эксперимента
 - b. для каждого из членов регрессии
 - c. для оставшихся членов
 - d. для отброшенных членов?
18. Дисперсионный анализ регрессионной модели
19. Постройте самостоятельно ортогональный центральный композиционный план для трехфакторного эксперимента.
20. Каков порядок статистической обработки эксперимента, построенного по центральному композиционному плану?
21. Как строится отсеивающий эксперимент в регрессионном анализе Чем он лучше и хуже аналогичных планов дисперсионного анализа?
22. Какова логика планов крутого восхождения?
23. Как идеи планируемого эксперимента используются при построении автоматических систем управления?
24. Основные идеи корреляционного анализа. Чем корреляционный анализ отличается от регрессионного?

Литература

1. Кумэ Х. Статистические методы повышения качества. М.: Финансы и статистика, 1990. - 304 с.
2. Дунин-Барковский И.В. Взаимозаменяемость, стандартизация и технические измерения. Учебник для вузов. М.: “Машиностроение”, 1976. - 352 с.
3. Арутюнов В. О. Электрические измерительные приборы и измерения. Учебное пособие. М-Л.: Госэнергоиздат, 1958. - 632 с.
4. Гухман А.А. Введение в теорию подобия. М.: “Высшая школа”, 1973. - 296 с.
5. Чижиков Ю. М. Теория подобия и моделирование процессов обработки металлов давлением. М.: Metallurgy, 1970. - 296 с.
6. Клайн С.Д. Подобие и приближенные методы. М.: Мир, 1968. - 302 с.
7. Кассандрова О.Н., Лебедев В.В. Обработка результатов наблюдений. М.: “Наука”, 1970. - 104 с.
8. Бродский В.З., Бродский Л.И., Голиков Т.И. и др. Таблицы планов экспериментов для факторных и полиномиальных моделей. - М.: Metallurgy, 1982. - 751 с.
9. Трусов В.С. Теория планирования эксперимента. Учебное пособие. - Томск: Изд-во Томск. ун-та, 1983 - 248 с.
10. Асатурян В.И. Теория планирования эксперимента. Учебное пособие. - Радио и связь, 1983 - 248 с.
11. Налимов В.В. Статистические методы описания химических и металлургических процессов. - М.: Metallurgizdat, 1963. - 60 с.
12. Новик. Ф.С., Арсов Я.Б. Оптимизация процессов технологии металлов методами планирования экспериментов. - М.: Машиностроение; София: Техника, 1980. - 304 с.

ПРИЛОЖЕНИЕ

Таблица П1

Функция Лапласа

v	$\Phi(v)$	v	$\Phi(v)$	v	$\Phi(v)$	v	$\Phi(v)$
0	0	0,50	0,1915	1,0	0,3413	2,0	0,4772
0,05	0,0199	0,55	0,2088	1,1	0,3643	2,1	0,4821
0,10	0,0398	0,60	0,2257	1,2	0,3849	2,2	0,4861
0,15	0,0596	0,65	0,2422	1,3	0,4032	2,3	0,4893
0,20	0,0793	0,70	0,2580	1,4	0,4192	2,4	0,4918
0,25	0,0987	0,75	0,2734	1,5	0,4332	2,5	0,4938
0,30	0,1179	0,80	0,2881	1,6	0,4452	2,7	0,4965
0,35	0,1368	0,85	0,3023	1,7	0,4554	3,0	0,4986
0,40	0,1554	0,90	0,3159	1,8	0,4641	4,0	0,49997
0,45	0,1736	0,95	0,3289	1,9	0,4713	5,0	0,499997

Таблица П2

Коэффициенты Стьюдента t_p при $p = 0,95$

f	2	3	4	5	6	7	8	10	15	20	100
t_p	4,3	3,18	2,78	2,57	2,45	2,36	2,31	2,23	2,13	2,09	1,98

Таблица П3

Коэффициенты γ_1^2 , γ_2^2 и критерий χ^2 - распределения
Пирсона при вероятности $p = 0,95$.

f	2	3	4	5	8	10	15	20
γ_1^2	0,271	0,321	0,359	0,390	0,456	0,488	0,546	0,585
γ_2^2	39,5	13,9	8,26	6,02	3,67	2,54	2,40	2,08
χ^2	6,0	7,8	9,5	11,1	15,5	18,3	25,0	31,4

Таблица П4

Критерий совместимости θ_{\max} при $p = 0,95$

n	3	4	5	6	8	10	15	20	30	40	50
θ_{\max}	1,41	1,69	1,87	2,00	2,17	2,29	2,49	2,62	2,79	2,90	2,99

Таблица П 5
Критические значения коэффициента корреляции.

$f = n - 2$	1	2	3	5	8	13	25	50	100
$p = 0,95$	0,988	0,900	0,805	0,669	0,549	0,441	0,323	0,231	0,164
$p = 0,99$	1,000	0,990	0,959	0,874	0,765	0,641	0,487	0,354	0,254

Таблица П6
Табличное значение критерия Фишера при $p = 0,95$

Чсс	Число степеней свободы (чсс) числителя f_1							
зн. f_2	2	3	4	5	6	12	24	∞
2	19,2	19,2	19,3	19,3	19,3	19,4	19,5	19,5
3	9,6	9,3	9,1	9,0	8,9	8,7	8,6	8,5
4	6,9	6,6	6,4	6,3	6,2	5,9	5,8	5,6
5	5,8	5,4	5,2	5,1	5,0	4,7	4,5	4,4
6	5,1	4,8	4,5	4,4	4,3	4,0	3,8	3,7
8	4,5	4,1	3,8	3,7	3,6	3,3	3,1	2,9
10	4,1	3,7	3,5	3,3	3,2	2,9	2,7	2,5
12	3,9	3,5	3,3	3,1	3,0	2,7	2,5	2,3
15	3,7	3,3	3,1	2,9	2,8	2,5	2,3	2,1
20	3,5	3,1	2,9	2,7	2,6	2,3	2,1	1,8
30	3,3	2,9	2,7	2,5	2,4	2,1	1,9	1,6

Таблица П7

Ранги Дункана при $p = 0,95$

a	R_1	R_2	R_3	R_4	R_5	R_7	R_{15}	R_{20}	R_{50}
3	4,50	4,50							
4	3,94	4,01	4,02						
5	3,64	3,74	3,79	3,83					
6	3,46	3,58	3,64	3,68					
7	3,35	3,47	3,54	3,58	3,60				
10	3,15	3,30	3,37	3,43	3,46	3,47			
15	3,01	3,16	3,25	3,31	3,36	3,43	3,46		
20	2,95	3,10	3,18	3,25	3,34	3,40	3,45	3,47	
30	2,84	3,00	3,09	3,15	3,26	3,34	3,42	3,47	3,50

Контрасты C_{it}

a	t	Номер уровня i фактора x_1				
		1	2	3	4	5
3	1	-1	0	+1		
	2	+1	-2	+1		
4	1	-3	-1	+1	+3	
	2	+1	-1	-1	+1	
	3	-1	+3	-3	+1	
5	1	-2	-1	0	+1	+2
	2	+2	-1	-2	-1	+2
	3	-1	+2	0	-2	+1
	4	+1	-4	+6	+4	+1

Критическое значение F -отношения при вероятности $p=0,95$

f_2	f_1									
	4	6	8	10	15	20	30	50	100	∞
6	4,53	4,28	4,15	4,06	3,94	3,87	3,81	3,75	3,71	3,67
8	3,84	3,58	3,44	3,34	3,21	3,15	3,08	3,03	2,98	2,93
10	3,48	3,22	3,07	2,97	2,84	2,77	2,70	2,64	2,59	2,54
14	3,11	2,85	2,70	2,60	2,46	2,39	2,31	2,24	2,19	2,13
18	2,93	2,66	2,51	2,41	2,27	2,19	2,11	2,04	1,98	1,92
22	2,82	2,55	2,40	2,30	2,15	2,07	1,98	1,91	1,84	1,78
26	2,74	2,47	2,32	2,22	2,07	1,99	1,90	1,82	1,76	1,69
30	2,69	2,42	2,27	2,16	2,01	1,93	1,84	1,76	1,69	1,62
35	2,64	2,37	2,22	2,11	1,96	1,88	1,79	1,70	1,63	1,56
40	2,61	2,34	2,18	2,07	1,92	1,84	1,74	1,66	1,59	1,51
50	2,56	2,29	2,13	2,02	1,87	1,78	1,69	1,60	1,52	1,44
100	2,46	2,19	2,03	1,92	1,77	1,68	1,57	1,48	1,39	1,28
∞	2,37	2,09	1,94	1,83	1,66	1,57	1,46	1,35	1,24	1,00

