

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ  
ЛУГАНСКОЙ НАРОДНОЙ РЕСПУБЛИКИ  
ГОСУДАРСТВЕННОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ  
ВЫСШЕГО ОБРАЗОВАНИЯ  
ЛУГАНСКОЙ НАРОДНОЙ РЕСПУБЛИКИ  
«ЛУГАНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ  
ИМЕНИ ВЛАДИМИРА ДАЛЯ»

Стахановский инженерно-педагогический институт  
Кафедра технологии производства и охраны труда

**КОНСПЕКТ ЛЕКЦИЙ**  
по дисциплине  
**«ТЕОРИЯ И ПРАКТИКА ЭКСПЕРИМЕНТА**  
для магистров направления подготовки  
44.04.04 Профессиональное обучение (по отраслям)

*Рекомендовано к изданию Учебно-методическим советом  
ГОУ ВО ЛНР «ЛГУ им. В. ДАЛЯ»  
(протокол № \_\_\_ от \_\_\_\_\_ 2021 г.)*

Курс лекций по дисциплине *«Теория и практика эксперимента»* для магистров направления подготовки **44.04.04 Профессиональное обучение (по отраслям)**. / Сост.: С.Н. Акинъшин – Стаханов: : ГОУ ВО ЛНР «ЛГУ им. В.ДАЛЯ», 2021. – 88 с.

Курс лекций составлен в соответствии с учебным планом магистерской программы по профилю «Безопасность технологических процессов и производств», «Горное дело. Подземная разработка пластовых месторождений», «Горное дело. Технологическая безопасность и горноспасательное дело».

Курс лекций «Теория и практика эксперимента» состоит из шести тем учебного курса. В лекция приведен математический аппарат необходимый для *обработки экспериментальных данных, обоснования расчетной модели, разработки рабочей гипотезы, проведения анализа рассматриваемых процессов и выводы, вытекающие из результатов экспериментально-аналитических исследований.*

Составитель:

ст.препод. Акинъшин С.Н.

Ответственный за выпуск:

доц. Черникова С.А.

Рецензент:

доц. Петров А.Г.

©Акинъшин С.Н.,2021

© ГОУ ВО ЛНР «ЛГУ им. В. ДАЛЯ», 2021

## Содержание

Тема 1. ЭКСПЕРИМЕНТ КАК ПРЕДМЕТ ИССЛЕДОВАНИЯ .....	5
1.1. Классификация видов экспериментальных исследований .....	5
1.2. Погрешности результатов исследований.....	6
Тема 2. КРАТКИЕ СВЕДЕНИЯ ИЗ ТЕОРИИ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ.....	8
2.1. Вероятность случайных событий, их характеристики .....	8
2.2. Нормальный закон распределения .....	12
Тема 3. ПРЕДВАРИТЕЛЬНАЯ ОБРАБОТКА ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ	16
3.1. Вычисление характеристик эмпирических распределений .....	16
3.2. Статистические гипотезы .....	18
3.3. Отсев грубых погрешностей .....	20
3.4. Определение доверительных интервалов .....	23
для исследуемых величин.....	23
3.4.1. Оценка доверительного интервала для математического ожидания .....	24
3.4.2. Оценка доверительного интервала для дисперсии .....	26
3.5. Сравнение двух рядов наблюдений.....	28
3.5.1. Сравнение средних значений .....	28
3.5.2. Сравнение двух дисперсий .....	30
3.5.3. Проверка однородности нескольких дисперсий .....	32
3.6. Определение необходимого количества измерений.....	35
3.7. Проверка гипотезы нормального распределения.....	36
3.8. Преобразование распределений к нормальному.....	43
Тема 4. АНАЛИЗ РЕЗУЛЬТАТОВ ПАССИВНОГО ЭКСПЕРИМЕНТА. ЭМПИРИЧЕСКИЕ ЗАВИСИМОСТИ.....	44
4.1. Характеристика видов связей между рядами наблюдений.....	44
4.2. Определение коэффициентов уравнения регрессии.....	47
4.3. Определение тесноты связи между случайными величинами .....	49
4.4. Линейная регрессия от одного фактора .....	51
4.5. Регрессионный анализ .....	55
4.5.1. Проверка адекватности модели.....	56
4.5.2. Проверка значимости коэффициентов уравнения регрессии .....	57
4.6. Линейная множественная регрессия .....	58
4.7. Нелинейная регрессия.....	60

Тема 5. ОЦЕНКА ПОГРЕШНОСТЕЙ РЕЗУЛЬТАТОВ НАБЛЮДЕНИЙ.....	61
5.1. Оценка погрешностей определения величин функций .....	61
5.2. Обратная задача теории экспериментальных погрешностей .....	63
5.3.Определение наивыгоднейших условий эксперимента .....	65
Тема 6. МЕТОДЫ ПЛАНИРОВАНИЯ ЭКСПЕРИМЕНТОВ_ЛОГИЧЕСКИЕ ОСНОВЫ .....	65
6.1. Основные определения и понятия .....	65
6.2. Пример хорошего и плохого эксперимента.....	67
6.3. Планирование первого порядка .....	69
6.3.1. Выбор основных факторов и их уровней.....	69
6.3.2.Планирование эксперимента .....	70
6.3.3. Определение коэффициентов уравнения регрессии.....	71
6.3.4. Статистический анализ результатов эксперимента .....	72
6.3.5. Дробный факторный эксперимент.....	74
6.4. Планы второго порядка.....	77
6.4.1. Ортогональные планы второго порядка.....	78
6.4.2. Ротатабельные планы второго порядка.....	80
6.5. Планирование экспериментов при поиске оптимальных условий .....	82
6.5.1. Метод покоординатной оптимизации (Гаусса - Зейделя) .....	83
6.5.2. Метод крутого восхождения (Бокса-Уилсона).....	84
6.5.3. Симплексный метод планирования .....	85

# Тема 1. ЭКСПЕРИМЕНТ КАК ПРЕДМЕТ ИССЛЕДОВАНИЯ

Общей чертой, объединяющей инженеров, социологов, биологов и др. является то, что они проводят эксперимент. Биологи на животных, инженеры на различных установках, вычислительных машинах и промышленных объектах.

**Под экспериментом (от латинского *experiment*, переводится как “проба”, “опыт”) мы в дальнейшем будем понимать систему операций и воздействий на объект, предназначенных для получения информации об объекте, явлении на основе результатов измерений.**

Эксперименты и экспериментаторы могут отличаться друг от друга, но фактически планирование, проведение и анализ всех экспериментов осуществляется в одинаковой последовательности. Хотя объекты исследований различны, однако методы экспериментальных исследований имеют много общего:

- каким бы простым эксперимент не был вначале ставится план его проведения;
- стремятся уменьшить число переменных, поскольку это упрощает его работу и делает ее более экономичной;
- все стараются исключить влияние внешних (случайных факторов);
- необходимо контролировать ход эксперимента;
- всех интересует точность измерительных приборов и точность получения данных;
- и наконец, в процессе любого эксперимента необходимо анализировать получаемые результаты и давать их интерпретацию, поскольку без этого решающего этапа весь процесс не имеет смысла.

Мы с вами будем рассматривать инженерный эксперимент, причем под объектом исследования будем понимать либо модель физическую, либо модель математическую, реализованную в виде программного продукта на ПЭВМ, либо реальный процесс, устройство, конструкцию.

## *1.1. Классификация видов экспериментальных исследований*

По форме представления результатов выделяются следующие виды экспериментов:

**Качественный эксперимент.** Устанавливается факт существования каких-либо явлений, но количественных характеристик при этом не дается. Любой эксперимент, каким бы сложным он не казался, заканчивается представлением результатов, формулировкой выводов, выдачей рекомендаций. Эта информация может быть представлена в виде графиков, чертежей, таблиц, формул, статистических данных или словесных описаний. Качественный эксперимент, как правило, предусматривает именно словесное описание. Однако словесное описание – самый неэффективный способ представления результатов, поскольку не позволяет дать количественные рекомендации, анализировать свойства объекта в иных условиях, решать задачи его управления. В инженерной практике основное содержание эксперимента должно представляться числом или количественными зависимостями.

**Количественный эксперимент.** Позволяет не только фиксировать существование того или иного явления, но и устанавливать количественные взаимосвязи между факторами, определяющими протекание процесса, а также устанавливать математическую модель влияния этих факторов на то или иное явление.

По условиям проведения различают:

**Лабораторный эксперимент.** В лаборатории меньше влияние случайных погрешностей, обеспечивается большая “стерильность” условий проведения опытов, осуществляется в большинстве случаев и более тщательная подготовка, одним словом выше “культура эксперимента”. Как правило, в лабораторных условиях экспериментатор может воспроизвести опыт “одинаково” значительно лучше, чем в промышленности. Это означает, что при прочих равных условиях для установления некоторого факта на заводе потребуется выполнить значительно больше опытов, чем в лаборатории. Другое важное отличие – различные ограничения на возможности варьирования факторами. Когда в лаборатории исследуется химическая реакция, температуру по желанию можно менять в широких пределах, а в металлургических печах, напротив, если ее и можно менять, то в значительно более узком диапазоне и с большей осторожностью.

В промышленных экспериментах эти условия обеспечить значительно сложнее. Усложняются измерения и сбор информации, значительно больше влияние различного рода помех на организацию и проведение эксперимента, измерительные приборы, поэтому особенно необходимо использовать специальные методы. Требуется по возможно меньшему числу измерений получить наиболее достоверные результаты. Заметим, что в современной математической статистике имеются специальные методы, которые при том же количестве измерений позволяют повысить точность или даже при их уменьшении получить более представительную информацию.

По виду воздействий на объект различают:

Активный эксперимент – это такой эксперимент, когда уровень (значение) фактора для каждого опыта задает исследователь. Он целенаправленно изменяет условия функционирования процесса и наблюдает результаты. Такой эксперимент можно планировать.

Пассивный – это эксперимент, в котором уровень факторов регистрирует исследователь в каждом опыте, но не задает его значение. В дальнейшем он обрабатывает результаты такой регистрации и пытается давать рекомендации.

Здесь использовано новое понятие – фактор, под которым будем понимать контролируемые переменные объекта (явления), влияющие на количественные характеристики последнего. Количественные характеристики будем оценивать выходными величинами, функцией отклика.

## **1.2. Погрешности результатов исследований**

Результаты опытов обычно не являются точными. По различным причинам результаты любых двух параллельных опытов отличаются друг от друга, за исключением случайных совпадений. Экспериментатор, в какой бы области он не работал, почти всегда придерживается более или менее регулярной последовательности: вначале производится планирование, затем приобретается оборудование, после этого производятся испытания и наконец выполняется анализ и составляется отчет. При планировании и приобретении оборудования анализ ошибок должен быть на одном из первых мест.

Под точностью эксперимента понимают его качество, отражающее близость полученных результатов к истинному значению искомой величины. Точность эксперимента тем выше, чем меньше его погрешность.

Абсолютная погрешность – это разность  $\Delta x$  между результатом эксперимента  $x$  и истинным значением искомой величины  $x^*$ :

$$\Delta = |x - x^*|. \quad (1.1)$$

Относительная погрешность

$$\Delta^* = \frac{|x - x^*|}{x^*} \cdot 100\% = \frac{\Delta}{x^*} \cdot 100\%. \quad (1.2)$$

Следует заметить, что истинное значение величины, определяемой в результате эксперимента, всегда остается неизвестным, поэтому и погрешности эксперимента могут быть оценены лишь приближенно.

Приведенной погрешностью называют отношение абсолютной погрешности  $\Delta x$  к нормирующему значению  $x_n$ , выраженному в процентах:

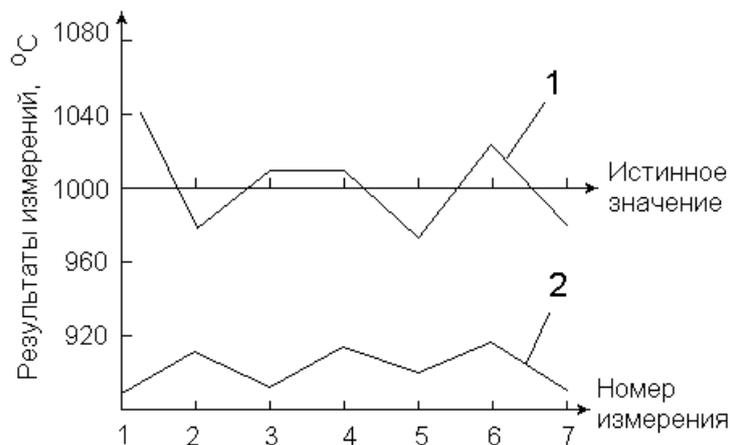
$$\gamma = \frac{\Delta x}{x_n} \cdot 100\%. \quad (1.2a)$$

В качестве нормирующего значения используют условно принятое значение измеряемой величины, выраженное в тех же единицах, в качестве которых, как правило, используют абсолютные значения разности верхнего и нижнего пределов шкалы.

При проведении эксперимента его погрешности принято условно разделять на систематические, случайные и грубые (промахи).

Систематической называется погрешность, которая при повторных экспериментах остается постоянной или изменяется закономерно. Наличие систематических погрешностей может быть обнаружено путем анализа условий измерения одного и того же значения измеряемой величины разными методами или приборами. Примером переменной систематической погрешности может

быть погрешность от измерения (закономерного) напряжения источника питания, если результат измерения зависит от напряжения (например, потенциометр). **Систематические погрешности нельзя уменьшить увеличением числа параллельных опытов. Должны устраняться вызывающие их причины.** Общим методом выявления причин систематических погрешностей является калибровка (поверка), которая представляет собой поверку прибора во всем диапазоне измеряемой величины с помощью известного эталона. Прибор может давать очень малый разброс показаний, но результат будет неверным вследствие наличия систематической ошибки. Пример: пирометр излучения дает показания, °С: 950, 952, 948, 950, 951 при истинном значении 1000 °С (рис.1.2). Можно выделить следующие источники систематических погрешностей.



**Рис. 1.2.** Пример данных, иллюстрирующий различие между случайной и систематической погрешностями: 1 – измерения характеризуются наличием случайной погрешности; 2 – измерения характеризуются наличием систематической погрешности.

- **инструментальные** (приборные или аппаратурные) погрешности средств измерений называются такие, которые принадлежат данному средству измерений, они могут быть определены при его испытаниях и занесены в его паспорт. Принято различать основную погрешность средств измерений, т.е. погрешность в условиях, принятых за нормальные, и дополнительную погрешность, вызванную отклонением влияющих параметров за пределы области нормальных значений (вибрации, влажности среды, инерцией и т.п.);
- **методические** погрешности – это погрешности, которые не могут быть приписаны данному прибору, не смогут быть указаны в его паспорте, т.е. связаны не с самим прибором, а с методикой проведения измерений. Очень часто причиной возникновения методической погрешности является то, что организуя измерения, измеряют или вынуждены измерять не ту величину, которую в принципе требуется измерять, а некоторую другую, близкую, но не равную ей. Например, для измерения температуры поверхности тела по его тепловому излучению, зависящего не только от этой температуры, но и приведенной степени черноты  $\epsilon_{пр}$  ( $q = \epsilon_{пр} C_0 T^4$ ). (Определение температурного поля термически массивного тела по температуре его поверхности). Отличительной особенностью методических погрешностей является то, что они могут быть определены лишь путем создания математической модели исследуемого объекта и не смогут быть найдены сколь угодно тщательным исследованием лишь самого измерительного прибора. Действительно, определить температурное поле тела по температуре его поверхности можно только располагая математической моделью нагрева металла, а определить температуру поверхности по показаниям радиационного пирометра только при заданной (рассчитанной) степени черноты этого тела.
- **субъективные** погрешности, обусловленные особенностями исследователя (совпадение яркости накала лампы и излучаемого тела в оптических пирометрах, определяемое наблюдателем, и т.п.).

Следует иметь в виду, что полностью исключить систематические погрешности невозможно, так как методы и средства, с помощью которых обнаруживаются и оцениваются систематические погрешности, сами имеют свои погрешности.

Случайной называется погрешность, обусловленная действием ряда причин, меняющихся случайным образом от эксперимента к эксперименту. Значение этой погрешности не может быть определено в каждом эксперименте и на нее невозможно оказать влияние. В то же время в результате большого числа экспериментов могут быть выявлены некоторые закономерности, присущие этому типу погрешностей. К случайным относятся непостоянные погрешности, причины возникновения которых неизвестны. Эти погрешности, как правило, вызываются сложной совокупностью изменяющихся факторов, обычно неизвестные экспериментатору и трудно поддающиеся анализу. Таким образом, **случайные погрешности представляют собой беспорядочные флуктуации показаний прибора относительно истинного значения измеряемой величины.** Для исследования случайных погрешностей, возникающих при проведении эксперимента, широко используются математическая статистика и теория вероятностей.

Грубые погрешности (промахи) возникают вследствие непредвиденного изменения условий эксперимента, качества измерений, поломок прибора, неправильной записи в рабочих журналах, механических ударах прибора, неправильном отчете показаний прибора, отключении источника питания и т.п. Результат, содержащий грубую ошибку, резко отличается по величине от остальных измерений. Такие результаты должны быть исключены из рассмотрения до математической обработки результатов эксперимента.

Вопросы к теме 1

1. Перечислите основные методы, используемые на этапе экспериментальных и теоретических исследований.
2. В чем отличие естественных экспериментов от искусственных?
3. Перечислите основные виды экспериментов.
4. В чем отличие лабораторных экспериментов от промышленных?
5. В чем отличие пассивного эксперимента от активного?
6. Перечислите основные направления теории эксперимента.
7. На какие вопросы отвечает теория подобия при планировании эксперимента?
8. Что такое объект исследования?
9. Что такое фактор?
10. Перечислите требования, предъявляемые к факторам.

## **Тема 2. КРАТКИЕ СВЕДЕНИЯ ИЗ ТЕОРИИ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ**

### ***2.1. Вероятность случайных событий, их характеристики***

При анализе погрешностей эксперимента и обработке результатов широко используется аппарат математической статистики и теории вероятностей, поэтому напомним некоторые основные понятия и определения теории вероятностей и математической статистики.

Случайная – это величина, принимающая в результате эксперимента некоторое значение, наперед неизвестное. Пусть проведено  $n$  измерений признака  $X$ , в результате получим ряд измерений (например, температура поверхности металла), отличающихся друг от друга

$$x_i \quad (i = 1, 2, \dots, n), \quad (2.1)$$

где  $x_i$  – " $i$ "-е измерение величины  $X$ ;  $x_1, x_2, \dots, x_n$  – реализация случайной величины  $X$ .

Случайная величина может быть непрерывной и дискретной. Случайная величина  $X$  называется непрерывной, если она может принимать любые значения в некотором интервале числовой оси (например, продолжительности выпуска чугуна и шлака из доменных печей, плавки в конвертерах и т.п.).

Дискретная случайная величина  $X$  принимает конечное или счетное, строго определенное число значений с определенной вероятностью (число слитков, печей, остановок печи, количество плавков в конвертере единицу времени, например, за месяц).

Полностью свойства случайной величины описываются законом распределения. **Под законом распределения понимают связь между возможными значениями случайной величины и соответствующими им вероятностями.** Мерой вероятности дискретной случайной величины может служить частота наступления событий:

$$P_i = \frac{m_i}{n}, \quad (2.2)$$

где  $P_i$  – вероятность " $i$ "-го события;  $m_i$  – число наступлений " $i$ "-го события в испытаниях;  $n$  – общее число испытаний.

При возрастании  $n \rightarrow N$  (где  $N$  – большое число) отношение  $m_i/n$  принимает все более устойчивое значение, т.е. становится статистически устойчивым. Предел, к которому стремится отношение  $m_i/n$  при неограниченном возрастании числа экспериментов (испытаний)  $n$ , называют вероятностью случайного события

$$P(X) = \lim_{n \rightarrow \infty} \left( \frac{m_i}{n} \right).$$

Говорят, что действует закон больших чисел. Таким образом, частота случайного события – это отношение числа появления этого события к общему числу произведенных испытаний. Нетрудно заметить, что  $0 < P_i < 1$ . Если  $P_i = 0$ , то событие невозможно,  $P_i = 1$  – достоверно.

Пример 2.1. Требуется определить вероятность числа остановок печи, исходные данные представлены в табл. 2.1, а результаты расчета – на рис.2.1.

Таблица 2.1

Число остановок печи по месяцам  
(общее число испытаний  $n=18$ )

Месяц	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Число Остановок	3	4	3	5	5	5	6	4	6	5	5	2	4	6	7	5	6	7

Число остановок	2	3	4	5	6	7
Количество соответствующих остановок печи, $m_i$	1	2	3	6	4	2
Вероятность, $P_i = m_i/n$	0,05	0,11	0,17	0,34	0,22	0,11

По графику, представленному на рис.2.1, можно сделать некоторые выводы. Так, наиболее вероятное число остановок печи равно 5, а вероятность числа остановок печи в диапазоне от 2 до 5 включительно составляет 0,67. Заметим, что эти выводы сделаны на основании весьма ограниченного числа наблюдений, равного 18. Как же оценить достоверность и надежность полученных результатов? Как изменится их достоверность при увеличении числа измерений? Сколько же измерений (наблюдений) при этом следует провести? Теория инженерного эксперимента, как мы увидим в дальнейшем, позволяет дать ответ на эти вопросы.

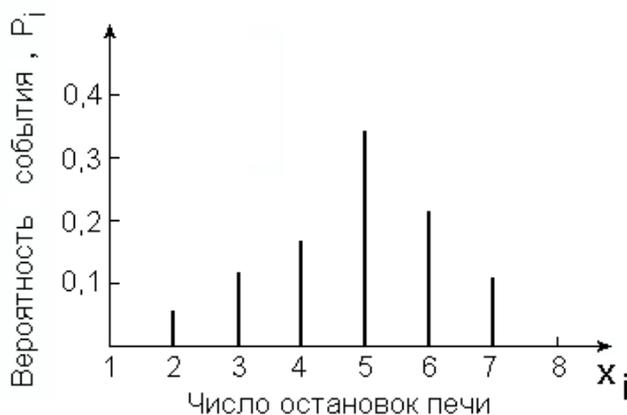


Рис.2.1. Вероятность остановок печи

Различают два вида описания законов распределения: интегральный и дифференциальный.

Для характеристики непрерывной случайной величины используется интегральная функция распределения  $F(x)$  – это вероятность того, что случайная величина  $X \leq x$ , т.е.  $-\infty \leq X \leq x$

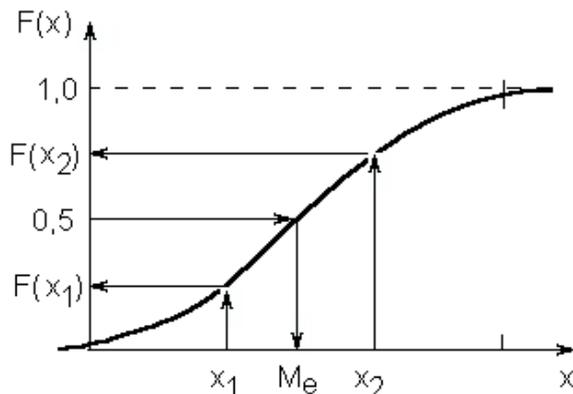
$$F(x) = P(X < x), \quad (2.3)$$

где  $x$  – текущая точка числовой оси.

Интегральная функция распределения имеет следующие свойства (рис.2.2):

$$1. 0 \leq F(x) \leq 1. \quad (2.4)$$

$$\lim_{x \rightarrow -\infty} F(x) = 0, \quad \lim_{x \rightarrow +\infty} F(x) = 1. \quad (2.5)$$



**Рис.2.2.** Интегральная функция распределения:

$$F(x_2) - F(x_1) = P(x_1 \leq x \leq x_2)$$

2. Она представляет собой монотонно возрастающую кривую. Если  $x_2 > x_1$ , то  $F(x_2) > F(x_1)$ . Ордината этой кривой, соответствующая точке  $x_1$ , представляет собой вероятность того, что случайная величина  $x$  будет меньше  $x_1$ .  $F(x_1) = P(x < x_1)$ . Таким образом,  $F(x)$  непрерывная и возрастающая функция.

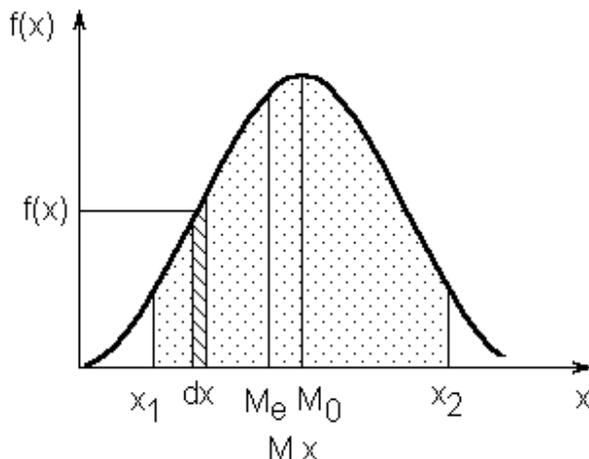
3. Ее приращение в промежутке  $(x_1; x_2)$  равно вероятности для величины  $X$  попасть в этот промежуток:

$$F(x_2) - F(x_1) = P(x_1 < x < x_2) \quad (2.6)$$

Дифференциальный закон распределения (плотность распределения) выражается соотношением (рис.2.3)

$$f(x) = \frac{dF(x)}{dx}; \quad F(x) = \int_{-\infty}^x f(Z) dZ, \quad (2.7)$$

то есть она характеризует приращение вероятности (ее скорость) при изменении значениях случайной величины  $X$  на единицу.



**Рис.2.3.** Дифференциальный закон распределения (плотность распределения)

Напомним основные свойства  $f(x)$ :

$$1) f(x) \geq 0. \quad (2.7a)$$

2) Площадь, заключенная под кривой плотности распределения, согласно правилу нормировки, равна единице, т.е. отражает вероятность всех возможных событий.

$$\int_{-\infty}^{+\infty} f(Z) dZ = 1. \quad (2.7б)$$

$$3) \lim_{x \rightarrow \infty} f(x) = 0. \quad (2.7в)$$

$$4) F(x_2) - F(x_1) = P(x_1 \leq x \leq x_2) = \int_{x_1}^{x_2} f(Z) dZ, \quad (2.7г)$$

Часто для характеристики случайной величины используют не сами функции распределения, а некоторые числовые параметры – параметры распределения (рис.2.4). Важнейшими параметрами распределения, характеризующими случайную величину  $x$ , являются ее математическое ожидание  $M_x$  (центр рассеивания) и дисперсия  $\sigma_x^2$  (степень рассеивания).

Рассмотрим характеристики положения центра распределения изучаемой случайной величины.

Для непрерывной случайной величины математическое ожидание определяется выражением

$$M_x = \int_{-\infty}^{+\infty} x \cdot f(x) dx, \quad (2.8)$$

где  $x$  – значение непрерывной случайной величины;  $f(x)$  – плотность вероятности. Важно знать геометрическую интерпретацию математического ожидания – это абсцисса центра тяжести кривой распределения плотности вероятности  $f(x)$ , т.е. дифференциального закона распределения.

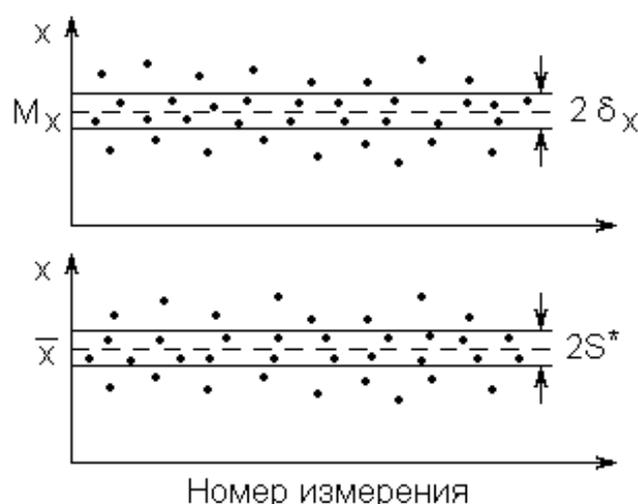
Для дискретной случайной величины

$$M_x = \sum_{i=1}^n x_i \cdot p(x_i), \quad n \rightarrow N, \quad (2.9)$$

где  $N$  – достаточно большое число;  $x_i$  – значение "i"-й дискретной случайной величины;  $P_i$  – вероятность ее реализации.

Сказанное проиллюстрируем на рис. 2.3, где видно, что произведение  $f(x)dx$  есть площадь элементарного участка под кривой  $f(x)$ , а  $x$  – абсцисса этого участка, т.е. расстояние от начала координат. Следовательно, интеграл (2.8) даст абсциссу центра тяжести всей площади под кривой  $f(x)$ .

Для рассмотренного ранее примера 2.1 величина  $M_x=4,88$  и не равна ни одному из значений дискретной случайной величины (следовательно, оно может быть дробным).



### Рис.2.4. К характеристикам рассеивания случайной величины

Модой  $M_o$  непрерывного распределения называют значение аргумента, при котором плотность распределения  $f(x)$  достигает максимума или это значение случайной величины, имеющей максимальную вероятность в том случае, когда случайная величина дискретная (для рассмотренного примера  $M_o=5$ ).

Медиана  $M_e$  – это значение аргумента, при котором число элементов совокупности со значением данного признака больше этой величины равно числу элементов со значением признака меньше ее. Для непрерывной случайной величины медиана определяется из решения уравнения:

$$F(M_e) = \int_{-\infty}^{M_e} f(x) dx = 0,5. \quad (2.10)$$

или для дискретной случайной величины

$$\sum_{i=1}^{M_e} P(x_i) = 0,5; \quad x_i \leq M_e. \quad (2.11)$$

Таким образом, для дифференциального закона распределения медиана есть такое значение непрерывной случайной переменной  $x$ , которая делит пополам площадь под кривой распределения  $f(x)$ .

Если общее число  $n$  значений дискретной случайной величины нечетно, то медиана равна значению случайной величины с индексом  $i=(k+1)/2$ , при четном  $n$  медиана равна  $M_e=(x_k+x_{k+1})/2$ . Обычно значения числовых характеристик  $M_x$ ,  $M_o$ ,  $M_e$  не совпадают.

Кроме характеристик положения центра, используются другие характеристики, описывающие рассеивание случайных величин. Важнейшей из них является дисперсия (см. рис. 2.4).

Дисперсия для непрерывной случайной величины определяется соотношением

$$\sigma_x^2 = \int_{-\infty}^{+\infty} (x - M_x)^2 \cdot f(x) dx, \quad (2.12)$$

где  $x$  – непрерывная случайная величина;  $f(x)$  – плотность вероятности.

Для дискретной случайной величины

$$\sigma_x^2 = \sum_{i=1}^n (x_i - M_x)^2 \cdot P(x_i), \quad n \rightarrow N, \quad (2.13)$$

где  $x_i$  – реализация "i"-й случайной величины;  $N$  – достаточно большое число.

**Дисперсия имеет размерность квадрата случайной величины и выражает как бы мощность рассеяния относительно среднего значения (математического ожидания).**

Положительное значение квадратного корня из дисперсии называется среднеквадратичным отклонением

$$\sigma_x = \sqrt{\sigma_x^2}. \quad (2.13a)$$

## 2.2. Нормальный закон распределения

Функции распределения  $F(x)$  и  $f(x)$  представляют собой математическую модель, которая описывает экспериментально наблюдаемые величины. Одной из задач статистической обработки данных является нахождение таких функций распределения, которые, с одной стороны, достаточно хорошо описывали наблюдаемые значения случайной величины, а с другой – были бы удобны для дальнейшего статистического анализа. Вид функции распределения предпочтительно выбирать на основе представлений о физической природе рассматриваемого явления, т.к. в этом случае исключаются возможные погрешности при распространении найденных закономерностей за пределы изучаемого интервала варьирования случайных величин.

Среди всех изученных до настоящего времени случайных величин наиболее важное место занимают случайные величины, имеющие так называемое нормальное (Гауссово) распределение

(рис.2.5). Нормальному закону подчиняются, как правило, результаты испытаний стали на прочность, производительность многих металлургических агрегатов, составы сырья, топлива, сплавов, массы слитков, отлитых в однотипные изложницы, случайные ошибки измерений и т.п., поэтому исследователи чаще всего используют это распределение при математической обработке результатов наблюдений.

Другие распределения, которые мы будем использовать в дальнейшем (Стьюдента, Фишера, Пирсона, Кохрена, а также критериальные таблицы) составлены на основе именно нормального распределения. Заметим, что в математической статистике сделана оценка для двух противоположных распределений: нормального и равномерного и показано, что оценки результатов исследований в технических задачах отличаются не более, чем на 20% .

Дифференциальная функция (плотность распределения) нормального распределения имеет вид:

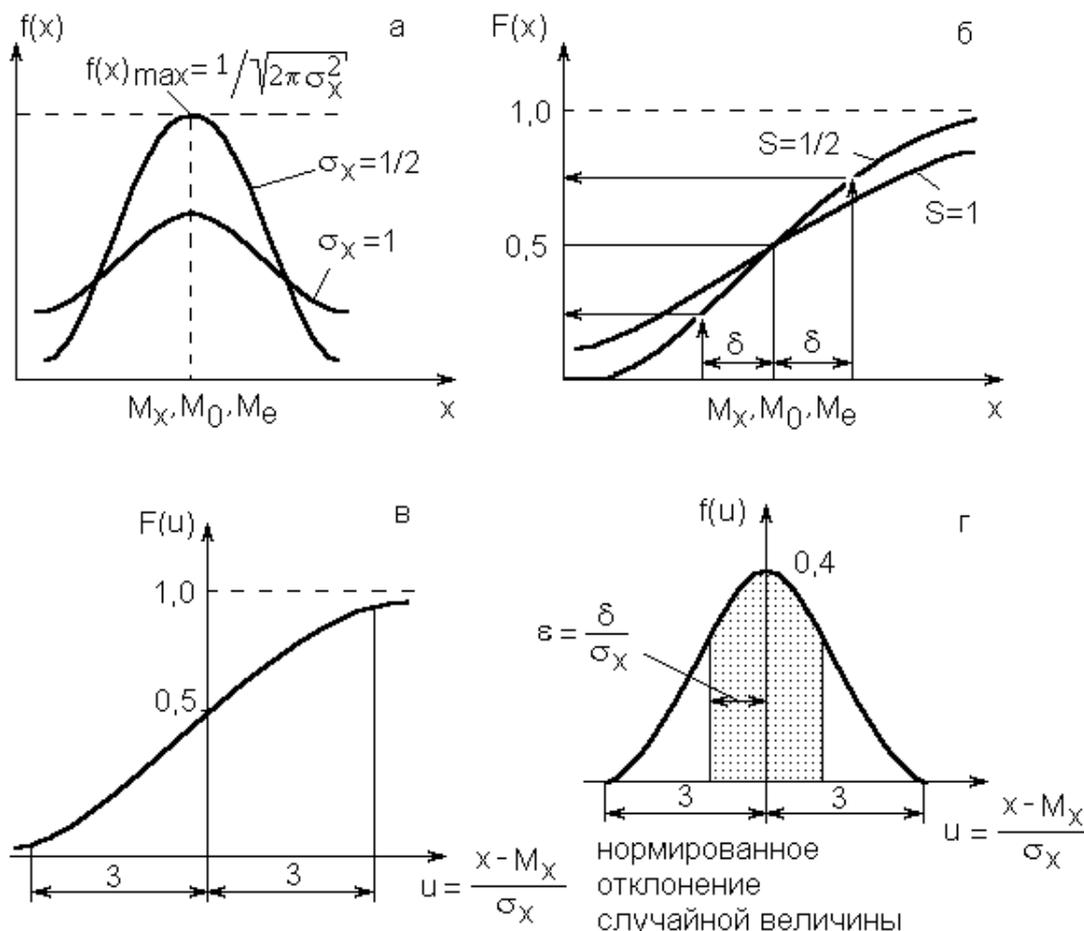
$$f(x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} \cdot e^{-\frac{[x-M_x]^2}{2\sigma_x^2}}, \quad (2.14)$$

Из уравнения (2.14) следует, что плотность распределения вероятностей для нормального распределения случайной величины определяется двумя параметрами  $M_x$  и  $\sigma_x^2$ .

Интегральная функция нормально распределенной величины имеет вид

$$F(x) = \int_{-\infty}^x f(Z) dZ = \frac{1}{\sqrt{2\pi\sigma_x^2}} \int_{-\infty}^x e^{-\frac{[Z-M_x]^2}{2\sigma_x^2}} dZ, \quad (2.15)$$

Отметим некоторые свойства нормального закона распределения.



**Рис.2.5.** Дифференциальная (а,г) и интегральная (б,в) функции при нормальном законе распределения случайных величин

1. Кривая плотности распределения симметрична относительно значения  $M_x$ , называемого иногда центром распределения.

2. При больших значениях  $\sigma_x^2$  кривая  $f(x)$  более пологая, т.е.  $\sigma_x^2$  является мерой величины рассеивания значения случайной величины около значений  $M_x$ . При уменьшении параметра  $\sigma_x^2$  кривая нормального распределения сжимается вдоль оси ОХ и вытягивается вдоль  $f(x)$ .

3. Максимум ординаты кривой плотности распределения определяется выражением

$$f_{\max} = \frac{1}{\sqrt{2\pi\sigma_x^2}}, \quad (2.16)$$

что при  $\sigma_x^2=1$  соответствует примерно 0,4.

4. Для нормального распределения среднее, мода и медиана совпадают

$$M_x = M_0 = M_e. \quad (2.17)$$

В ряде случаев рассматривается не сама случайная величина, а ее отклонение от математического ожидания

$$\Delta x = x - M_x.$$

Такая случайная величина называется центрированной. Очевидно, что  $M_x=0$ .

Введем нормированное отклонение (значение) случайной величины, выраженное в долях среднеквадратичного отклонения

$$u = \frac{x - M_x}{\sigma_x} = \frac{\Delta x}{\sigma_x}. \quad (2.18)$$

Для нормированного отклонения, распределенного по нормальному закону, получаем

$$f(u) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{u^2}{2}}; \quad (2.19)$$

$$F(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^u e^{-\frac{z^2}{2}} dz. \quad (2.20)$$

Графики этих функций показаны на рис. 2.5в, г.

Значение плотности вероятности и интегральной функции этого распределения, называемого нормированным нормальным распределением, табулированы и приведены в различных учебниках и справочниках по математической статистике.

В ряде случаев важно знать вероятность того, что случайная величина  $X$  не будет отличаться от своего среднего значения  $M_x$  больше, чем на величину  $\delta$ , т.е. нормально распределенная случайная величина примет значение на интервале  $[a;b]$  (рис.2.6). Эта вероятность называется **доверительной вероятностью (P)** и она показывает вероятность того, что результат измерений отличается от истинного значения на величину, не большую  $\delta$ :

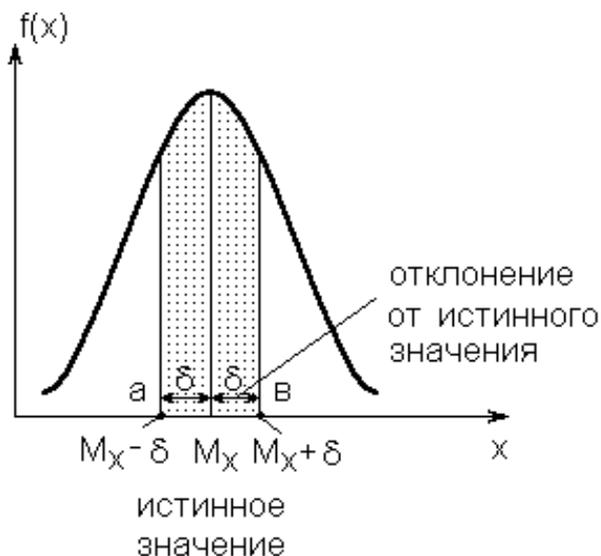
$$\begin{aligned} F(M_x + \delta) - F(M_x - \delta) &= P(M_x + \delta \leq x \leq M_x - \delta) = \\ &= \frac{1}{\sqrt{2\pi\sigma_x^2}} \int_{M_x - \delta}^{M_x + \delta} e^{-\frac{(z - M_x)^2}{2\sigma_x^2}} dz. \end{aligned} \quad (2.21)$$

**Интервал от  $(M_x - \delta)$  до  $(M_x + \delta)$  называется доверительным интервалом.** В дальнейшем уточним эти понятия. Заметим, что доверительная вероятность равна площади, заштрихованной на рис.2.6. Действительно, площадь, ограниченная интервалом  $\pm\delta$ , равна вероятности того, что случайная величина находится в этих пределах. Обозначим значение доверительного интервала, выраженного в долях среднеквадратичного отклонения, как

$$\varepsilon = \delta / \sigma_x. \quad (2.21a)$$

Тогда, в случае использования нормированного значения случайной величины, выражение (2.21) примет вид

$$P(M_x - \delta \leq x \leq M_x + \delta) = \frac{1}{\sqrt{2\pi}} \int_{-\varepsilon}^{+\varepsilon} e^{-\frac{z^2}{2}} dz. \quad (2.22)$$



**Рис.2.6.** К понятию доверительного интервала

Эта функция называется нормированной функцией Лапласа и для облегчения расчетов эта функция представлена таблицами и приведена в справочной литературе. Так, доверительному интервалу, равному значению среднеквадратичного отклонения ( $\delta = \varepsilon \sigma_x = 1 \cdot \sigma_x$ ), соответствует доверительная вероятность 0,68. Иными словами, при нормальном законе распределения примерно 2/3 всех значений случайной величины (наблюдений) лежит в площади, отсекаемым перпендикуляром к оси ОХ ( $M_x = \pm \sigma_x$ ). Доверительному интервалу равному  $\delta = 1,96 \sigma_x \approx 2 \sigma_x$ , соответствует доверительная вероятность 0,95. Доверительному интервалу  $3 \sigma_x$  соответствует доверительная вероятность 0,997 (см. рис. 2.5в,г).

Следовательно, отклонение истинного значения случайной величины от математического ожидания не превосходит утроенного среднего квадратичного отклонения с вероятностью 0,997. Это свойство в математической статистике носит названия правила трех сигм.

Отметим дополнительно, что 90% значений случайной величины лежат в диапазоне между  $-1,64 \sigma_x$  и  $+1,64 \sigma_x$ . Таким образом, чем больше величина доверительного интервала  $\delta$ , тем с большей вероятностью величина  $x$  попадает в этот интервал.

Рассмотрим небольшой пример

**Пример 2.2.** Предположим, что математическое ожидание содержания кремния в чугуне равно  $M_{Si} = 0,6\%$ , а среднеквадратичное отклонение  $\sigma_{Si} = 0,15\%$ . В этом случае, мы можем быть уверены в том, что величина измеренного содержания кремния в чугуне будет находиться в интервалах:

$$0,6 \pm 0,68 \times 0,15 = 0,6 \pm 0,10 \text{ с вероятностью } 68\%;$$

$$0,6 \pm 1,64 \times 0,15 = 0,6 \pm 0,29 \text{ с вероятностью } 90\%;$$

$$0,6 \pm 1,96 \times 0,15 = 0,6 \pm 0,29 \text{ с вероятностью } 95\%;$$

$$0,6 \pm 3,00 \times 0,15 = 0,6 \pm 0,45 \text{ с вероятностью } 99,7\%,$$

т.е. из 1000 проб только 3 пробы по содержанию кремния в чугуне будут выходить из диапазона от 0,15 до 1,05%.

Заметим однако, что мы предполагали нормальность закона распределения измерений, а также то, что изначально были известны математическое ожидание  $M_x$  и среднеквадратичное отклонение  $\sigma_x$ , т.е. выполнено большое (в пределе бесконечное) число измерений. Как же быть в реальных условиях, когда число измерений весьма ограничено? Рассмотрим в дальнейшем методологию решения этой задачи.

Вопросы к теме 2

1. В чем отличие математической модели от статистической?

2. Опишите процедуру построения математической модели.
3. Укажите виды подобия.
4. В чем отличие геометрического подобия от кинематического (динамического)?
5. Что такое критерии подобия?
6. Как определяется число опытов для полного факторного эксперимента?
7. Какой эксперимент называется экстремальным?
8. В каких случаях используется симплекс-планирование эксперимента?

### **Тема 3. ПРЕДВАРИТЕЛЬНАЯ ОБРАБОТКА ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ**

Предварительная обработка результатов измерений и наблюдений необходима для того, чтобы в дальнейшем с наибольшей эффективностью, а главное корректно, использовать для построения эмпирических зависимостей статистические методы и корректно проанализировать полученные результаты.

Содержание предварительной обработки состоит в отсеивании грубых погрешностей, оценке достоверности результатов измерений. Другим важным моментом предварительной обработки данных является проверка соответствия распределения результатов измерения закону нормального распределения и определения параметров распределения. Если эта гипотеза неприемлема, то следует определить какому закону распределения подчиняются опытные данные и, если это возможно, преобразовать данное распределение к нормальному.

#### ***3.1. Вычисление характеристик эмпирических распределений***

При рассмотрении основных положений теории вероятностей и математической статистики, определении параметров распределения мы исходили из предположения, что осуществляется достаточно большое, в пределе бесконечное число испытаний  $n \rightarrow N$  ( $N \rightarrow \infty$ ), что практически осуществить невозможно. Однако имеются методы, которые позволяют оценить эти параметры по выборке (части) случайных событий.

Генеральной называется совокупность всех мыслимых значений наблюдений, которые мы могли бы сделать при данном комплексе условий. Другими словами все возможные реализации случайной величины, теоретически в пределе их может быть бесконечное число ( $N \rightarrow \infty$ ). Часть этой совокупности  $n \in N$ , т.е. результаты ограниченного ряда наблюдений  $x_1, x_2, \dots, x_n$  случайной величины, можно рассматривать как выборочное значение случайной величины (например, при определении химического состава сплавов, их механической прочности и т.п.). Если все слитки данной марки стали, чугуна, сплава разделить на образцы и исследовать их химический состав, механическую прочность и другие физические характеристики, то имели бы генеральную совокупность наблюдений. Фактически доступно, возможно (целесообразно), исследовать свойства весьма ограниченного числа образцов – это и есть выборка их генеральной совокупности.

По результатам такого ограниченного числа наблюдений можно определить точечные оценки законов распределения и их параметров. Оценкой (или выборочной статистикой)  $\Theta^*$  какого-либо параметра  $\Theta$  называется произвольная функция  $\Theta^* = \Theta^*(x_1, x_2, \dots, x_n)$  наблюдаемых значений  $x_1, x_2, \dots, x_n$ , в той или иной степени отражающая действительное значение параметра  $\Theta$ .

Если говорить о характеристиках распределений вероятностей, то характеристики теоретических распределений ( $M_x, \sigma_x^2, M_0, M_e$ ) можно рассматривать как характеристики, существующие в генеральной совокупности, а характеризующие эмпирическое распределение – как выборочные их характеристики (оценки). Числовые параметры для оценки  $M_x, \sigma_x^2$  и др. – называются иногда статистиками.

**Для оценки математического ожидания используется среднеарифметическое (среднее значение) ряда измерений по выборке:**

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad i = 1, 2, \dots, n, \quad (3.1)$$

где  $x_i$  – реализация либо дискретной, либо отдельная точка для непрерывной случайной величины;  $n$  – объем выборки.

Для характеристики разброса случайной величины используется оценка теоретической дисперсии – выборочные дисперсии (см.рис.2.4):

$$\tilde{S}_X^2 = \frac{\sum_{i=1}^n (x_i - M_X)^2}{n}; \quad (3.2a)$$

$$S_X^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}. \quad (3.2b)$$

Неотрицательное значение квадратного корня из выборочной дисперсии – это выборочное стандартное отклонение (выборочное среднеквадратичное) отклонение

$$\tilde{S}_X = \sqrt{\tilde{S}_X^2}; \quad (3.3a)$$

$$S_X = \sqrt{S_X^2}. \quad (3.3b)$$

Следует отметить, что в любой задаче, связанной с выполнением измерений, возможны два способа получения оценки значения  $\sigma_x^2$ .

При использовании первого способа снимается последовательность показаний прибора и путем сравнения полученных результатов с известным или калиброванным значением измеряемой величины находится последовательность отклонений. Затем полученная последовательность отклонений используется для вычисления среднего квадратичного отклонения по формуле (3.3a).

Второй способ получения оценки значения  $\sigma_x^2$  состоит в определении среднего арифметического  $\bar{x}$ , т.к. в этом случае действительное (точное) значение измеряемой величины неизвестно. В этом случае целесообразно использовать другую, формулу для нахождения среднеквадратичного отклонения (3.2б, 3.3б). Деление на  $(n-1)$  производится по той причине, что наилучшая оценка, получаемая путем усреднения массива  $X$ , будет отличаться от точного значения на некоторую величину, если рассматривается выборка, а не вся генеральная совокупность. В этом

случае сумма квадратов отклонений  $\sum_{i=1}^n (x_i - \bar{x})^2$  будет несколько меньше, чем при использовании

истинного среднего  $\sum_{i=1}^n (x_i - M_X)^2$ . При делении на  $(n-1)$  вместо  $n$  эта погрешность будет частично

скорректирована. В некоторых руководствах по математической статистике рекомендуется при вычислении выборочного среднеквадратичного отклонения всегда делить на  $(n-1)$ , хотя иногда этого делать не следует. **Нужно делить на  $(n-1)$  лишь в тех случаях, когда истинное значение не было получено независимым способом.**

Выборочное значение коэффициента вариации  $v$ , являющееся мерой относительной изменчивости случайной величины, вычисляют по формуле

$$v = \frac{S_X}{\bar{x}} \quad (3.4a)$$

или в процентах

$$v = \frac{S_X}{\bar{x}} \cdot 100\%. \quad (3.4b)$$

Та из выборок имеет большее рассеяние, у которой вариация больше.

К оценкам  $\bar{x}$ ,  $S_x^2$  предъявляются требования состоятельности, несмещенности и эффективности.

Оценка параметра  $\Theta^*$  называется состоятельной, если по мере роста числа наблюдений  $n$  (т.е.  $n \rightarrow N$  в случае конечной генеральной совокупности объема  $N$  и при  $n \rightarrow \infty$  в случае бесконечной генеральной совокупности) она стремится к оцениваемому теоретическому значению параметра  $\lim_{n \rightarrow \infty} \Theta^*(n) = \Theta$ .

Например, для дисперсии

$$\lim_{n \rightarrow \infty} S^2(n) = \sigma_x^2. \tag{3.5}$$

Оценка параметра  $\Theta^*$  называется несмещенной, если ее математическое ожидание  $M(\Theta^*)$  при любом  $n$  асимптотически стремится к истинному значению  $M(\Theta^*) = \Theta$ . Удовлетворение требованию несмещенности устраняет систематическую погрешность оценки параметра, которая зависит от объема выборки  $n$  и в случае состоятельности стремится к нулю при  $n \rightarrow \infty$ . Выше было определены две оценки для дисперсии  $\tilde{S}^2$  и  $S^2$ . В случае неизвестного значения математического ожидания (истинного значения измеряемой величины) обе оценки состоятельны, но только вторая (3.2б), (3.3б), как было показано ранее, является несмещенной. Требование несмещенности особенно важно при малом числе наблюдений, так как при  $n \rightarrow \infty$   $\tilde{S}^2 \rightarrow S^2$ .

Оценка параметра  $\Theta_1^*$  называется эффективной, если среди прочих оценок того же параметра  $\Theta_2^*$ ,  $\Theta_3^*$  она обладает наименьшей дисперсией.

$$M\{(\Theta_1^* - \Theta)^2\} = \min \tag{3.6}$$

или

$$M\{(\Theta_1^* - \Theta)^2\} \leq M\{(\Theta_i^* - \Theta)^2\} \tag{3.6a}$$

где  $\Theta_i^*$  – любая другая оценка.

Так, если имеется выборка  $x_1, x_2, \dots, x_n$  из генеральной совокупности, то среднее математическое ожидание можно оценить двумя способами:

$$\bar{x}_A = \frac{1}{n} \sum_{i=1}^n x_i, \quad \bar{x}_B = \frac{x_{\max}(n) + x_{\min}(n)}{2}, \tag{3.7}$$

где  $x_{\max}(n)$ ,  $x_{\min}(n)$  – соответственно максимальное и минимальное значения случайной величины из выборки  $n$ .

Обе оценки обладают свойствами состоятельности и несмещенности, однако можно показать, что дисперсия при первом способе оценки равна  $S_x^2/n$ , а во втором  $\pi^2 S_x^2/[24 \ln(n)]$ , т.е. существенно больше. Таким образом, первый способ оценки математического ожидания является состоятельным, несмещенным и эффективным, а второй – только состоятельным и несмещенным. Заметим, что из всех несмещенных и состоятельных оценок следует предпочесть такую, которая оказывается наиболее близкой к оцениваемому параметру.

Заметим, что все сказанное относится к равноточным измерениям, т.е. к измерениям, которые содержат только случайную погрешность, подчиняющуюся нормальному закону распределения.

### 3.2. Статистические гипотезы

Как уже отмечалось, в ходе предварительной обработки экспериментальных данных решаются следующие задачи:

- отсев грубых ошибок (промахов) наблюдений;
- доверительная оценка измеряемых величин;
- проверка соответствия распределения результатов измерений закону нормального распределения.

Проверка и решение этих задач осуществляется с помощью статистических гипотез. **Статистическими гипотезами (H)** называются предположения о свойствах генеральной совокупности, т.е. относительно закона распределения ( $f(x)$ ,  $F(x)$ ) и их параметров ( $M_x$ ,  $\sigma_x^2$  и др.).

Например, специалиста интересует, удалось ли добиться повышения в среднем механической прочности окатышей при использовании новой технологии их обжига. Тогда он формулирует гипотезу: "Механическая прочность окатышей в среднем увеличилась". Для проверки этой гипотезы необходимо сформулировать другую гипотезу: "Изменения механической прочности в среднем не произошло". Эту последнюю гипотезу называют нулевой  $H_0$  или гипотезой отсутствия изменения (последствия). Задача исследователя заключается в том, чтобы на основе анализа ограниченной выборки принять ту или иную гипотезу.

Таким образом, основная выдвинутая гипотеза является нуль-гипотезой  $H_0$ . Весьма часто она формулируется о том, что оценки  $\Theta_A^*$  и  $\Theta_B^*$ , полученные для выборок А и В, принадлежат к одной генеральной совокупности, т.е. разница между  $\Theta_A^*$  и  $\Theta_B^*$  фактически равна нулю и возникла в силу случайности отбора элементов и ограниченности объема выборки. Противоречащие ей гипотезы  $H_1$  называются альтернативными, или конкурирующими.

Правильность этих гипотез проверяется путем вычисления некоторых числовых характеристик по данным наблюдений (измерений) и сравнения их с теми, которые должны быть при условии, что проверяемая гипотеза истинна, а наблюдаемые отклонения объясняются случайными колебаниями в выборках и их ограниченностью. Такие характеристики называют критериями проверки статистических гипотез  $G_r$ . Для этого строится случайная величина – степень рассогласования теоретического и экспериментального распределения. По величине рассогласования можно проверить является ли это расхождение незначительным или существенным. Если  $G_{r_{\text{эксп}}}$  и  $G_{r_{\text{теор}}}$  отличаются существенно, то гипотеза отвергается, если мало, то гипотеза принимается.

Таким образом, статистические гипотезы носят вероятностный характер. Это говорит о том, что в ряде случаев можно ошибиться. Для количественной характеристики степени ошибки используется показатель  $\alpha$ , который называется уровнем значимости. **Произведение  $\alpha \cdot 100\%$  показывает в скольких случаях из 100 можно ошибаться, т.е. сколько процентов отвергается гипотеза.** Тогда величина  $P=1-\alpha$  является доверительной вероятностью (надежностью). Уровень значимости – мера наших требований к ответу: чем больших гарантий мы будем требовать, тем менее определенным станет ответ. Ответ одновременно очень точный и очень надежный, как правило, стоит очень дорого. Опыт использования статистики в разнообразных ситуациях в течение нескольких десятилетий показал, что обычно в практических ситуациях определяющим значением  $\alpha$  является 0,05. Такое значение, называемое иногда 5%-ным уровнем риска, соответствует вероятности верного ответа, т.е. его надежности при проверке гипотез  $P=1-0,05=0,95$  или 95%. При этом говорят, что в среднем только в 5 случаев из 100 возможна ошибка. Конечно, никогда нельзя дать 100%-й гарантии. Подобное желание приводит к необходимости выполнять бесконечное много опытов, что абсурдно.

Поскольку проверка гипотез ведется при ограниченной информации (по выборке), то могут возникнуть ошибки двух родов. Если будет отвергнута правильная гипотеза, то совершается ошибка первого рода, если будет допущена неправильная гипотеза, то совершается ошибка второго рода. Очевидно, что вероятность допустить ошибку первого рода, равна  $\alpha$ . Область, отвечающая вероятности  $\alpha$ , называется критической, а дополняющая ее область, вероятность попадания в которую  $P(\Theta_\alpha)=1-\alpha$ , называется допустимой.

Вероятность ошибки второго рода обозначается  $\beta$ , а величина  $P(\Theta_\beta)=1-\beta$  называется мощностью критерия. Чем больше мощность критерия, тем меньше вероятность совершить ошибку второго рода. Так, при контроле качества продукции величина  $\alpha$  показывает риск производства, отвергая правильную гипотезу, он бракует годную продукцию. Величина  $\beta$  характеризует риск потребителя, допускает ложную гипотезу, он принимает и использует фактически непригодную продукцию. Ситуации, возникающие при статистической проверке гипотез представлены в табл.3.1.

Таблица 3.1

Общий подход решения задач при использовании статистических гипотез состоит в проверке нулевой гипотезы  $H_0$ , т.е. об отсутствии различия между теоретическими и экспериментальными результатами, разброс которых объясняется случайными факторами. Проверка гипотезы – это правило, по которому она принимается или отвергается.

Правильность нулевой гипотезы можно проверить следующим образом. Предположив справедливость нулевой гипотезы, т.е. отсутствия реального различия, вычисляется вероятность

того, что вследствие случайной выборки расхождение может достигнуть фактически величины, которая установлена в результате наблюдения. Если эта вероятность окажется очень малой, то нулевая гипотеза отвергается (т.е. маловероятно, что расхождение вызвано случайными величинами, а не реальным различием). Вероятность  $P=1-\alpha$ , которую принимают за основу при статистической оценке гипотезы, определяют уровнем значимости  $\alpha$ .

Возможные варианты принятия решения о правильности гипотез

	Критерий $G_r$ рекомендует допустить нуль-гипотезу $H_0$	Критерий $G_r$ рекомендует отклонить нуль-гипотезу $H_0$ (т.е. допустить альтернативную $H_1$ )
Фактически истина нуль-гипотеза $H_0$	Решение правомерно: гипотеза $H_0$ допускается	Решение ложно: совершена ошибка первого рода, т.к. отклонена верная гипотеза $H_0$
Фактически истинна альтернативная гипотеза $H_1$	Решение ложно: совершена ошибка второго рода, т.к. допускается ложная гипотеза $H_0$ вместо истинной $H_1$	Отклоняется $H_0$ , решение истинно, т.к. допущена гипотеза $H_1$

Однако чаще всего в инженерной практике изначально задаются критическим значением уровня значимости  $\alpha$  или доверительной вероятности  $P=1-\alpha$ . Далее находят экспериментальное значение соответствующего критерия  $G_{r_{\text{эксп}}}$ . По специальным таблицам или с использованием пакетов прикладных программ и интегрированных сред для ПЭВМ находят теоретическое значение критерия  $G_{r_{\text{теор}}}$ , зависящее, как правило, от уровня значимости  $\alpha$  (доверительной вероятности  $P$ ) и степени свободы  $m$ . В итоге сравнивают  $G_{r_{\text{эксп}}}$  и  $G_{r_{\text{теор}}}$  и делают выводы (заключение) о принятии или отклонении нуль-гипотезы.

Напомним, что число степеней свободы  $m$  – это понятие, которое учитывает в статистических ситуациях связи, ограничивающие свободу изменения случайных величин. **Число степеней свободы вычисляется как разность между числом экспериментальных точек  $n$  и числом связей  $f$ , ограничивающих свободу изменения случайной величины.** Так, при вычислении выборочной

дисперсии  $S_X^2 = \sum_{i=1}^n (x_i - \bar{x})^2 / (n-1)$  наблюдается одна связь, определяемая уровнем средней

арифметической величины  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ , поэтому число степеней свободы будет равно  $m=n-1$ , а для

дисперсии  $\tilde{S}_X^2 = \sum_{i=1}^n (x_i - M_X)^2 / n$  число степеней свободы равно числу испытаний  $m=n$ , так как  $M_X$

определено независимым способом.

Понятие о степени свободы поясним еще на примере решения системы линейных алгебраических уравнений. Допустим, что мы имеем систему из  $n$  линейных алгебраических уравнений с  $n$  неизвестными  $x_1, x_2, \dots, x_n$ . Очевидно, решение такой системы (при линейной независимости уравнений) будет ограниченным, т.е. такая система не будет иметь ни одной степени свободы. Но если для  $n$  неизвестных переменных мы имеем  $l$  уравнений, то такая система уравнений будет иметь число степеней свободы  $m=n-l$ .

### 3.3. Отсев грубых погрешностей

В литературе можно встретить большое количество различных рекомендаций для проведения отсева грубых погрешностей наблюдений. Наиболее распространенным и теоретически обоснованным является метод, основанный на доверительной вероятности.

Мы с вами установили (см. п. 2.1), что при нормальном законе распределения случайная величина не должна отличаться от своего математического ожидания с вероятностями: 90% на  $1,64 \sigma_x$ ; 95% на  $1,96 \sigma_x$  и т.п. Но это теоретические расчеты, в предположении, что известно математическое ожидание  $M_x$ , дисперсия  $\sigma_x$ , а результаты измерений строго подчиняются нормальному закону распределения при числе испытаний  $n \rightarrow N$ . Таким образом, в случае нормального закона распределения, при достаточно большом числе наблюдений (как показывает практический опыт, более 30), если какое-либо из измеренных значений  $x_{\max}$  отличается от его математического ожидания более чем на  $1,96 \cdot \sigma_x$ , то с вероятностью 95% его можно отбросить. Следовательно условие, что измеренная величина  $x_{\max}$  является грубой погрешностью можно выразить соотношением

$$\left| \frac{x_{\max} - M_x}{\sigma_x} \right| \geq \varepsilon(\alpha). \quad (3.8)$$

На практике, как правило, число измерений конечно и в большинстве случаев не превышает 15–30. При таком малом числе наблюдений мы можем определить только оценки математического ожидания  $M_x$  и дисперсии  $\sigma_x$ , т.е. рассчитать  $\bar{x}$  и  $S_x$ . Измерения при малом числе наблюдений чаще всего дают меньшее значение среднеквадратичной погрешности  $S_x$  по сравнению с погрешностью для достаточно большого ряда тех же измерений (в пределе всей генеральной совокупности)  $\sigma_x$ . Поэтому при неизвестных действительных значениях  $M_x$ ,  $\sigma_x$ , ограниченном числе испытаний, используют распределение Стьюдента и довольствуются весьма приближенными методами. Подробно оно рассмотрено в учебниках по математической статистике. Ограничимся здесь лишь толкованием идеологии этого распределения и методологии применения его при анализе результатов измерений. Распределение Стьюдента, упрощенно говоря, учитывает это обстоятельство. Иными словами, вероятность появления, например, одинаково больших погрешностей в распределении Стьюдента, т.е. при малом числе измерений, больше.

Стьюдент – псевдоним У.С. Госсета (1876-1937) – химика, работавшего в одной из пивоваренных фирм Великобритании. Он почти самостоятельно разработал статистику малых выборок. Поскольку в современной технике чаще всего исследуются небольшие по объему выборки (менее 30), то работа Стьюдента имеет большое практическое значение.

Рассмотрим случайную величину  $t$ , равную отношению случайной величины  $(x - \bar{x})$  и  $S_x$

$$t = \frac{(x - \bar{x})}{S_x}. \quad (3.9)$$

При этом предполагается, что случайная величина  $(x - \bar{x})$  распределена по нормальному закону. Если обозначить вероятность появления того или иного значения  $t$  в пределах  $[t - (1/2)dt; t + (1/2)dt]$  через  $f(t)dt$ , то как строго доказывается в курсах теории вероятностей и математической статистике, плотность распределения вероятности появления величины  $t$  имеет вид:

$$f(t) = \frac{\Gamma(n/2)}{\sqrt{\pi} \sqrt{n-1} \cdot d \left( \frac{n-1}{2} \right) \cdot \left( 1 + \frac{t^2}{n-1} \right)^{n/2}}. \quad (3.10)$$

Это распределение названо распределением Стьюдента. Здесь  $\Gamma(x)$  – гамма-функция, являющаяся обобщением понятия факториала и обладающая рекуррентным свойством:  $d(x+1) = x \cdot d(x)$ . Для целых чисел  $n$  справедливо  $d(n+1) = n!$  Множители при  $1 / \left[ 1 + t^2 / (n-1) \right]^{n/2}$  в  $f(t)$  выбраны так, чтобы площадь под любой кривой  $f(t)$  равнялась единице.

На рис. 3.1. приведено распределение Стьюдента для различных значений  $n$ . При  $n \rightarrow \infty$  (практически при  $n \geq 30$ ) распределение Стьюдента переходит в нормальное распределение с единичной дисперсией. Распределение Стьюдента позволяет оценить величину надежности  $P$  по

заданному значению  $(x - \bar{x})$  или, наоборот, по заданной величине надежности (доверительной вероятности) найти величину погрешности результата  $(x - \bar{x})$ .

Действительно, если взять на оси  $t$  некоторое значение  $t_{\alpha,n}$ , то величина надежности будет определяться площадью, ограниченной осью  $t$ , ординатами  $-t_{\alpha}$  и  $+t_{\alpha}$  и кривой  $f(t)$ . Следовательно, при недостаточно большом числе измерений ( $n \geq 30$ ) при расчете  $(x - \bar{x})$  при заданном уровне надежности  $P$ , необходимо вводить вместо коэффициента  $\varepsilon(\alpha)$ , коэффициент Стьюдента  $t_{\alpha,n}$ .

Процедура отсева грубых погрешностей измерений заключается в следующем:

1. По результатам наблюдений (измерений) и объему выборки  $n$  рассчитываются оценки математического ожидания  $\bar{x}$  и дисперсии  $S_x^2$ .

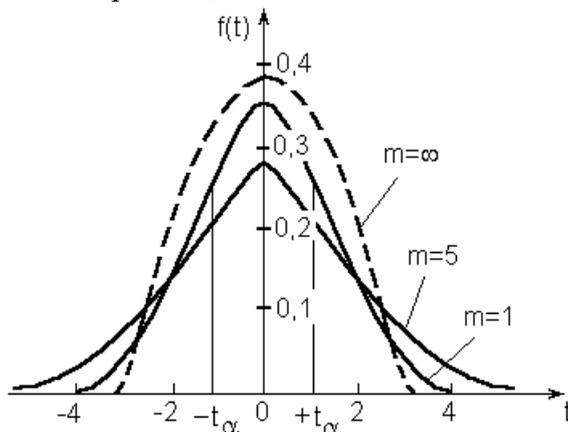


Рис.3.1. Кривая  $t$ -распределения Стьюдента

2. Из всего ряда наблюдений выбирается наблюдение (измерение), имеющее наибольшее отклонение от среднеарифметического значения  $x_{\max}$ .

3. Формулируется нуль-гипотеза  $H_0$ : отклонение  $x_{\max}$  от  $\bar{x}$  несущественно с доверительной вероятностью  $P$  (уровнем значимости  $\alpha$ ).

4. Для оценки этой гипотезы рассчитывается максимальное относительное (по отношению к среднеквадратичному) отклонение:

$$t_{\text{эксп}} = \frac{|x_{\max} - \bar{x}|}{S_x}, \quad (3.11)$$

где  $x_{\max}$  – "выскакивающее" значение.

5. В качестве критерия проверки статистической нуль-гипотезы используется теоретическое значение критерия Стьюдента  $t$ , которое зависит от уровня значимости  $\alpha$  или доверительной вероятности  $P=1-\alpha$ .  $t_{\alpha,m=n-1}$  – представляет собой допустимое отклонение случайной величины, выраженное в долях оценки среднеквадратичного отклонения и учитывает ограниченность объема выборки ( $n$ ) и заданную доверительную вероятность ( $P$ ). Данные  $t_{\alpha,m}$  для различных значений  $\alpha$  представлены в справочной литературе. При наличии современных ПЭВМ можно воспользоваться пакетами прикладных программ или интегрированными средами, например, электронными таблицами Microsoft Excel (см. статистическую функцию СТЬЮДРАСПОБР в п.7.1).

6. Если  $t_{\text{эксп}} > t_{\alpha,m}$ , то имеется достаточно основания с вероятностью  $P$  исключить "выскакивающее" значение как грубую ошибку и отвергнуть нуль-гипотезу. В противном случае  $t_{\text{эксп}} < t_{\alpha,m}$ , нуль гипотеза  $H_0$  принимается и от отсева "выскакивающего значения" лучше воздержаться с вероятностью  $P$ .

Рассмотрим небольшой пример.

Пример 3.1. Пирометром измеряется температура поверхности нагретого тела. Будем предполагать, что температура видимой поверхности нагретого тела во всех точках одинакова. Различными исследователями было проведено шесть измерений температуры и получены следующие их значения: Температура,  $^{\circ}\text{C}$ : 925, 950, 975, 1000, 1025, 1050 ( $n=6$ ).

Имеются ли среди этих измерений грубые погрешности? Предварительно вычислим оценки  $\bar{x}$  и  $S$ :

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = 987,5 \text{ } ^\circ\text{C}; S_x = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}} = 46,8 \text{ } ^\circ\text{C}.$$

Для определения  $S_x$  использовали  $(n-1)$ , т.к. истинное значение измеряемой температуры нам не известно. Заметим, что здесь это важно, т.к. сделано мало измерений (всего  $n=6$ ).

Выберем измерения, имеющие наибольшее отклонение от среднеарифметического значения. Таких значений оказалось два:  $925 \text{ } ^\circ\text{C}$  и  $1050 \text{ } ^\circ\text{C}$ .

Предварительно вычислим

$$t_{\text{эксп}} = \frac{1050 - 987,5}{46,8} = 1,34.$$

При  $\alpha=0,05$  и  $m=n-1=5$  определяем  $t_{0,05;5}=2,57$  (например, с помощью функции СТЬЮДРАСПОБР(0,05;5)=2,57 из электронных таблиц Excel).

Так как  $t_{\text{эксп}} < t_{\alpha,m}$ , то от отсева выделяющихся наблюдений лучше воздержаться.

Попытайтесь самостоятельно оценить, каким должно быть значение измеряемого параметра, чтобы его можно отнести к грубой погрешности при доверительной вероятности 0,997.

Заметим дополнительно, что если бы число наблюдений было достаточно большим и было бы известно действительное значение измеряемой температуры, при условии нормального закона распределения  $t^*_{0,05}=1,96$ , что соответствует теоретическому значению при доверительной вероятности равной 0,95. В нашем случае табличное значение  $t_{\text{эксп}}=2,57$  было существенно выше 1,96, т.к. оно учитывало ограниченность экспериментальных данных.

### 3.4. Определение доверительных интервалов для исследуемых величин

Постановка задачи заключается в следующем: требуется оценить точность и достоверность  $M_x$  и  $\sigma_x^2$ . Эти параметры нам, как правило, неизвестны, а имеются только их оценки  $\bar{x}$  (среднеарифметическое) и  $S_x^2$  (выборочная дисперсия). Насколько же близки оценки и их истинные значения? В связи с этим требуется оценить истинное значение  $M_x$ , т.е. указать границы интервала  $(\bar{x} - \delta; \bar{x} + \delta)$ , в который с заданной вероятностью  $P_{M_x}$  попадает действительное значение  $M_x$  (рис.3.2).

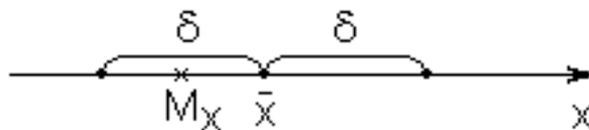


Рис.3.2. К понятию доверительного интервала

**Доверительный интервал ( $\delta$ )** – это случайный интервал, полностью определяющийся результатами опытов, который с вероятностью  $P_{M_x}$  покрывает (накрывает) скалярную статистическую характеристику (в данном случае  $M_x$ ). Здесь  $P_{M_x}$  – доверительная вероятность для этой характеристики.

Доверительный интервал в дальнейшем мы будем определять и для многих других статистических характеристик (дисперсии, среднеквадратичного отклонения, коэффициентов уравнений и т.п.), поэтому дадим этому понятию более общее и строгое определение.

Предположим, что для оценки параметра  $\Theta$  удалось найти две функции  $\Theta_1^*(x_1, x_2, \dots, x_n)$  и  $\Theta_2^*(x_1, x_2, \dots, x_n)$ , такие, что при всех  $(x_1, x_2, \dots, x_n)$  и при любых значениях  $\Theta$  выполняется условие

$$\left. \begin{aligned} \Theta_1^* &< \Theta_2^* \\ P\left\{\Theta_1^*(x_1, x_2, \dots, x_n) \leq \Theta \leq \Theta_2^*(x_1, x_2, \dots, x_n)\right\} &= 1 - \alpha. \end{aligned} \right\} \quad (3.12)$$

Это означает, что действительное значение параметра  $\Theta$  находится в интервале значений  $(\Theta_1^*; \Theta_2^*)$  с вероятностью  $P$ .

Интервал  $(\Theta_1^*; \Theta_2^*)$  называют доверительным интервалом для неизвестного параметра  $\Theta$ , соответствующего доверительной вероятности (или надежности)  $P=1-\alpha$ .

Для симметричного доверительного интервала его ширина  $2\delta$  определяется условием

$$P\{|\Theta - \bar{\Theta}^*| \leq \delta\} = 1 - \alpha, \quad (3.13)$$

где  $\bar{\Theta}_X^*$  – оценка параметра  $\Theta$ ;  $\delta$  – точность оценки. Чем меньше для данного  $\alpha$  будет  $\delta$ , тем точнее оценивается  $\Theta$ .

Вероятностное утверждение  $P\{\Theta_1^* \leq \Theta \leq \Theta_2^*\}$  следует понимать следующим образом:  $\Theta$  не есть случайная величина, лежащая с вероятностью  $P$  между  $\Theta_1^*$  и  $\Theta_2^*$ ;  $\Theta$  – есть константа нам неизвестная, но принимающая фиксированное значение. Пределы  $\Theta_1^*$  и  $\Theta_2^*$  будут случайными переменными и утверждение  $P\{\Theta_1^* \leq \Theta \leq \Theta_2^*\} = P$  означает, что для данного интервала, выбранного из совокупности интервалов, вероятность содержать значение  $\Theta$  равна  $P$ . Таким образом, доверительный интервал – это интервал, который с заданной доверительной вероятностью  $P$  накрывает оцениваемый параметр.

#### 3.4.1. Оценка доверительного интервала для математического ожидания

В качестве примера рассмотрим сначала оценку доверительного интервала для математического ожидания, т.к. именно такие задачи наиболее часто встречаются в инженерной практике.

Предварительно рассмотрим, как рассчитывается выборочное среднеквадратичное отклонение среднеарифметического значения  $\bar{S}_X$ . Физически это означает, например, что требуется определить среднеквадратичное отклонение (аналог точности) определения параметра при неоднократных его измерениях.

Из курса теории вероятностей и математической статистики Вам известно, что если измеряемая величина ( $y = x_1 \pm x_2$ ) является суммой или разностью двух случайных независимых величин  $x_1$  и  $x_2$ , то справедливо равенство

$$\sigma_y^2 = \sigma_{x_1}^2 + \sigma_{x_2}^2. \quad (3.14)$$

Дисперсия произведения случайной переменной  $X$  и постоянной величины  $C$  равна:

$$\delta^2(C \cdot x) = C \cdot \delta_x^2. \quad (3.15)$$

Закон сложения дисперсий сохраняется при любом числе слагаемых. Учитывая, что

$\bar{x} = \sum_{i=1}^n x_i / n$ , и  $S_x^2$  – дисперсия величины  $x$ , а также соотношения (3.14) и (3.15) имеем

$$\bar{S}_x^2 = S_x^2 \cdot \frac{1}{n}; \quad \bar{\sigma}_x^2 = \sigma_x^2 \cdot \frac{1}{n}; \quad (3.16)$$

$$\bar{S}_x = S_x \cdot \frac{1}{\sqrt{n}}, \quad \bar{\sigma}_x = \sigma_x \cdot \frac{1}{\sqrt{n}}, \quad (3.17)$$

где, как уже указывалось,  $\bar{S}_x$  – среднеквадратичное отклонение среднеарифметического значения.

Таким образом, **средняя квадратичная случайная ошибка среднеарифметического  $\bar{S}_x$  меньше, чем квадратичная ошибка единичного измерения  $S_x$  в  $\sqrt{n}$  раз, где  $n$  – число измерений.**

Анализ выражений (3.16, 3.17) позволяет сделать важный практический вывод: **Если точность результата измерений определяется случайной погрешностью, то повышение этой точности возможно в результате не только увеличения точности применяемого способа, т.е. уменьшения  $S_x$ , но и увеличения числа измерений  $n$ .** В то же время очевидно, что для повышения точности результата есть смысл увеличивать число измерений  $n$  только до тех пор, пока случайная погрешность является преобладающей. Дальнейшее увеличение числа измерений не будет приводить к заметному повышению точности результатов измерений, поскольку преобладающей становится систематическая погрешность, значение которой от числа измерений не зависит.

Если заранее известна дисперсия  $\sigma_x^2$  (или другая, связанная с ней характеристика точности измерений), то доверительный интервал для оценки  $M_x$ , учитывая соотношение (3.17), рассчитывается очень просто:

$$\delta = \varepsilon(\alpha) \cdot \overline{\sigma_x} = \varepsilon(\alpha) \frac{\sigma_x}{\sqrt{n}}, \quad (3.18)$$

где  $\varepsilon(\alpha)$  – значение доверительного интервала, выраженного в долях среднеквадратичной погрешности и зависящего от доверительной вероятности (надежности). Таким образом, при известной дисперсии, с вероятностью, например, 0,95 величина  $M_x$  будет находиться в интервале

$$\bar{x} - 1,96 \frac{\sigma_x}{\sqrt{n}} \leq M_x \leq \bar{x} + 1,96 \frac{\sigma_x}{\sqrt{n}}, \quad (3.19)$$

а с вероятностью 0,90 в интервале

$$\bar{x} - 1,67 \frac{\sigma_x}{\sqrt{n}} \leq M_x \leq \bar{x} + 1,67 \frac{\sigma_x}{\sqrt{n}}. \quad (3.20)$$

Пример.3.2. Предположим, что 50 пассажиров пользуются автобусом и средняя продолжительность ожидания составила  $\bar{x}=10$  мин., а  $\sigma_x=1,5$  мин. В таком случае мы могли бы на 95% быть уверенными, что среднее время ожидания автобуса находится в интервале

$$10 - 1,96 \frac{1,5}{\sqrt{50}} \leq M_x \leq 10 + 1,96 \frac{1,5}{\sqrt{50}},$$

т.е. в диапазоне примерно от 9,5 до 10,5 мин.

В этом примере мы предполагали, что продолжительность ожидания автобуса подчиняется нормальному закону распределения и считали, что нам известно действительное значение дисперсии продолжительности ожидания  $\sigma_x^2$ . В данном случае такие допущения вполне оправданы, т.к. проведено относительно большое количество измерений (50) с относительно высокой точностью (время можно измерить в данном случае в принципе с любой требуемой точностью).

На практике, как правило, число измерений (отбора проб шихты, чугуна, стали и других материалов) конечно и не превышает 10-30 и менее. При таком малом числе наблюдений фактическая дисперсия неизвестна, а можно получить только ее оценку  $S_x^2$ . При условии, что погрешности отдельных измерений подчиняются нормальному закону распределения можно опять же воспользоваться t-критерием Стьюдента, который, как уже указывалось, учитывает ограниченность объема выборки и невозможность определить действительное значение дисперсии  $\sigma_x^2$ .

В этом случае следует заменить  $\varepsilon(\alpha)$  на  $t_{\alpha,m}$ , а  $\sigma_x$  – на  $S_x$  и выражение (3.18) примет вид:

$$\delta = \frac{t_{\alpha,m} \cdot S_x}{\sqrt{n}} = t_{\alpha,m} \cdot \overline{S_x}, \quad (3.21)$$

где коэффициент  $t_{\alpha,m}$  – теоретическое значение коэффициента Стьюдента при уровне значимости  $\alpha$  или доверительной вероятности  $P=1-\alpha$  и числе степеней свободы  $m$ ;  $n$  – объем выборки.

Число степеней свободы в данном случае рассчитывается как  $m=n-1$ , так как при определении доверительной вероятности используется  $S_x^2$ , т.е. на случайные величины накладывается одна связь (более строгое доказательство приводится в учебниках по математической статистике). Теоретические значения критерия Стьюдента приводятся в справочной литературе по математической статистике (см., например, Большев Л.Н., Смирнов Н.В. Таблицы математических статистик. М.: Наука, 1982), или, как уже отмечалось выше, его можно определить, воспользовавшись статистической функцией СТЬЮДРАСПОБР из электронных таблиц Microsoft Excel (см. п.7.1).

Сравнивая выражения (3.18) и (3.21) нетрудно догадаться, что коэффициент  $t_{\alpha,m}$  представляет собой значение доверительного интервала, выраженного в долях выборочной среднеквадратичной погрешности и учитывает ограниченность объема выборки.

Естественно, что при  $n \rightarrow \infty$   $t_{\alpha,m} = \varepsilon(\alpha)$ . Однако уже при числе наблюдений  $n$  более 30 значения  $\varepsilon(\alpha)$  и  $t_{\alpha,m}$  практически совпадают. Так, например,  $t_{0,05;30} = 2,04$ , а  $\varepsilon(0,05) = 1,96$ .

**Пример.3.3.** Было отобрано 16 проб одного и того же чугуна, в которых определено содержание кремния. В дальнейшем рассчитали  $\bar{x}=0,65\%$  и  $S_x=0,11\%$ . Теоретическое значение критерия Стьюдента составило  $t_{0,05;15}=2,13$ . Тогда

$$\delta = \frac{2,13 \cdot 0,11}{\sqrt{16}} = 0,06.$$

Таким образом, с вероятностью 95% можно утверждать, что математическое ожидание (среднее значение) содержания кремния в чугуне лежит в диапазоне  $0,59\% \leq M_{Si} \leq 0,71\%$ .

### 3.4.2. Оценка доверительного интервала для дисперсии

Для оценки доверительного интервала дисперсии используется  $\chi^2$  (читается: "хи-квадрат-критерий"), предложенный английским математиком и биологом К.Пирсоном. Если обозначить через

$$\chi^2 = \sum_{i=1}^n \left( \frac{x_i - \bar{x}}{\sigma_x} \right)^2 = \frac{n-1}{\sigma_x^2} \cdot S_x^2, \quad (3.22)$$

то величина  $\chi^2$ , связанная с дисперсией  $\sigma_x$ , будет иметь следующий закон распределения:

$$f(\chi^2) = \frac{1}{2^{m/2} \cdot \Gamma(m/2)} \cdot (\chi^2)^{\frac{m-2}{2}} \cdot e^{-\chi^2/2}, \quad 0 \leq \chi^2 \leq \infty, \quad (3.23)$$

где  $m$  – число степеней свободы этого распределения ( $m=n-1$ ).

На рис.3.3 приведены кривые  $f(\chi^2)$  для различных значений  $m$ . Эти кривые асимметричны, причем асимметрия особенно резко выражена при малых значениях параметра  $m$ . При  $m=1$  кривая уходит в бесконечность при  $\chi^2=0$ , при  $m=2$  она достигает максимального значения, равного 0,5; при  $m>2$  кривые имеют максимум при  $\chi^2_{\max}=m-2$ . При больших значениях  $m$   $\chi^2$ -распределение переходит в нормальное со средним значением  $f(\chi^2) = \sqrt{2m-1}$  и дисперсией  $\sigma_{\chi^2}=1$ .

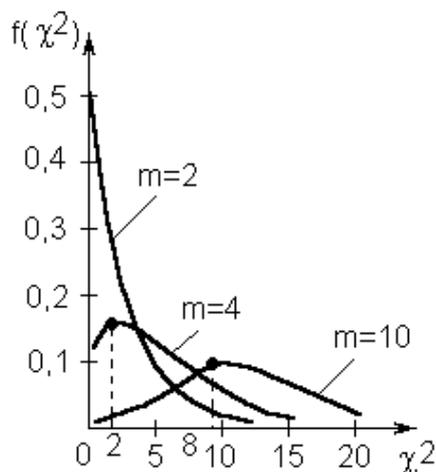


Рис.3.3. Кривые  $\chi^2$ -распределения для различных степеней свободы  $m$

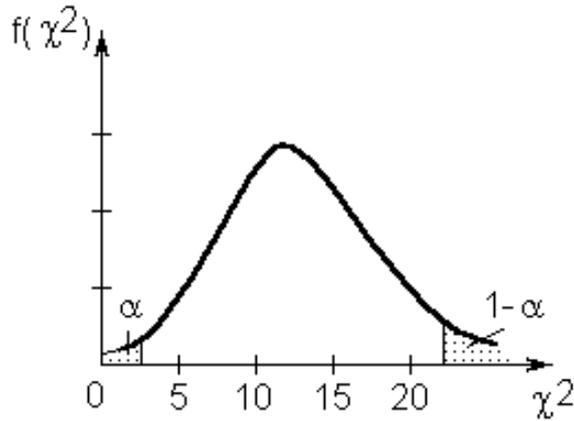


Рис.3.4. Определение границ доверительного интервала для оценки дисперсии  $\gamma^2$  при значении надежности  $P=1-2\alpha_1$

Применяя соотношение (3.22), можно записать

$$\sigma_x^2 = \frac{m}{\chi^2} \cdot S_x^2 = \gamma^2 \cdot S_x^2,$$

где  $\gamma^2 = m/\chi^2$ , но и после этого величина  $\gamma^2$  остается неизвестной, т.к. значение  $\chi^2$ , входящее в это равенство, а следовательно, и значение коэффициента  $\gamma^2$  остается неизвестным. Тем не менее с помощью  $\chi^2$ -распределения можно найти границы доверительных интервалов для  $\sigma_x^2$ . Для этой цели необходимо определить такие значения  $\chi_1^2(m)$  и  $\chi_2^2(m)$ , которые выделяют  $P$  – долю площади под соответствующей  $f(\chi^2)$  кривой (рис.3.4). Эти значения  $\chi_2^2(m)$  и  $\chi_1^2(m)$  определяют величины соответствующих им  $\gamma_1^2$  и  $\gamma_2^2$ . Таким образом, мы получили два неравенства, обуславливающих одинаковую надежность определения границ доверительного интервала для  $\gamma_x^2$  с двух сторон (слева и справа):

$$\chi^2 = \frac{m \cdot S_x^2}{\gamma_x^2} \leq \chi_1^2 = \frac{m}{\gamma_1^2} \quad (3.24)$$

с надежностью  $\alpha_1 = (1-P)/2$ ,

$$\chi^2 = \frac{m \cdot S_x^2}{\gamma_x^2} \leq \chi_2^2 = \frac{m}{\gamma_2^2} \quad (3.25)$$

с надежностью  $\alpha_2 = 1 - \alpha_1 = 1 - (1-P)/2 = (1+P)/2$ .

Из неравенств (3.24) и (3.25) получим, что вероятность попадания  $\chi^2$  в интервал  $(\chi_1^2; \chi_2^2)$ , т.е.

$$\gamma_1^2 \cdot S_x^2 \leq \sigma_x^2 \leq \gamma_2^2 \cdot S_x^2 \quad \text{или} \quad \frac{n-1}{\chi_1^2} \cdot S_x \leq \sigma_x \leq \frac{n-1}{\chi_2^2} \cdot S_x,$$

равна разности вышеупомянутых площадей  $\alpha_2 - \alpha_1 = P$ . Значения  $\gamma_1^2$  и  $\gamma_2^2$  при различных значениях  $m$  и надежности  $P$  рассчитаны заранее и приведены в справочниках. Аналогичным образом можно

оценить и дисперсию окончательного результата. Так как  $\frac{-2}{\gamma_x^2} = \frac{\sigma_x^2}{n}$ ,  $\frac{-2}{S_x^2} = \frac{S_x^2}{n}$ , то  $\gamma_1 \cdot \bar{S}_x \leq \bar{\sigma}_x \leq \gamma_2 \cdot \bar{S}_x$ .

**Пример 3.4.** В серии из десяти измерений ( $n=10$ ) было получено  $\bar{S}_x^2 = 5,4 \cdot 10^{-3} \text{ мм}^2$ . Требуется определить, в каком интервале находится фактическое значение дисперсии  $\sigma$ .

При  $P=0,9$ ,  $m=n-1=9$ , поэтому на основе (3.24) и (3.25) имеем

$$\gamma_1^2 = \frac{m}{\chi_1^2} = \frac{9}{16,919} = 0,532 \quad \text{и} \quad \gamma_2^2 = \frac{m}{\chi_2^2} = \frac{9}{3,325} = 2,707.$$

Значения  $\chi_1^2$  с надежностью  $\alpha_1=(1-P)/2=0,05$  и  $\chi_2^2$  с надежностью  $\alpha_2=(1+P)/2=0,95$  табулированы и получены нами с помощью статистической функции ХИ2ОБР из электронных таблиц Microsoft Excel (см. п. 7.1). Следовательно,  $0,532 \cdot 5,4 \cdot 10^{-3} \leq \sigma \leq 2,707 \cdot 5,4 \cdot 10^{-3}$  или  $2,87 \cdot 10^{-3} \leq \sigma \leq 1,46 \cdot 10^{-2}$ . Аналогично для  $P=0,95$  получаем  $2,55 \cdot 10^{-3} \leq \sigma \leq 1,80 \cdot 10^{-2}$ .

### 3.5. Сравнение двух рядов наблюдений

При проведении и анализе результатов экспериментальных исследований часто в инженерной практике приходится сравнивать две партии изделий, показания двух или нескольких приборов, анализировать результаты работы однотипных металлургических агрегатов и печей, сравнивать результаты исследований проб материалов и др. Вот некоторые из них.

1. Требуется определить правильно ли показывают приборы, если одновременно осуществляется измерение эталонным прибором, а затем измерение этой же величины осуществили другими приборами и получили ряды наблюдений. Равно ли их математическое ожидание действительному значению измеряемого параметра?

2. Необходимо сравнить показания двух приборов, измеряющих одну и ту же величину. Получено два ряда наблюдений. Одинакова ли точность измерения одного и того же параметра разными приборами?

3. Два агрегата выпускают одну и ту же продукцию. Необходимо сделать вывод о том, какой из них лучше в каком-либо смысле.

Решение задач осуществляется с использованием статистических гипотез.

#### 3.5.1. Сравнение средних значений

Задачи 1 и 2 решаются с помощью гипотезы об единичном среднем.

**Задача 1.** Сформулируем нуль-гипотезу  $H_0: M_x=M_{\text{э}}$ , т.е. математическое ожидание по выборке  $M_x$  равно математическому ожиданию эталона  $M_{\text{э}}$ .

Решение осуществляется с помощью t-критерия Стьюдента. Для этого рассчитывается экспериментальное (эмпирическое) значение этого критерия по выражению

$$t_{\text{эксп}} = \frac{\bar{x} - M_{\text{э}}}{S_x} \cdot \sqrt{n}. \quad (3.26)$$

В дальнейшем определяется теоретическое значение этого критерия при заданном уровне значимости  $\alpha$  и числе степеней свободы  $m=n-1$ . Для этого опять же можно воспользоваться таблицами из статистических справочников или пакетами прикладных программ для ЭВМ. Характеристики отдельных таких программ, их возможности мы будем изучать в дальнейшем (см. главу 7), а навыки их использования осваивать на практических занятиях.

Если  $t_{\text{эксп}} \geq t_{\alpha,m}$ , то нуль гипотеза отвергается. Физически это означает, что  $\bar{x}$  не покрывается доверительным интервалом  $M_x$  с вероятностью  $P=1-\alpha$ .

Таким образом, алгоритм решения задачи следующий:

1. По выборочным данным объема выборки  $\bar{x}$  находятся  $S_x^2$ .
2. Рассчитывается эмпирическое значение критерия Стьюдента  $t_{\text{эксп}}$  по выражению (3.26).
3. При  $\alpha=0,05$  и  $m=n-1$  определяется теоретическое значение критерия  $t_{\alpha,m}$ .
4. Если  $|t_{\text{эксп}}| \geq t_{\alpha,m}$ , то гипотеза не принимается.

Для решения подобных задач удобно использовать компьютерные пакеты прикладных программ, которые содержат разнообразные тесты по оценке статистических гипотез (см. главу 7). В частности, в пакет электронных таблиц Microsoft Excel входит специальная функция ТТЕСТ, которая определяет вероятность, что две выборки взяты из генеральных совокупностей, имеющих одно и то же среднее.

**Пример 3.5.** При проверке Ph-метра с помощью эталонного раствора, имеющего Ph=9,0, получили следующие результаты: 8,7; 9,2; 9,1; 9,0; 9,4; 9,6; 9,7; 8,9; 8,8; 8,7; 9,8; 9,3; 9,8; 8,8, т.е.  $n=14$ . Обладает ли Ph-метр систематической погрешностью?

Решение задачи. Предварительно рассчитаем среднеарифметическое  $\bar{x}$  и выборочное среднеарифметическое отклонение  $S_x$

$$\bar{x} = \frac{14}{\sum_{i=1}^{14} x_i} / 14 = 9,2; S_x = \sqrt{\frac{14}{\sum_{i=1}^{14} (x_i - M_x)^2} / 14} = 0,439.$$

В дальнейшем рассчитывается эмпирическое значение критерия

$$t_{\text{эксп}} = \frac{9,2 - 9,0}{0,439} \sqrt{14} = 1,7$$

и определяется табличное значение критерия  $t_{0,05;13} = 2,16$ . Так как  $t_{\text{эксп}} < t_{\alpha,m}$ , то с вероятностью 95% можно считать, что имеющееся различие между показаниями Ph-метра и эталонным значением вызвана случайными причинами и в частности ограниченностью числа измерений.

Аналогичный результат можно получить, если воспользоваться статистической функцией TTEST из электронных таблиц Microsoft Excel (см. п. 7.1).

**Задача 2.** Сравниваются показания двух рядов измерений  $x_i^{(1)}$  и  $x_i^{(2)}$ , например двух приборов, измеряющих один и тот же параметр при одинаковом числе наблюдений  $n$  для каждого из приборов. Требуется определить значимо ли их различие в среднем результатов измерений?

Вспользуемся нуль-гипотезой об единичном среднем, для чего выдвигается гипотеза  $H_0$ : "Математическое ожидание разности показаний двух приборов должно быть равно нулю".

Алгоритм решения состоит в следующем:

1. Строится новая выборка из  $n$  элементов, определяемая как разность значений первой и второй выборок:  $x_i = x_i^{(1)} - x_i^{(2)}$ .

2. По выборке рассчитываются оценки математического ожидания  $\bar{x}$  и среднеквадратичного отклонения  $S_x$ :

$$\bar{x} = \frac{1}{n} \cdot \sum_{i=1}^n x_i; S_x = \sqrt{\frac{1}{n-1} \cdot \sum_{i=1}^n (x_i - \bar{x})^2}.$$

3. Составляется гипотеза об единичном среднем  $H_0: M_x = 0$ .

4. Рассчитывается экспериментальное значение критерия Стьюдента. Учитывая, что  $M_x = 0$ , выражение (3.26) примет вид:

$$t_{\text{эксп}} = \frac{\bar{x} - 0}{S_x} \sqrt{n} = \frac{\bar{x}}{S_x} \sqrt{n}. \quad (3.27)$$

В дальнейшем решение осуществляется так же, как и в предыдущем случае.

**Пример 3.6.** Проведены испытания механической прочности проб окатышей (кг/окатыш) при использовании новой и старой технологий их обжига. Всего исследовано по 8 проб. Результаты испытаний представлены в таблице 3.2.

Таблица 3.2

Результаты испытаний прочности окатышей, изготовленных по старой и новой технологии

Новая технология $x_i^{(1)}$ , кг/окатыш	Старая технология $x_i^{(2)}$ , кг/окатыш	$x_i = x_i^{(1)} - x_i^{(2)}$ , кг/окатыш
215	220	-5
217	223	-6
223	220	+3
210	205	+5
222	230	-8
217	220	-3
220	207	+13
227	218	+9
$\bar{x}^{(1)} = 218,9$	$\bar{x}^{(2)} = 217,9$	$\bar{x} = 1,0$

1. Сформулируем нуль-гипотезу  $H_0: M_x = 0$ .

2. Оценка среднеквадратичного отклонения выборки  $S_x$  составит:

$$S_x = \sqrt{\frac{(-5-1)^2 + (-6-1)^2 + (3-1)^2 + (5-1)^2 + (-8-1)^2 + (-3-1)^2 + (13-1)^2 + (9-1)^2}{8-1}} = 7,65 \frac{\text{кг}}{\text{окатыш}}.$$

3. Экспериментальное значение критерия Стьюдента равно:

$$t_{\text{эксп}} = \frac{1-0}{7,65} \sqrt{8} = 0,37.$$

4. Теоретическое его значение при уровне значимости 0,05 и числе степеней свободы  $m=8-1=7$  определяем по таблицам математических статистик  $t_{0,05;7}=2,36$  (см. также функцию СТЬЮДРАСПОБР из электронных таблиц Microsoft Excel в п. 7.1).

Так как  $t_{\text{эксп}} < t_{\alpha,m}$ , то нуль-гипотеза принимается. Таким образом с вероятностью 95% можно утверждать, что новая технология по среднему значению механической прочности не лучше старой.

Решение данной задачи с использованием функции проверки статистических гипотез ТТЕСТ, входящей в электронные таблицы Microsoft Excel, рассмотрено в п.7.1.

**Задача 3.** Осуществляется оценка двух независимых рядов наблюдений, когда число наблюдений различно. Методику решения подобных задач поясним на примере.

**Пример 3.7.** В процессе проведения исследований из партии бетона, замешанной в понедельник, взято восемь проб ( $n_1=8$ ) и они подвергнуты испытаниям на сжатие. Получены следующие данные прочности на сжатие,  $x_i^{(1)}$ , кг/см<sup>2</sup>: 305,6; 270,8; 298,0; 218,6; 273,3; 270,8; 219,4; 265,8.

Из партии бетона, замешанной в среду, имелась возможность взять 17 проб ( $n_2=17$ ) и после испытаний получены следующие результаты,  $x_i^{(2)}$ , кг/см<sup>2</sup>: 298,0; 263,4; 288,2; 300,7; 327,9; 303,1; 278,2; 296,0; 316,3; 290,7; 318,0; 270,0; 305,6; 320,5; 293,2; 285,5; 316,3. Утверждается, что состав бетона и методика испытаний не изменялись. Выборки 1 и 2 независимы и не взаимосвязаны.

1. Определяются среднеарифметические значения параметра и оценки их дисперсий по каждой из выборок:

$$\bar{x}^{(1)} = 265,3; \bar{x}^{(2)} = 298,3; S_x^2{}^{(1)} = 1015,1; S_x^2{}^{(2)} = 329,1.$$

2. Определяется средневзвешенная дисперсия (суммарная, выборочная дисперсия)  $(S_x^2)_{\Sigma}$ :

$$(S_x^2)_{\Sigma} = \frac{(n_1 - 1) \cdot S_x^2{}^{(1)} + (n_2 - 1) \cdot S_x^2{}^{(2)}}{n_1 + n_2 - 2} = 537,9; (S_x)_{\Sigma} = 23,2 \text{ кг/см}^2. \quad (3.28)$$

3. Проверяется нуль-гипотеза "Математические ожидания первой и второй выборок одинаковы".  $H_0: M_x^{(1)} = M_x^{(2)}$ .

4. Рассчитывается эмпирическое значение критерия Стьюдента

$$t_{\text{эксп}} = \frac{\bar{x}^{(2)} - \bar{x}^{(1)}}{(S_x)_{\Sigma}} \sqrt{\frac{n_1 \cdot n_2}{n_1 + n_2}} = \frac{298,3 - 265,3}{23,2} \sqrt{\frac{8 \cdot 17}{8 + 17}} = 3,32. \quad (3.29)$$

4. Число степеней свободы определяются по выражению

$$m_{\Sigma} = n_1 + n_2 - 2 = 23 \quad (3.30)$$

и определяется теоретическое значений критерия Стьюдента  $t_{0,05;23}=2,07$  (по статистической справочной таблице или с помощью функции СТЬЮДРАСПОБР из электронных таблиц Microsoft Excel).

5. Так как  $t_{\text{эксп}} > t_{\alpha,m}$ , то с вероятностью 95% гипотеза не принимается.

Таким образом, справедливость гипотезы, согласно которой обе партии бетона и методика исследования одинаковы по их средним значениям на сжатие, весьма сомнительна (с вероятностью 95%).

Решение данной задачи с использованием функции проверки статистических гипотез ТТЕСТ, входящей в электронные таблицы Microsoft Excel, рассмотрено в п. 7.1.

### 3.5.2. Сравнение двух дисперсий

При выполнении измерений в различных условиях возникает задача сравнения не только средних значений измеряемых параметров, но и их разброса. Так для последнего примера при исследовании свойств бетона мы установили, что  $S_x^{2(1)}=1015,1$ , а  $S_x^{2(2)}=329,1$ . Существенно ли это различие?

Этот же подход можно применять и для проверки однородности ряда дисперсий, т.е. проверки того, что все эмпирические дисперсии  $S_1^2, S_2^2, \dots, S_k^2$  относятся к выборкам из совокупности с одной и той же теоретической дисперсией  $\sigma_x^2$ . В этих случаях приходится сравнивать не только средние значения, но и изменчивость, или "размах" двух или большего числа выборочных данных.

Для ответа на эти вопросы можно воспользоваться путем применения так называемого критерия F (обозначенного так по первой букве фамилии английского математика Р.Фишера), который называют иногда дисперсионным отношением.

**Критерий F – это отношение двух дисперсий (большей к меньшей), вычисленных или полученных различными способами:**

$$F_{\text{эксп}} = \frac{S_{\text{больш.}}^2}{S_{\text{меньш.}}^2}. \quad (3.31)$$

Очевидно, что его значения всегда не меньше единицы. Так для последнего примера он равен  $F_{\text{эксп}}=1015,1/329,1=3,08$ .

Вероятности (P) получения любого данного значения F, если в действительности две дисперсии ( $S_x^{2(1)}, S_x^{2(2)}$ ) не являются различными, рассчитаны Фишером и представлены в виде таблиц в справочной литературе по математической статистике как функции числа степеней свободы для двух выборок данных; для их определения можно воспользоваться и пакетами прикладных программ для ПЭВМ (см. главу 7).

Иными словами, **имеются значения критерия  $F_{\alpha;m_1;m_2}$ , определенные Фишером, которые показывают во сколько раз максимум могут отличаться дисперсии двух рядов наблюдений при данных числе степеней свободы  $m_1$  и  $m_2$  и уровня значимости  $\alpha$ , когда можно считать, что между этими дисперсиями нет значимого с доверительной вероятностью ( $P=1-\alpha$ ) различия.**

Проиллюстрируем применение критерия Фишера на примере анализа разброса прочности на сжатие образцов бетона, результаты измерений которой приведены ранее (см. пример 3.7).

1. Выдвигается нуль-гипотеза "Дисперсии проб бетона в первой и второй партии одинаковы". Заметим, что сами дисперсии нам не известны, знаем только их весьма приближенные (грубые) оценки  $S_x^{2(1)}$  и  $S_x^{2(2)}$ ; число наблюдений, выборки для первого и второго ряда относительно малы:  $n_1=8, n_2=17$ . Таким образом, основная гипотеза  $H_0: \sigma_x^{2(1)}=\sigma_x^{2(2)}$ . Альтернативная ей гипотеза  $H_1: \sigma_x^{2(1)} \neq \sigma_x^{2(2)}$ .

2. Определяется число степеней свободы  $m_1$  и  $m_2$  по соотношениям

$$m_1 = n_1 - 1, m_2 = n_2 - 1. \quad (3.32)$$

Заметим, что число степеней свободы  $m_1$  относится к большей выборочной дисперсии  $S_x^{2(1)}$ , а  $m_2$  – к меньшей выборочной дисперсии  $S_x^{2(2)}$ .

3. Определяется теоретическое значение критерия Фишера при заданном уровне значимости  $\alpha=0,05$  (надежности  $P=0,95$ ) и числе степеней свободы  $m_1$  и  $m_2$ , т.е. величина  $F_{\alpha;m_1;m_2}$ .

Для примера с бетоном теоретическое значение критерия Фишера мы определим с помощью статистической функции ФРАСПОБР из электронных таблиц Microsoft Excel (см. п. 7.1):  $F_{0,05;7;16}=\text{ФРАСПОБР}(0,05;7;16)=2,66$ .

4. Если  $F_{\text{эксп}} < F_{\alpha;m_1;m_2}$ , то нуль-гипотеза выполняется, т.е. дисперсии выборок однородны, в противном случае нуль-гипотеза об однородности дисперсий отклоняется.

В данном случае  $F_{\text{эксп}} > F_{\alpha;m_1;m_2}$ , т.е. имеются основания сомневаться в том, что эти две дисперсии соответствуют одной и той же совокупности. Другими словами, с вероятностью 95% можно утверждать, что или марка бетона, или методики отбора и испытания проб были различны.

Рассмотрим еще один пример.

**Пример 3.8.** Пусть измеряем одну и ту же величину (температуру, давление, состав газа и т.п.). Старым измерительным прибором проведено 200 измерений, которые дали выборочную дисперсию  $S_x^{2(1)}=3,82$ , а вторым (новым) выполнено 15 измерений при выборочной дисперсии

$S_x^{2(2)}=2,00$ . Можно ли считать, что новый прибор по разбросу показаний дает существенно лучшую точность, чем старый?

Выдвигается гипотеза о равенстве дисперсий.  $H_0: \sigma_x^{2(1)}=\sigma_x^{2(2)}$ . Альтернативная ей гипотеза  $H_1: \sigma_x^{2(1)}>\sigma_x^{2(2)}$ . Далее определяем:

$$F_{\text{эксп}}=3,82/2,00=1,91; m_1=199; m_2=14; F_{0,05;199;14}=2,16; F_{\text{эксп}}<F_{\text{теор}}$$

Таким образом, с вероятностью 95% нет оснований считать, что результаты измерений нового прибора лучше старого.

Как изменится наш вывод, если мы увеличим число измерений новым прибором до 50 при той же выборочной дисперсии?

Теоретическое значение критерия Фишера в этом случае будет равно  $F_{0,05;199;49}=1,49$  и  $F_{\text{эксп}}>F_{\text{теор}}$ , т.е. результаты измерений новым прибором лучше, чем старым! Следовательно, принимается альтернативная гипотеза  $H_1$ .

### 3.5.3. Проверка однородности нескольких дисперсий

Критерий Фишера используется для сравнения двух дисперсий, однако в реальной инженерной практике приходится сравнивать не две, а несколько дисперсий. Пусть среди нескольких серий измерений обнаружена такая серия, выборочная дисперсия которой  $S_{x_{\text{max}}}^2$  заметно больше всех остальных. Задача заключается в том, чтобы выяснить, можно ли считать отличие выделенной дисперсии  $S_{x_{\text{max}}}^2$  не существенным. Другими словами, можно ли считать все эти дисперсии однородными и использовать их в дальнейшем для анализа результатов измерений. Если нет, то серию измерений с максимальной выборочной дисперсией следует отбросить для дальнейшего анализа или выяснить причины таких отклонений.

Произведя в каждой из  $n$  серий одинаковое число  $m^*$  измерений (иногда называемых параллельными измерениями), определим эмпирические (выборочные) дисперсии  $S_{x_1}^2, S_{x_2}^2, \dots, S_{x_n}^2$  для каждой серии измерений

$$S_{x_i}^2 = \frac{\sum_{j=1}^{m^*} (x_{ij} - \bar{x}_i)^2}{m^* - 1}, \quad (3.33)$$

где  $\bar{x}_i = \frac{\sum_{j=1}^{m^*} x_{ij}}{m^*}$  и выделим из них максимальную  $S_{x_{\text{max}}}^2$  (рис.3.5).

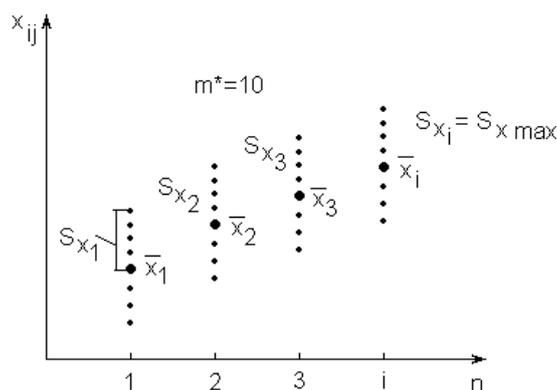


Рис.3.5. К понятию дисперсий эксперимента

Для оценки однородности ряда дисперсий используется статистический критерий Кохрена  $G$ , который рассчитывается как отношение  $S_{x_{\text{max}}}^2$  к сумме всех дисперсий:

$$G_{\text{эж}} = \frac{S_{x_{\text{max}}}^2}{\sum_{i=1}^n S_{x_i}^2}. \quad (3.34)$$

В дальнейшем определяется теоретическое значение этого критерия, которое зависит от уровня значимости  $\alpha$ , числа степеней свободы  $m$  и числа серии измерений  $n$  (числа приборов) –  $G_{\alpha;m;n}$ .

Число степеней свободы рассчитывается как  $m=m^*-1$ , где  $m^*$  – число параллельных измерений в каждой серии.

Если  $G_{\text{эксп}} < G_{\alpha;m;n}$ , то отличия не существенны и ряды дисперсий однородны, их можно использовать для дальнейшего анализа результатов эксперимента.

Так, при однородности дисперсий можно определить оценку дисперсии, характеризующей воспроизводимость всего эксперимента. Она вычисляется по данным параллельных измерений и называется дисперсией воспроизводимости. Можно встретить и другой термин "суммарная дисперсия воспроизводимости":

$$S_{\text{восп}}^2 = \frac{\sum_{i=1}^n S_{xi}^2}{n} \quad (3.35)$$

и вычислить среднеквадратичное отклонение измеряемого параметра в эксперименте при наличии параллельных измерений.

$$S_{\text{восп}} = \sqrt{S_{\text{восп}}^2} = \sqrt{\frac{\sum_{i=1}^n S_{xi}^2}{n}} \quad (3.36)$$

**Пример 3.9.** Пусть шестью ( $n=6$ ) приборами произведено по семь измерений ( $m^*=7$ ) одного и того же параметра, которые дали следующие дисперсии,  $S_{xi}^2$ : 3,82; 1,7; 1,3; 0,92; 0,78; 0,81.

Тогда экспериментальное значение критерия Кохрена в соответствии с уравнением (3.34) составит:

$$G_{\text{эксп}} = \frac{3,82}{3,82 + 1,7 + 1,3 + 0,92 + 0,78 + 0,81} = 0,409.$$

Теоретическое значение этого критерия при уровне значимости  $\alpha=0,05$ , числе степеней свободы  $m=7-1=6$  и серии измерений  $n=6$  составит  $G_{\alpha;m;n}=0,418$ .

Так как  $G_{\text{эксп}} < G_{\alpha;m;n}$ , отклонение дисперсии  $S_{x_{\text{max}}}^2=3,82$  от остальных с вероятностью 95% несущественно, а дисперсии однородны. Дисперсия воспроизводимости (ее оценка) равна

$$S_{\text{восп}}^2 = \frac{\sum_{i=1}^n S_{xi}^2}{n} = \frac{9,33}{6} = 1,56,$$

а среднее квадратичное отклонение измерений составит

$$S_{\text{восп}} = \sqrt{1,56} = 1,25.$$

Попытайтесь самостоятельно оценить при каком значении  $S_{xi}^2$  дисперсия этого измерения будет неоднородна.

Критерий Кохрена можно использовать только в тех случаях, когда во всех точках имеется одинаковое число повторных экспериментов. Если же число измерений в различных сериях неодинаково, то для проверки однородности дисперсий можно использовать критерий Бартлета. При необходимости с процедурой его использования можно познакомиться в литературе по теории вероятности и математической статистике.

В заключении этого параграфа рассмотрим комплексный пример для оценки как средних значений, так и разброса значений двух рядов наблюдений.

**Пример.3.10.** После 100 испытаний материала на прочность было установлено, что  $\bar{x}^{(1)}=150$  кг/см<sup>2</sup>,  $S_x^{(1)}=20$  кг/см<sup>2</sup>. Техническим отделом завода, предложено ввести некоторые изменения в рецептуру материала, технологию его обработки для увеличения его прочности (увеличения  $\bar{x}$ ) и однородности (снижения  $S_x$ ). По новым рекомендациям было отобрано 5 образцов материала, результаты которых представлены в таблице 3.3. Требуется оценить доверительный интервал

изменения прочности материала и ответить на вопрос: обладает ли новый материал лучшими свойствами по прочности и их разбросу по сравнению со старым?

Таблица 3.3

Результаты испытаний на прочность различных образцов материала

Номер образца	Прочность, кг/см <sup>2</sup> , $x_i^{(2)}$	$x_i^{(2)} - \bar{x}^{(2)}$	$(x_i^{(2)} - \bar{x}^{(2)})^2$
1	160	+5	25
2	150	-5	25
3	140	-15	225
4	155	0	0
5	170	+15	225
$\Sigma$	775	0	500

$$\bar{x}^{(2)} = \frac{1}{5} \sum_{i=1}^5 x_i^{(2)} = 155; \quad (S_x^2)^{(2)} = \frac{1}{4} \sum_{i=1}^4 (x_i^{(2)} - 155)^2 = 500/4 = 125;$$

$$S_x^{(2)} = \sqrt{(S_x^2)^{(2)}} = \sqrt{125} = 11,2 \text{ кг/см}^2.$$

1. Для вычисления доверительного интервала определим теоретическое (табличное) значение критерия Стьюдента:  $t_{\alpha; m} = t_{0,05; 5-1} = 2,78$ .

Тогда

$$\delta = \frac{t_{\alpha; m} \cdot S_x^{(2)}}{\sqrt{n_2}} = \frac{2,78 \cdot 11,2}{\sqrt{5}} = 14 \text{ кг/см}^2.$$

С вероятностью 95% можно утверждать, что математическое ожидание прочности материала находится в диапазоне:  $141 \leq M_x^{(2)} \leq 169 \text{ кг/см}^2$ .

2. Сформулируем нуль-гипотезу: дисперсия в первой выборке  $(\sigma_x^2)^{(1)}$  равна дисперсии во второй выборке  $(\sigma_x^2)^{(2)}$ , т.е.  $(\sigma_x^2)^{(1)} - (\sigma_x^2)^{(2)} = 0$ .

$H_0: (\sigma_x^2)^{(1)} = (\sigma_x^2)^{(2)}$ . Альтернативная гипотеза:  $H_1: (\sigma_x^2)^{(1)} > (\sigma_x^2)^{(2)}$ .

Экспериментальное значение критерия Фишера составит:

$$F_{\text{эксп}} = \left( S_x^2 \right)^{(1)} / \left( S_x^2 \right)^{(2)} = 20^2 / 11,2^2 = 3,2.$$

При числе степеней свободы  $m_1=99$  и  $m_2=4$ , уровне значимости  $\alpha=0,05$  теоретическое значение критерия Фишера равно  $F_{0,05; 99; 4} = 5,66$ .

Поскольку  $F_{\text{эксп}} < F_{\alpha; m_1; m_2}$  приходится признать, что с вероятностью 95% данные выборки не дают основания считать новый материал более однородным по механической прочности.

Оценим достоверность повышения в среднем механической прочности. Сформулируем нуль-гипотезу "Математическое ожидание прочности металла при старой и новой технологии не изменилось", т.е.  $H_0: M_{x1} = M_{x2}$ . Альтернативная гипотеза "Математическое ожидание прочности металла по новой технологии возросло", т.е.  $H_1: M_{x1} > M_{x2}$ . Определим средневзвешенную оценку дисперсии (суммарную выборочную дисперсию)

$$S_{x\Sigma}^2 = \frac{(n_1 - 1) \left( S_x^2 \right)^{(1)} + (n_2 - 1) \left( S_x^2 \right)^{(2)}}{n_1 + n_2 - 2} = \frac{99 \cdot 20^2 + 4 \cdot 11,2^2}{100 + 5 - 2} = 389;$$

$$S_{x\Sigma} = \sqrt{S_{x\Sigma}^2} = \sqrt{389} = 19,7.$$

Рассчитаем эмпирическое значение критерия Стьюдента

$$t_{\text{эксп}} = \frac{\left| \bar{x}^{(1)} - \bar{x}^{(2)} \right|}{S_{x\Sigma}} \sqrt{\frac{n_1 \cdot n_2}{n_1 + n_2}} = \frac{|150 - 155|}{19,7} \sqrt{\frac{100 \cdot 5}{100 + 5}} = 0,55.$$

Суммарное число степеней свободы составит  $m_\Sigma = n_1 + n_2 - 2 = 103$ .

Теоретическое значение критерий Стьюдента при уровне значимости 0,05 и числе степеней свободы 103 равно  $t_{0,05;103}=1,98$ .

Так как  $t_{\text{эксп}} < t_{\alpha,m}$ , то можно с вероятностью 95% утверждать, что в среднем механическая прочность материала не повысилась.

Таким образом, нельзя с вероятностью 95% утверждать улучшение прочностных свойств материала как в среднем, так и по их разбросу при использовании новой технологии.

### 3.6. Определение необходимого количества измерений

Увеличение количества измерений (числа проб, образцов и т.п., как видно из выражений (3.18), (3.21) даже при неизменной их точности ( $\sigma_x = \text{const}$ ,  $S_x = \text{const}$ ) может увеличить надежность доверительной оценки (P) или сузить доверительный интервал для определения действительного значения измеряемой величины (математического ожидания)  $\delta$ .

Необходимое количество измерений (образцов, проб и т.п.)  $n$  для достижения требуемой точности  $\delta$  и заданной надежности P можно определить заранее в том случае, когда известно действительное значение среднеквадратичного отклонения  $\sigma_x$ , а экспериментальные данные (измерения) подчиняются нормальному закону распределения.

Действительно, при этих допущениях число измерений можно определить из выражения (3.18)

$$n \geq \left( \frac{\varepsilon(\alpha) \cdot \sigma_x}{\delta} \right)^2 = \varepsilon^2(\alpha) \cdot \left( \frac{\sigma_x}{\delta} \right)^2. \quad (3.37)$$

Таким образом, число измерений  $n$  определяется требуемой надежностью P (доверительной вероятностью  $\alpha$ ) и относительным (по отношению к среднеквадратичному отклонению) значением доверительного интервала ( $\sigma_x/\delta$ ), т.е. требуемой точностью определения измеряемой величины. Так при  $P=0,95$ ,  $\varepsilon(\alpha)=1,96$  и при  $\delta=\sigma_x$  число измерений равно 4. При увеличении необходимой точности измерений в 2 раза, т.е. сужении доверительного интервала до величины  $\delta=(1/2)\sigma_x$ , необходимое число измерений составит 16. **Нетрудно заметить, что необходимое число измерений с увеличением точности возрастает в квадратичной зависимости.**

Как правило, действительное значение среднеквадратичной ошибки ( $\sigma_x$ ) неизвестно, а имеется только ее оценка ( $S_x$ ). В этом случае следует воспользоваться соотношением (3.21), т.е. критерием Стьюдента. Необходимое число измерений следует определять из следующего соотношения

$$n \geq \frac{t_{\alpha,m}^2 \cdot S_x^2}{\delta^2} = t_{\alpha,m}^2 \cdot \left( \frac{S_x}{\delta} \right)^2. \quad (3.38)$$

При расчетах по этому уравнению следует иметь в виду, что теоретическое значение критерия Стьюдента зависит не только от доверительной вероятности  $\alpha$ , но и числа степеней свободы  $m$ , последние же определяются числом измерений. В связи с этим уравнение (3.38) следует решать методом последовательных приближений. В качестве начального приближения можно задать, в частности, число измерений, рассчитанных по формуле (3.37). Так, если решить последнее уравнение методом последовательных приближений, то можно показать, что при  $P=0,95$  ( $\alpha=0,05$ ) для определения доверительного интервала с точностью  $\delta=S_x$  требуется 7 измерений, а с точностью  $\delta=0,5S_x$  – 19. С повышением необходимой точности различие в числе измерений, рассчитанных по соотношениям (3.37) и (3.38) уменьшается и, как показывают расчеты, при величине  $\delta \leq 0,2S_x$  они практически совпадают.

В порядке упражнений постарайтесь определить, какое количество проб чугуна следует отобрать, чтобы повысить надежность доверительного интервала содержания кремния в чугуне до 0,99 и понизить до 0,90? Какое количество проб следует отобрать, чтобы при той же надежности повысить точность (сократить доверительный интервал) определения содержания кремния в чугуне в два раза? Необходимое количество измерений различных значениях величины доверительного интервала  $\delta/S_x$  и  $\delta/\sigma_x$  (в скобках) приведено в табл. 3.4.

Таблица 3.4  
Необходимое количество измерений

$\delta/S_x$	P=0,90	P=0,95	P=0,99
1	5	7 (4)	11
0,5	13	19 (16)	31
0,4	19	27 (24)	46
0,3	32	46 (48)	78
0,1	273	387 (384)	668

### 3.7. Проверка гипотезы нормального распределения

При построении таблиц распределения все экспериментально полученные значения сначала располагают в порядке возрастания  $x_1 < x_2 < \dots < x_n$ .

Если число различных значений опытных данных более 15-20, то при группировке все эти данные целесообразно сводить в определенное число разрядов  $k$ , принимая ширину разряда одинаковой.

Построение таблиц ведется в следующей последовательности.

1. Находят наибольшее ( $x_{\max}$ ) и наименьшее ( $x_{\min}$ ) значения случайной величины и вычисляют размах параметра  $R = x_{\max} - x_{\min}$ .

2. Задаются некоторым числом интервалов  $k$ , принимая его нечетным. При общем числе измерений  $n > 100$ ,  $k = 9 \div 15$ ;  $n < 100$ ,  $k = 7$ . Иногда число интервалов определяется по приближенной формуле Штюргеса  $k = 1 + 3,32 \lg(n)$  с округлением числа интервалов до ближайшего целого значения.

3. Определяют ширину интервала  $h = R/k$ , для упрощения расчетов полученные значения округляют в любую сторону.

4. Устанавливают границы интервалов и подсчитывают число событий (замеров) в каждом интервале  $\bar{m}_i$ ,  $1 \leq i \leq k$ .

5. Определяют вероятность для каждого интервала как  $P_i = \bar{m}_i / n$ . Результаты вычислений представлены в таблице 3.5.

Графической формой представления непрерывной случайной величины является гистограмма (рис.3.6).

Последовательность построения гистограмм следующая:

1. Определяется величина ординаты  $f_i = P_i/h$ , где  $P_i$  – вероятность появления случайной величины в  $i$ -м интервале.

2. В системе координат  $f_i=f(x)$  на ширине интервала  $h$  откладывают величины  $f_i$  как высоты и строятся прямоугольники.

Очевидно, что площадь элементарного прямоугольника

$$S_i = h \cdot f_i = h \cdot \frac{P_i}{h} = P_i = \frac{\bar{m}_i}{n}$$

равна отношению числа опытов  $\bar{m}_i$ , при которых случайная величина оказалась внутри этого интервала, к общему числу опытов  $n$ .

Таблица 3.5  
Построение распределения экспериментальных данных

Интервал	Число замеров в каждом интервале, $\bar{m}_i$	Вероятность интервала $P_i = \bar{m}_i/n$
$x_1 \div x_2$	$\bar{m}_1$	$\bar{m}_1/n$
$x_2 \div x_3$	$\bar{m}_2$	$\bar{m}_2/n$
...	...	...
$x_i \div x_{i+1}$	$\bar{m}_i$	$\bar{m}_i/n$
...	...	...
$x_k \div x_{k+1}$	$\bar{m}_k$	$\bar{m}_k/n$
Проверка	$\sum_{i=1}^k \bar{m}_i = n$	$\sum_{i=1}^k P_i = 1$

Площадь всей гистограммы  $S = \sum_{i=1}^k S_i = \sum_{i=1}^k P_i = 1$ . Следовательно, площадь, ограниченная

гистограммой, равна единице.

3. Построение гистограммы интегральной функции распределения осуществляется

суммированием вероятностей:  $F(x) = \sum_{i=1}^k P_i$ .

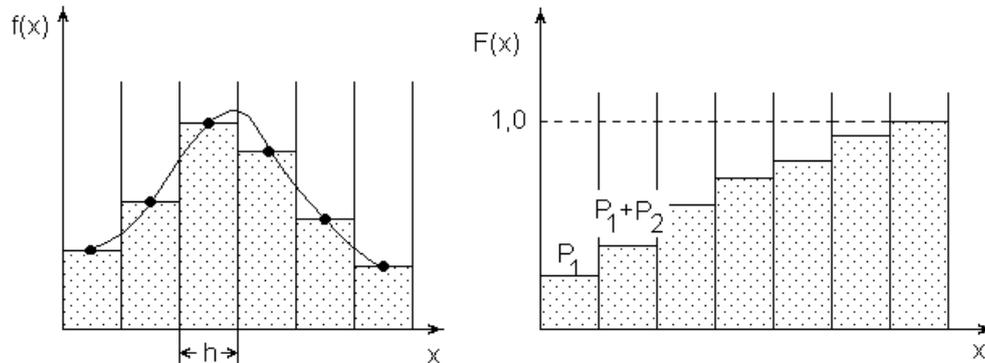


Рис.3.6. К построению гистограммы случайной величины

В дальнейшем осуществляется сравнение экспериментально полученного распределения случайной величины с некоторым видом теоретического распределения. Для этой цели используются различные критерии согласия:  $\chi^2$  (хи-квадрат) Пирсона, Колмогорова–Смирнова и др.

Рассмотрим методику проверки гипотезы нормального распределения по  $\chi^2$  критерию Пирсона. Этот критерий кроме определения доверительного интервала для дисперсии нередко **используется для проверки согласованности распределений, полученных по данным выборки с некоторой теоретической плотностью распределения.**

В данном случае применение критерия  $\chi^2$  предполагает использование свойств стандартного нормального распределения. Напомним, что уравнение кривой плотности нормального распределения имеет вид:

$$f(u) = \frac{1}{\sqrt{2\pi}} e^{-u^2/2} \approx 0,4 \cdot e^{-u^2/2}; \quad u = \frac{x - M_x}{\sigma_x}.$$

Тогда теоретическая вероятность попадания случайной величины в интервал  $\Delta u = u_2 - u_1$  в случае нормального распределения можно определить по формуле:

$$P_i^* = F(u_2) - F(u_1) = \frac{1}{2\pi} \int_{u_1}^{u_2} e^{-z^2/2} dz, \quad u = \frac{x - M_x}{\sigma_x}.$$

Таблицы нормированной функции нормального распределения приведены в справочной литературе, можно воспользоваться также компьютерными пакетами прикладных программ (см. главу 7).

Отличие оценки закона распределения  $P$  от теоретического закона распределения  $P^*$  можно охарактеризовать величиной

$$\chi^2 = \sum_{i=1}^k C_i (P_i - P_i^*)^2, \quad (3.39)$$

где  $P_i$  и  $P_i^*$  – оценка и теоретическая вероятность случайной величины для  $i$ -го интервала;  $C_i$  – весовые коэффициенты, которые с большим весом учитывают отклонения для меньших  $P_i$ .

Пирсон выбрал весовые коэффициенты следующим образом

$$C_i = \frac{n}{P_i^*} \quad (3.40)$$

и показал, что при таком выборе  $C_i$  закон распределения  $\chi^2$  слабо зависит от  $n$  и  $P(x)$ , а определяется в основном числом разрядов  $k$ .

Следовательно,

$$\chi_{\text{эксп}}^2 = n \sum_{i=1}^k \frac{(P_i - P_i^*)^2}{P_i^*} = n \sum_{i=1}^k \frac{(\bar{m}_i/n - P_i^*)^2}{P_i^*} = \sum_{i=1}^k \frac{(\bar{m}_i - n \cdot P_i^*)^2}{n \cdot P_i^*}. \quad (3.41)$$

Очевидно, что в идеальном соответствии экспериментальных данных нормальному закону распределения экспериментальное значение критерия Пирсона будет равно нулю, т.к.  $P_i = P_i^*$ .

В выражении (3.41) стоит сумма квадратов  $k$  случайных величин, однако они не являются независимыми, так как на них накладывается некоторое число связей. Одной из таких связей является требование, чтобы площадь под кривой оценки закона распределения равнялась единице

$\sum_{i=1}^k P_i = 1$ . Иногда требуют, чтобы среднее значение  $\bar{x}$  совпадало с математическим ожиданием  $M_x$ , а

выборочная дисперсия  $S_x^2$  с дисперсией  $\sigma_x^2$ . Поэтому число степеней свободы чаще всего определяется как:

$$m = k - 2. \quad (3.42)$$

Теоретическое значение критерия Пирсона определяется по справочным данным или с использованием пакетов для ПЭВМ при заданном уровне значимости  $\alpha$  и числе степеней свободы  $m$ , т.е.  $\chi_{\text{теор } \alpha; m}^2$  (см. функцию ХИ2ОБР из электронных таблиц Microsoft Excel в п. 7.1). Алгоритм использования критерия Пирсона заключается в следующем.

1. Выдвигается нуль-гипотеза  $H_0$ : "Отличие экспериментальных данных от нормального закона распределения не существенно" и альтернативная ей гипотеза  $H_1$ : "Отличие экспериментальных данных от нормального закона распределения существенно, т.е. экспериментальные данные не подчиняются закону нормального распределения".

2. По результатам экспериментальных измерений и предположению нормального закона их распределения определяется расчетное значение критерия Пирсона  $\chi_{\text{эксп}}^2$ .

3. Определяют число степеней свободы  $m$ , задаются уровнем значимости  $\alpha$  и определяют теоретическое значение критерия Пирсона  $\chi_{\text{теор}}^2$ .

4. Если  $\chi_{\text{эксп}}^2 < \chi_{\text{теор}}^2$ , то нуль-гипотеза  $H_0$  о нормальном законе распределения экспериментальных данных принимается с доверительной вероятностью  $P=1-\alpha$ . В противном случае нуль-гипотеза отвергается и принимается альтернативная гипотеза  $H_1$ .

Отметим важные рекомендации по использованию критерия  $\chi^2$ .

Если распределение отличается от нормального, то при достаточно большом числе измерений  $\chi_{\text{эксп}}^2$  превысит соответствующее критическое значение  $\chi_{\alpha; m}^2$ . Поэтому, если при произвольном числе измерений критерий  $\chi_{\text{эксп}}^2$  дал малую надежность, но сомнение в нормальности распределения осталось, то следует, если имеется возможность, увеличить число измерений в несколько раз и повторить анализ по этому же критерию.

Число степеней свободы  $m=k-2$  относится к такому случаю, когда оба параметра нормального закона распределения определяются по результатам измерений, т.е. когда вместо точных измерений значений  $M_x$  и  $\sigma_x$  применяют их эмпирические значения (оценки)  $\bar{x}$  и  $S_x$ . Если же значение  $M_x$  точно известно (например, при измерении эталона), то число степеней свободы равно  $k=n-1$ ; если известны оба параметра  $M_x$  и  $\sigma_x$ , то число степеней свободы равно  $k=n$ . На практике такая ситуация встречается относительно редко, и поэтому для получения числа степеней свободы не менее пяти желательно брать число интервалов не менее семи (иногда девяти).

Рассмотрим использование критерия Колмогорова–Смирнова для проверки гипотезы нормальности распределения случайных величин. Данная процедура также предполагает построение таблицы распределения экспериментально полученных значений с группировкой данных в определенное число разрядов  $k$ . Дополнительно в таблицу необходимо включить следующие колонки:

– колонку с накопительной суммой  $\sum_{p=1}^i m_p$  для каждого  $i$ -го интервала сгруппированных данных

как показано в табл.3.6.

– колонку с накопительной суммой  $n \sum_{p=1}^i P_p^*$ , где  $P^*$  – теоретическая вероятность попадания случайной величины в  $i$ -й интервал, значения которой, как уже отмечалось ранее, табулированы и приводятся в статистических справочниках.

На основании данных табл.3.6 вычисляют экспериментальное значение критерия согласия Колмогорова–Смирнова

$$D_{\text{эксп}} = \frac{\max \left| \sum_{p=1}^i \bar{m}_p - n \sum_{p=1}^i P_p^* \right|}{n}$$

Далее экспериментальное значение критерия Колмогорова–Смирнова сравнивают с теоретическим  $D_{n;\alpha}$ , которое определяют из статистических таблиц в соответствии с объемом выборки  $n$  и требуемым уровнем значимости  $\alpha$ . Если  $D_{\text{эксп}} < D_{n;\alpha}$ , то гипотеза о нормальном распределении результатов замеров принимается с вероятностью  $P=1-\alpha$ .

Таблица 3.6

Процедура вычисления критерия Колмогорова–Смирнова

Интервал	Число замеров в каждом интервале, $\bar{m}_i$	$\sum_{p=1}^i \bar{m}_p$	Теоретическая вероятность $P_i^*$	$n \sum_{p=1}^i P_p^*$	$\left  \sum_{p=1}^i \bar{m}_p - n \sum_{p=1}^i P_p^* \right $
$X_1 \div X_2$	$\bar{m}_1$	$\bar{m}_1$	$P_1^*$	$nP_1^*$	$ \bar{m}_1 - nP_1^* $
$X_2 \div X_3$	$\bar{m}_2$	$\bar{m}_1 + \bar{m}_2$	$P_2^*$	$n(P_1^* + P_2^*)$	$ \bar{m}_1 + \bar{m}_2 - n(P_1^* + P_2^*) $
...	...	...	...	...	...
$X_i \div X_{i+1}$	$\bar{m}_i$	$\sum_{p=1}^i \bar{m}_p$	$P_i^*$	$n \sum_{p=1}^i P_p^*$	$\left  \sum_{p=1}^i \bar{m}_p - n \sum_{p=1}^i P_p^* \right $
...	...	...	...	...	...
$X_k \div X_{k+1}$	$\bar{m}_k$	$\sum_{p=1}^k \bar{m}_p$	$P_k^*$	$n \sum_{p=1}^k P_p^*$	$\left  \sum_{p=1}^k \bar{m}_p - n \sum_{p=1}^k P_p^* \right $

Критерий согласия Колмогорова–Смирнова для проверки нормальности распределения результатов наблюдений входит в пакет статистической обработки данных STATISTICA, с которым мы познакомимся в главе 7 (см. п. 7.2).

Рассмотрим использование рассмотренных критериев  $\chi^2$  Пирсона и Колмогорова–Смирнова для оценки нормальности распределения данных на следующем примере.

**Пример 3.11.** В таблице 3.7 приведено содержание кремния в чугуна при выплавке передельного чугуна в печи, которое изменяется в пределах от 0,32% до 0,95%. Всего было отобрано 50 проб чугуна.

Таблица 3.7

Содержание кремния в чугуна по результатам отбора 50 проб

№ пробы	1	2	3	4	5	6	7	8	9	10
[Si],%	0,32	0,35	0,45	0,43	0,41	0,51	0,52	0,53	0,57	0,58
№ пробы	11	12	13	14	15	16	17	18	19	20
[Si],%	0,59	0,56	0,56	0,58	0,54	0,57	0,61	0,62	0,63	0,64
№ пробы	21	22	23	24	25	26	27	28	29	30

[Si],%	0,65	0,66	0,67	0,68	0,69	0,61	0,65	0,62	0,63	0,67
№ пробы	31	32	33	34	35	36	37	38	39	40
[Si],%	0,65	0,62	0,68	0,71	0,72	0,78	0,75	0,72	0,79	0,72
№ пробы	41	42	43	44	45	46	47	48	49	50
[Si],%	0,73	0,72	0,79	0,73	0,84	0,82	0,87	0,90	0,95	0,93

Требуется оценить, подчиняется ли содержание кремния в пробах нормальному закону распределения? Если да, определить медиану, моду, среднее, выборочную дисперсию, доверительный интервал для математического ожидания, среднее квадратичное отклонение. Для вычислений воспользуемся статистическими функциями из электронной таблицы Microsoft Excel (см. п. 7.1).

Предварительно вычислим с использованием статистических функций СРЗНАЧ, ДИСП и СТАНДОТКЛОН среднее значение  $\bar{x}$ , выборочную дисперсию  $S_x^2$  и стандартное отклонение  $S_x$ , которые оказались равны соответственно  $\bar{x}=0,65$ ,  $S_x^2=0,01853$  и  $S_x=0,1361$ .

Примем число интервалов, равным 7. Тогда величина интервала составит  $h=(0,95-0,32)/7=0,09=0,1$ . Результаты группировки исходных данных и вычислений приведены в табл. 3.8. Вычисление  $F(x_i)$  проводили с использованием статистической функции НОРМРАСП.

Таким образом, экспериментальное значение критерия Пирсона равно  $\chi_{\text{эксп}}^2 = \sum_{i=1}^k \chi_i^2 = 1,96$ , а

теоретическое при уровне значимости  $\alpha=0,05$  и числе степеней свободы  $m_1=7-2=5$  составляет  $\chi_{0,05;5}^2=11,07$ , что значительно больше экспериментального значения. Следовательно, весьма уверенно можно утверждать, что содержание кремния в пробах чугуна подчиняется нормальному закону распределения. Предлагаем читателям самостоятельно оценить доверительный интервал математического ожидания. Здесь можно отметить, что, как показывают расчеты, доверительный интервал с вероятностью 95% равен 0,030 (учитывая близость распределения к нормальному и достаточно большое число экспериментальных точек, можно воспользоваться электронными таблицами Microsoft Excel (функция ДОВЕРИТ), т.е. действительное среднее содержание кремния в чугуне (математическое ожидание) лежит в интервале от 0,62% до 0,68%.

Таблица 3.8

Процедура вычисления критерия  $\chi^2$  Пирсона по данным примера 3.11

Интервал $x_{i-1} \div x_i$	$\bar{m}_i$	$F(x_i)_{0,05}$	$P_i^* = F(x_i) - F(x_{i-1})$	$nP_i^*$	$\bar{m}_i - nP_i^*$	$\chi_i^2 = \frac{(\bar{m}_i - nP_i^*)^2}{nP_i^*}$
0,3÷0,4	2	0,033	0,033	1,7	0,4	0,07
0,4÷0,5	3	0,135	0,102	5,1	-2,1	0,87
0,5÷0,6	11	0,356	0,221	11,1	-0,1	0,00
0,6÷0,7	17	0,642	0,286	14,3	2,7	0,51
0,7÷0,8	11	0,864	0,222	11,1	-0,1	0,00
0,8÷0,9	4	0,967	0,103	5,2	-1,2	0,26
0,9÷1,0	2	0,995	0,028	1,4	0,6	0,26

Процедура проверки гипотезы нормального распределения данных из примера с использованием критерия Колмогорова–Смирнова (D) представлена в табл.3.9. На основании результатов табл.3.9 определяем максимальное значение из последней колонки и по нему рассчитываем экспериментальное значение критерия  $D_{\text{эксп}}$ :

$$D_{\text{эксп}} = \frac{\max \left| \sum_{p=1}^i \overline{m}_p - n \sum_{p=1}^i P_p^* \right|}{n} = \frac{1,8}{50} = 0,036.$$

Таблица 3.9

Процедура вычисления критерия Колмогорова–Смирнова  
по данным примера 3.11

Интервал $x_{i-1} \div x_i$	$\overline{m}_i$	$\sum_{p=1}^i \overline{m}_p$	$F(x_i)$	$P_i^* = F(x_i) - F(x_{i-1})$	$nP_i^*$	$n \sum_{p=1}^i P_p^*$	$\left  \sum_{p=1}^i \overline{m}_p - n \sum_{p=1}^i P_p^* \right $
0,3÷0,4	2	2	0,033	0,033	1,7	1,7	0,4
0,4÷0,5	3	5	0,135	0,102	5,1	6,8	1,8
0,5÷0,6	11	16	0,356	0,221	11,1	17,8	1,8
0,6÷0,7	17	33	0,642	0,286	14,3	32,1	0,9
0,7÷0,8	11	44	0,864	0,222	11,1	43,2	0,8
0,8÷0,9	4	48	0,967	0,103	5,2	48,4	0,3
0,9÷1,0	2	50	0,995	0,028	1,4	49,8	0,3

Экспериментальное значение критерия сравниваем в теоретическим  $D_{\text{теор}}$ , взятым из статистической таблицы (см. книгу Львовский Е.Н. Статистические методы построения эмпирических формул. М.: Высшая школа, 1988. 239 с.) с учетом объема выборки  $n=50$  и уровнем значимости  $\alpha=0,05$ . Величина  $D_{\text{теор}}=0,177 > 0,036 = D_{\text{эксп}}$ , поэтому можно сделать тот же вывод, что и ранее: гипотеза нормального распределения результатов измерения содержания кремния в чугуна принимается с вероятностью 95%.

Пример 3.12. В таблице 3.10 приведено содержание оксида железа в шлаке перед раскислением металла при выплавке стали марки 0,8 КП в 200-тонной мартеновской печи, работающей с применением кислорода для интенсификации горения топлива и прямого окисления примесей ванны. Всего отобрано 56 проб шлака.

Таблица 3.10

Исходные данные к примеру 3.12

1	$x_i$	$\overline{m}_i$	$\sum_{i=1}^l \overline{m}_i$
1	10	1	1
2	12	5	6
3	13	3	9
4	14	6	15
5	15	6	21
6	16	2	23
7	17	8	31
8	18	6	37
9	19	6	43
10	20	4	47
11	22	1	48
12	23	4	52
13	24	1	53
14	25	1	54
15	26	1	55

16	30	1	56
17	31	0	56

$$\bar{x} = \frac{1}{17} \frac{\sum_{i=1}^{17} x_i \cdot m_i^*}{\sum_{i=1}^{17} m_i} = 17,375;$$

$$S_x^2 = \frac{1}{\sum_{i=1}^{17} m_i - 1} \sum_{i=1}^{17} m_i \cdot (x_i - \bar{x})^2 = 16,0;$$

$$S_x = 4,0.$$

Из табл. 3.10 видно, что содержание оксида железа в шлаке колеблется от 10% до 31%. Примем число интервалов равным  $k=7$ . Тогда величина интервала составит  $h=(31-10)/7=3$ . Результаты группировки исходных данных и вычислений приведены в таблице 3.11. Расчетное значение критерия Пирсона составило  $\chi^2_{\text{эсп}}=4,8$ . Его следует сравнить с табличным значением  $\chi^2_{\alpha;m}$  для 5% уровня значимости при числе степеней свободы равном  $m=5$  составляет  $\chi^2_{0,05;5}=11,07$ .

Процедура вычисления критерия  $\chi^2$  по данным примера 3.12

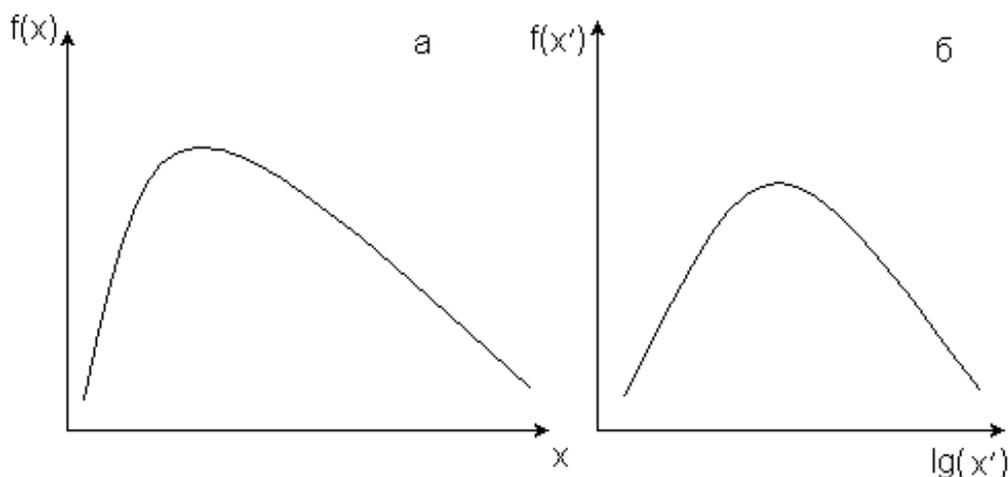
Интервал $x_{i-1} \div x_i$	$\bar{m}_i$	$F(x_i)_{0,05}$	$P_i = F(x_i) - F(x_{i-1})$	$nP_i$	$m_i - nP_i$	$(m_i - nP_i)^2$	$\chi_i^2 = \frac{(m_i - nP_i)^2}{nP_i}$
10÷13	6	0,0326 0,137	0,104	5,824	0,176	0,031	0,005
13÷16	15	0,366	0,229	12,824	2,176	4,735	0,369
16÷19	16	0,658	0,292	16,352	0,352	0,124	0,008
19÷22	10	0,876	0,218	12,208	2,208	4,875	0,399
22÷25	6	0,972	0,096	5,376	0,624	0,389	0,072
25÷28	2	0,996	0,024	1,344	0,656	0,430	0,320
28÷31	1	1,000	0,004	0,224	0,776	0,602	2,688
Сумма	56	-	-	-	-	-	$\Sigma=3,861$

Таким образом, опытные данные не противоречат гипотезе о том, что содержание FeO в конечном шлаке подчиняется закону нормального распределения. Оцените самостоятельно моду, медиану, доверительный интервал для математического ожидания при вероятностях 90, 95 и 99,7%.

### 3.8. Преобразование распределений к нормальному

Если исследователь использовал методы, изложенные в предыдущем параграфе, убедился, что гипотеза нормальности распределения не может быть принята, то вполне может быть, что с помощью существующих методов удастся так преобразовать исходные данные, что их распределение будет подчиняться нормальному закону распределения. Для пояснения идеи преобразований рассмотрим качественный пример. Пусть кривая распределения  $f(x)$  имеет вид, представленный на рис. 3.7, т.е. имеется очень крутая левая ветвь и пологая правая. Такое распределение отличается от нормального.

Для выполнения операций преобразования каждое наблюдение трансформируется с помощью логарифмического преобразования  $x' = \lg(x)$ . При этом левая ветвь кривой распределения сильно растягивается и распределение принимает приближенно нормальный вид. Если при преобразовании получаются значения, расположенные между 0 и 1, то все наблюдаемые значения для удобства расчетов и во избежание получения отрицательных параметров необходимо умножить на 10 в соответствующей степени, чтобы все вновь полученные, преобразованные значения были больше единицы, т.е. необходимо выполнить преобразования  $x'' = \lg(x \cdot 10^a)$ .



*Рис. 3.7. Преобразование функции  $f(x)$  к нормальному распределению*

Асимметричное распределение с одной вершиной приводится к нормальному преобразованием  $x' = \lg(x \pm a)$ . В отдельных случаях можно применять и другие преобразования:

а) обратная величина  $x' = 1/x$ ;

б) обратное значение квадратных корней  $x' = 1/\sqrt{x}$ .

Преобразование "обратная величина" является наиболее "сильным". Среднее положение между логарифмическим преобразованием и "обратной величиной" занимает преобразование "обратное значение квадратных корней".

Для нормализации смещенного вправо распределения служат, например, степенные преобразования  $x' = x^a$ . При этом для  $a$  принимают значения:  $a=1,5$  при умеренном и  $a=2$  при сильно выраженном правом смещении. Рекомендуем читателю придумать такие преобразования, которые удовлетворяли бы исследователя в том или ином случае.

Вопросы к теме 3

1. Вычисление характеристик эмпирических распределений.
2. Виды статистических гипотез.
3. Методы отсева грубых погрешностей.
4. Определение доверительных интервалов для исследуемых величин.
5. Оценка доверительного интервала для математического ожидания.
6. Оценка доверительного интервала для дисперсии.
7. Сравнение двух рядов наблюдений.
8. Сравнение средних значений.
9. Сравнение двух дисперсий.
10. Проверка однородности нескольких дисперсий.
11. Определение необходимого количества измерений.
12. Проверка гипотезы нормального распределения.
13. Преобразование распределений к нормальному.

## **Тема 4. АНАЛИЗ РЕЗУЛЬТАТОВ ПАССИВНОГО ЭКСПЕРИМЕНТА. ЭМПИРИЧЕСКИЕ ЗАВИСИМОСТИ**

### ***4.1. Характеристика видов связей между рядами наблюдений***

На практике сама необходимость измерений большинства величин вызывается тем, что они не остаются постоянными, а изменяются в функции от изменения других величин. В этом случае целью проведения эксперимента является установление вида функциональной зависимости  $y=f(X)$ . Для этого должны одновременно определяться как значения  $X$ , так и соответствующие им значения  $y$ , а задачей эксперимента является установление математической модели исследуемой зависимости. Фактически речь идет об установлении связи между двумя рядами наблюдений (измерений).

Определение связи включает в себя указание вида модели и определения ее параметров. В теории экспериментов независимые параметры  $X=(x_1, \dots, x_n)$  принято называть факторами, а зависимые переменные  $y$  – откликами. Координатное пространство с координатами  $x_1, x_2, \dots, x_i, \dots, x_n$  называется факторным пространством. Эксперимент по определению вида функции

$$y = f(x), \quad (4.1)$$

где  $x$  – скаляр, называется однофакторным. Эксперимент по определению функции вида

$$y = f(X), \quad (4.1a)$$

где  $X=(x_1, x_2, \dots, x_i, \dots, x_k)$  – вектор, – многофакторным.

Геометрическое представление функции отклика в факторном пространстве является поверхностью отклика. При однофакторном эксперименте  $k=1$  поверхность отклика представляет собой линию на плоскости, при двухфакторном  $k=2$  – поверхность в трехмерном пространстве.

Связи в общем случае являются достаточно многообразными и сложными. Обычно выделяют следующие виды связей.

**Функциональные связи** (или зависимости). Это такие связи, когда при изменении величины  $X$  другая величина  $Y$  изменяется так, что каждому значению  $x_i$  соответствует совершенно определенное (однозначное) значение  $y_i$  (рис.4.1а). Таким образом, если выбрать все условия эксперимента абсолютно одинаковыми, то повторяя испытания получим одну и ту же зависимость, т.е. кривые идеально совпадут для всех испытаний.

К сожалению, таких условий в реальности не встречается. На практике не удается поддерживать постоянство условий (например, колебания физико-химических свойств шихты при моделировании процессов тепломассопереноса в металлургических печах). При этом влияние каждого случайного фактора в отдельности может быть мало, однако в совокупности они существенно могут повлиять на результаты эксперимента. В этом случае говорят о стохастической (вероятностной) связи между переменными.



**Рис.4.1.** Виды связей: а – функциональная связь, все точки лежат на линии; б – связь достаточно тесная, точки группируются возле линии регрессии, но не все они лежат на ней; в – связь слабая

**Стохастичность связи** состоит в том, что одна случайная переменная  $Y$  реагирует на изменение другой  $X$  изменением своего закона распределения (см. рис. 4.1б). Таким образом, зависимая переменная принимает не одно конкретное значение, а некоторое из множества значений. Повторяя испытания мы будем получать другие значения функции отклика и одному и тому же значению  $x$  в различных реализациях будут соответствовать различные значения  $y$  в интервале  $[x_{\min}; x_{\max}]$ . Искомая зависимость  $y = f(x)$  может быть найдена лишь в результате совместной обработки полученных значений  $x$  и  $y$ .

На рис.4.1б эта кривая зависимости, проходящая по центру полосы экспериментальных точек (математическому ожиданию), которые могут и не лежать на искомой кривой  $Y = f(x)$ , а занимают некоторую полосу вокруг нее. Эти отклонения вызваны погрешностями измерений, неполнотой модели и учитываемых факторов, случайным характером самих исследуемых процессов и другими причинами.

При анализе стохастических связей можно выделить следующие основные типы зависимостей между переменными.

1. Зависимости между одной случайной переменной  $X$  от другой случайной переменной  $Y$  и их условными средними значениями называются **корреляционными**. Условное среднее  $\bar{y}_i$  – это среднее арифметическое для реализации случайной величины  $Y$  при условии, что случайная величина  $X$  принимает значение  $\bar{x}_i$ .

2. Зависимость случайной переменной  $Y$  от неслучайной переменной  $X$  или зависимость математического ожидания  $M_y$  случайной величины  $Y$  от детерминированного значения  $X$  называется **регрессионной**. Приведенная зависимость характеризует влияние изменений величины  $X$  на среднее значение величины  $Y$ .

Стохастические зависимости характеризуются формой, теснотой связи и численными значениями коэффициентов уравнения регрессии.

Форма связи устанавливает вид функциональной зависимости  $y=f(X)$  и характеризуется уравнением регрессии. Если уравнение связи линейное, то имеем линейную многомерную регрессию, в этом случае зависимости  $Y$  от  $X$  описываются уравнением прямой линии в  $k$ -мерном пространстве

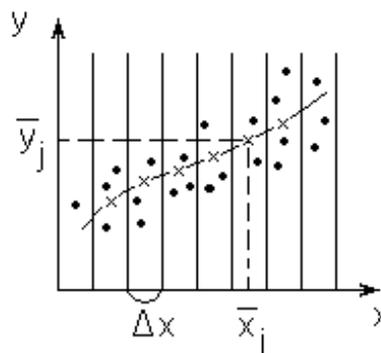
$$\hat{y} = b_0 + \sum_{j=1}^k b_j x_j, \quad (4.2)$$

где  $b_0, \dots, b_j, \dots, b_k$  – коэффициенты уравнения. Для пояснения существа используемых методов ограничимся сначала случаем, когда  $x$  скаляр. В общем случае виды функциональных зависимостей в технике достаточно многообразны: показательные  $y = b_0 x^{b_1}$ , логарифмические  $y = b_0 \lg(x)$  и т.д.

Заметим, что задача выбора вида функциональной зависимости – задача неформализуемая, т.к. одна и та же кривая на данном участке примерно с одинаковой точностью может быть описана самыми различными аналитическими выражениями. Отсюда следует важный практический вывод. **Даже в наш век ПЭВМ принятие решения о выборе той или иной математической модели остается за исследователем.** Только экспериментатор знает, для чего будет в дальнейшем использоваться эта модель, на основе каких понятий будут интерпретироваться ее параметры.

**Крайне желательно при обработке результатов эксперимента вид функции  $\hat{y}=f(X)$  выбирать исходя из условия соответствия физической природе изучаемых явлений или имеющимся представлениям об особенностях поведения исследуемой величины. К сожалению, такая возможность не всегда имеется, так как эксперименты чаще всего проводятся для исследования недостаточно или неполно изученных явлений.**

При изучении зависимости  $\hat{y}=f(X)$  от одного фактора при заранее неизвестном виде функции отклика для приближенного определения вида уравнения регрессии полезно предварительно построить эмпирическую линию регрессии (рис.4.2). Для этого весь диапазон изменения  $x$  на поле корреляции разбивают на равные интервалы  $\Delta x$ . Все точки, попавшие в данный интервал  $\Delta x_j$ , относят к его середине  $\bar{x}_j$ . Для этого подсчитывают частные средние для каждого интервала



*Рис.4.2. Корреляционное поле*

$$\bar{y}_j = \frac{\sum_{i=1}^{n_j} y_{ji}}{n_j}. \quad (4.3)$$

Здесь  $n_j$  – число точек в интервале  $\Delta x_j$ , причем  $\sum_{j=1}^{k^*} n_j = n$ , где  $k^*$  – число интервалов разбиения,  $n$  – объем выборки.

Затем последовательно соединяют точки  $(\bar{x}_j; \bar{y}_j)$  отрезками прямой. Полученная ломаная называется эмпирической линией регрессии. По виду эмпирической линии регрессии можно в первом приближении подобрать вид уравнения регрессии  $\hat{y}=f(X)$ .

Под теснотой связи понимается степень близости стохастической зависимости к функциональной, т.е. это показатель тесноты группирования экспериментальных данных относительно принятого уравнения модели (см. рис. 4.1б). В дальнейшем уточним это положение.

#### 4.2. Определение коэффициентов уравнения регрессии

Будем полагать, что вид уравнения регрессии уже выбран и требуется определить только конкретные численные значения коэффициентов этого уравнения  $\mathbf{b}=\{b_0, \dots, b_j, \dots, b_k\}$ . Отметим предварительно, что если выбор вида уравнения регрессии, как это уже отмечалось, процесс неформальный и не может быть полностью передан ПЭВМ, то **расчет коэффициентов выбранного уравнения регрессии – операция достаточно формальная и ее следует решать с использованием ПЭВМ**. Это трудный и утомительный расчет, в котором человек не застрахован от ошибок, а ПЭВМ выполнит его значительно быстрее и качественнее.

Существуют два основных различных подхода к нахождению коэффициентов  $b_j$ . Выбор того или иного из них определяется целями и задачами, стоящими перед исследователем, точностью полученных результатов, их количеством и т.д.

**Первый подход – интерполирование, базируется на удовлетворении условию, чтобы функция  $\hat{y}=(X,b)$  совпадала с экспериментальными значениями в некоторых точках, выбранных в качестве опорных (основных, главных)  $y_i$ .**

В этом случае для определения  $k+1$  неизвестных значений параметров  $b_j$  используется система  $k+1$  уравнений

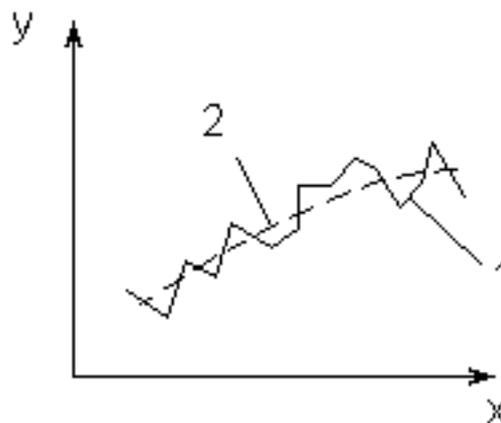
$$f(x_i, b_0, \dots, b_j, \dots, b_k)=y_i, \quad 1 \leq i \leq n. \quad (4.4)$$

В данном случае число независимых уравнений системы равно числу опорных точек, в пределе –  $n$  поставленных опытов. С другой стороны, для определения  $k+1$  коэффициентов необходимо не менее  $k+1$  независимых уравнений. Но если число  $n$  поставленных опытов и число независимых уравнений равно числу искомых коэффициентов  $k+1$ , то решение системы может быть единственно, а следовательно, точно соответствует случайным значениям исходных данных. Таким образом, в предельном случае, когда число коэффициентов уравнения регрессии равно числу экспериментальных точек  $n=k+1$ , все экспериментальные точки будут совпадать с их расчетными значениями. Следует заметить, что добиваться такого точного совпадения путем значительного увеличения числа коэффициентов уравнения регрессии часто просто неразумно, поскольку экспериментальные результаты получены с большей или меньшей погрешностью, и такая функция может просто не отражать действительного характера изменения исследуемой величины в силу влияния помех (возмущений) (рис.4.3).

Таким образом, задача в конечном счете сводится к решению системы  $k+1$  уравнений с  $k+1$  неизвестными. Основная сложность такого решения состоит в нелинейности системы, хотя в принципе при использовании ПЭВМ она преодолима.

При числе опытов  $n$ , большем, чем  $k+1$  искомых коэффициентов, число независимых уравнений системы избыточно. Избыточность информации можно использовать по-разному.

После определения численных значений параметров  $k+1$  проверяется качество аппроксимации путем сопоставления значений функции и экспериментальных данных в оставшихся, не использованных точках. Если обнаруженные между ними расхождения превышают допустимые



**Рис.4.3.** Аппроксимация функции с большим (1) и небольшим (2) числом коэффициентов  $b_i$

по условиям точности, то процедуру определения коэффициентов  $b_j$  можно повторить, приняв в качестве опорных (основных) другие точки.

Из этих уравнений в разных комбинациях можно составить несколько систем уравнений, каждая из которых в отдельности даст свое решение. Но между собой они будут несовместимыми. Каждое решение будет соответствовать своим значениям коэффициентов  $b_j$ . Если все их построить на графике, то получим целый пучок аппроксимирующих кривых.

Это открывает при  $n > k + 1$  совершенно новые возможности. Во-первых, этот пучок кривых показывает форму и ширину области неопределенности проведенного эксперимента. Во-вторых, может быть произведено усреднение всех найденных кривых и полученная усредненная кривая будет гораздо точнее и достовернее описывать исследуемое явление, так как она в значительной степени освобождена от случайных погрешностей, приводивших к разбросу отдельных экспериментальных точек. Поясним суть этого подхода на примере двух методов.

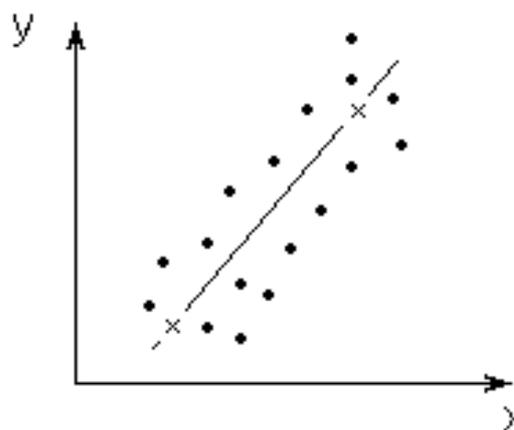


Рис.4.4. Метод избранных точек (x – избранные точки)

1. Метод избранных точек (рис. 4.4). На основании анализа данных выдвигают гипотезу о виде (форме) зависимости  $f(X)$ . Предположим, что она линейная, т.е. статистическая связь – это линейная одномерная регрессия

$$\hat{y} = b_0 + b_1 x. \quad (4.5)$$

Выбирают две наиболее характерные по мнению исследователя точки, через которые и проходит линия регрессии (см. рис. 4.4). Задача вычисления коэффициентов  $b_0$  и  $b_1$  модели в этом случае тривиальная. Если предполагается, что уравнение регрессии более высокого порядка, то соответственно увеличивают число избранных точек. Недостатки такого подхода очевидны. Так, избранные точки выбираются субъективно, а подавляющая часть экспериментального материала не используется для определения параметров (коэффициентов) уравнения регрессии, хотя ее можно использовать в дальнейшем для оценки надежности полученного уравнения.

2. Метод медианных центров. Сущность этого метода поясняет рис.4.5. Обведенное контуром поле точек делят на несколько частей, число которых равно числу определяемых коэффициентов уравнения регрессии. В каждой из этих частей находят медианный центр, т.е. пересечение вертикали и горизонтали слева и справа, выше и ниже которых оказывается равное число точек. Затем через эти медианные центры проводят плавную кривую и из решения системы уравнений определяют коэффициенты регрессии  $b_j$ . Так, в случае линейной зависимости (4.5) поле делится на две группы и определяют средние значения  $\bar{x}_I, \bar{y}_I; \bar{x}_{II}, \bar{y}_{II}$  для каждой из групп, а неизвестные коэффициенты  $b_0, b_1$  определяют из решения системы уравнений:

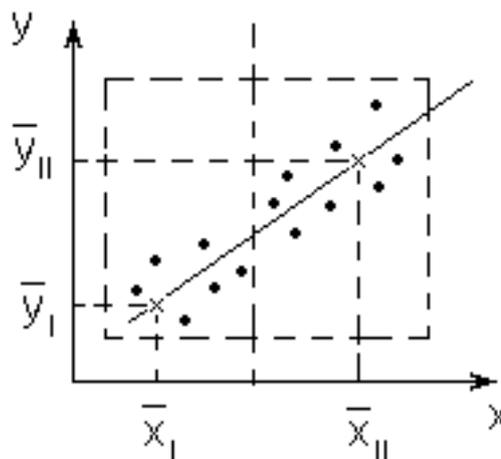


Рис.4.5. Метод медианных точек

$$\bar{y}_I = b_0 + b_1 \bar{x}_I;$$

$$\bar{y}_{II} = b_0 + b_1 \bar{x}_{II}.$$

$$(4.5a)$$

Если при выборе вида уравнения регрессии число его коэффициентов  $b_j$  окажется больше числа уравнений (имеющихся результатов измерений)  $k + 1 > n$ , система (4.4) не будет иметь

однозначного решения, в этом случае необходимо либо уменьшить число определяемых коэффициентов  $k+1$ , либо увеличить число опытов  $n$ , другого выхода здесь нет.

### **Второй подход. Метод наименьших квадратов.**

Усреднение несовместимых решений избыточной системы уравнений  $n > k$  может быть преодолено методом наименьших квадратов, который был разработан еще Лежандром и Гауссом. Таким образом, метод наименьших квадратов – это "новинка" почти 200 летней давности. Сегодня благодаря возможностям ПЭВМ этот метод поступил, по существу, в полосу своего ренессанса.

**Определение коэффициентов  $b_j$  основано на выполнении требования, чтобы сумма квадратов отклонений экспериментальных точек от соответствующих значений уравнения регрессии была минимальна.** Заметим, что в принципе можно оперировать и суммой других четных степеней этих отклонений, но тогда вычисления будут сложнее. Однако руководствоваться суммой невязок нельзя, так как она может оказаться малой при больших отклонениях отрицательного знака.

Математическая запись приведенного выше требования имеет вид

$$\Phi(b_0, b_1, \dots, b_j, \dots, b_k) = \sum_{i=1}^n [f(x_i, b_0, b_1, \dots, b_j, \dots, b_k) - y_i]^2 \rightarrow \min_{b_j} \quad (4.6)$$

где  $n$  – число экспериментальных точек в рассматриваемом интервале изменения аргумента  $x$ .

Необходимым условием минимума функции  $\Phi(b_0, b_1, \dots, b_j, \dots, b_k)$  является выполнение равенства

$$\partial \Phi / \partial b_j = 0, \quad 0 \leq j \leq k \quad (4.7)$$

или

$$\sum_{i=1}^n [f(x_i, b_0, b_1, \dots, b_j, \dots, b_k) - y_i] \frac{\partial f(x_i)}{\partial b_j} = 0, \quad 0 \leq j \leq k. \quad (4.7a)$$

После преобразований получим

$$\sum_{i=1}^n [f(x_i, b_0, b_1, \dots, b_j, \dots, b_k) \frac{\partial f(x_i)}{\partial b_j} - \sum_{i=1}^n y_i \frac{\partial f(x_i)}{\partial b_j}] = 0. \quad (4.8)$$

Система уравнений (4.8) содержит столько же уравнений, сколько неизвестных коэффициентов  $b_0, b_1, \dots, b_k$  входит в уравнение регрессии, и называется в математической статистике системой нормальных уравнений.

Поскольку  $\Phi \geq 0$  при любых  $b_0, \dots, b_k$ , величина  $\Phi$  обязательно должна иметь хотя бы один минимум. Поэтому, если система нормальных уравнений имеет единственное решение, оно и является минимумом для этой величины.

При  $n > k+1$  система имеет единственное решение, при  $n = k+1$  численные значения коэффициентов уравнения регрессии по первому и второму подходам идентичны, а все опытные точки совпадают с уравнением регрессии.

Очевидно, что при  $k+1 > n$  система уравнений (4.8) переопределена и имеет множество решений, преодолеть эту проблему можно, как уже отмечалось, либо увеличением числа наблюдений, либо уменьшением числа неизвестных коэффициентов  $b_j$ .

Расчет коэффициентов по методу наименьших квадратов можно применять при любых статистических данных, распределенных по любому закону.

### **4.3. Определение тесноты связи между случайными величинами**

Определив уравнение теоретической линии регрессии, необходимо дать количественную оценку тесноты связи между двумя рядами наблюдений. Уравнения линий регрессии, проведенные на рис. 4.1бв, одинаковы, однако на рис. 4.1б точки значительно ближе (теснее) расположены к теоретической линии регрессии, чем на рис. 4.1в.

При корреляционном анализе предполагается, что факторы и отклики носят случайный характер и подчиняются нормальному закону распределения.

Тесноту связи между случайными величинами характеризуют корреляционным отношением  $\rho_{xy}$ . Остановимся подробнее на физическом смысле этого показателя. Для определения физического смысла этого показателя введем новые понятия.

Остаточная дисперсия (дисперсия адекватности)  $S_{y\text{ост}}^2$  характеризует разброс экспериментально наблюдаемых точек относительно линии регрессии и представляет собой показатель ошибки предсказания параметра  $Y$  по уравнению регрессии (рис. 4.6):

$$S_{y\text{ост}}^2 = \frac{1}{n-1} \sum_{i=1}^n [y_i - \hat{y}_i]^2 = \frac{1}{n-1-k} \sum_{i=1}^n [y_i - f(x_i, b_0, b_1, \dots, b_k)]^2, \quad (4.9)$$

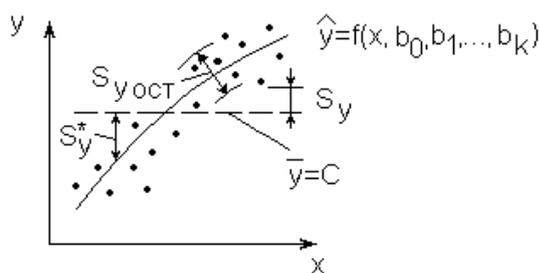
где  $l=k+1$  – число коэффициентов уравнения модели.

Общая дисперсия (дисперсия выходного параметра)  $S_y^2$  характеризует разброс экспериментально наблюдаемых точек относительно среднего значения  $\bar{y}$ , т.е. линии  $C$  (см. рис. 4.6):

$$S_y^2 = \frac{1}{n-1} \sum_{i=1}^n [y_i - \bar{y}]^2, \quad (4.10)$$

где  $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$ .

Средний квадрат отклонения линии регрессии от среднего значения линии  $\bar{y} = C$  (см. рис.4.6):



$$S_y^{*2} = \frac{1}{n-1} \sum_{i=1}^n [y_i - \bar{y}]^2 = \frac{1}{n-1} \sum_{i=1}^n [f(x_i, b_0, b_1, \dots, b_k) - \bar{y}]^2. \quad (4.11)$$

Очевидно, что общая дисперсия  $S_y$  (сумма квадратов относительно среднего значения  $\bar{y}$ ) равна остаточной дисперсии  $S_{y\text{ост}}^2$  (сумме квадратов относительно линии регрессии) плюс средний квадрат отклонения линии регрессии  $S_y^{*2}$  (сумма квадратов, обусловленная регрессией).

$$S_y^2 = S_{y\text{ост}}^2 + S_y^{*2}. \quad (4.11a)$$

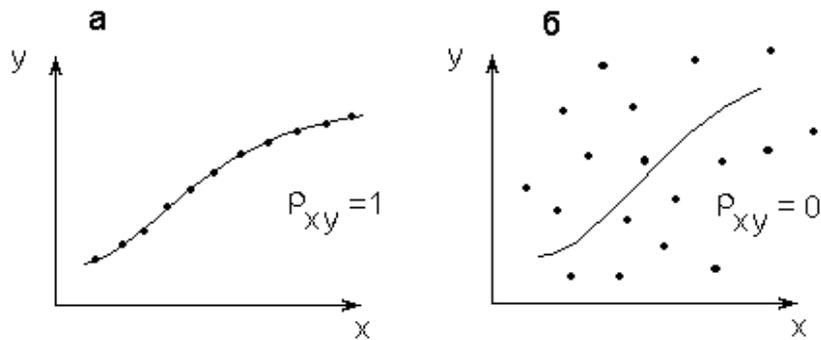
**Разброс экспериментально наблюдаемых точек относительно линии регрессии характеризуется безразмерной величиной – выборочным корреляционным отношением, которое определяет долю, которую привносит величина  $X$  в общую изменчивость случайной величины  $Y$ .**

$$\rho_{xy}^* = \sqrt{\frac{S_y^2 - S_{y\text{ост}}^2}{S_y^2}} = \sqrt{\frac{S_y^{*2}}{S_y^2}} = \frac{S_y^*}{S_y} \quad (4.12)$$

Проанализируем свойства этого показателя.

1. В том случае, когда связь является не стохастической, а функциональной, корреляционное отношение равно 1, т.к. все точки корреляционного поля оказываются на линии регрессии, остаточная дисперсия равна  $S_{y\text{ост}}^2 = 0$ , а  $S_y^{*2} = S_y^2$  (рис.4.7а).

2. Равенство нулю корреляционного отношения указывает на отсутствие какой-либо тесноты связи между величинами  $x$  и  $y$  для данного уравнения регрессии, поскольку разброс экспериментальных точек относительно среднего значения и линии регрессии одинаков, т.е.  $S_{y\text{ост}}^2 = S_y^2$  (рис.4.7б).



**Рис. 4.7.** Значения выборочного корреляционного отношения  $\rho_{xy}$ :  
*а* – функциональная связь; *б* – отсутствие связи

3. Чем ближе расположены экспериментальные данные к линии регрессии, тем теснее связь, тем меньше остаточная дисперсия и тем больше корреляционное отношение.

Следовательно, корреляционное отношение может изменяться в пределах от 0 до 1.

Учитывая, что на ПЭВМ имеются пакеты программ для статистической обработки результатов исследований, рассмотрим методологию этого подхода на примере простейших линейных и одномерных задач (см. уравнение 4.5). Идеология этого способа для решения более сложных задач принципиально не отличается. Более того, как мы увидим в дальнейшем, многие нелинейные зависимости можно свести к линейным.

#### 4.4. Линейная регрессия от одного фактора

Уравнение линии регрессии на плоскости в декартовых координатах имеет вид выражения (4.5).

Задачу метода наименьших квадратов аналитически можно выразить следующим образом:

$$\Phi(b_0, b_1) = \sum_{i=1}^n [y_i - (b_0 + b_1 x_i)]^2 \rightarrow \min_{b_0, b_1}. \quad (4.13)$$

Для решения этой задачи, как известно из математического анализа, необходимо вычислить частные производные функции  $\Phi$  по коэффициентам  $b_0, b_1$  и приравнять их нулю:

$$\frac{\partial \Phi(b_0, b_1)}{\partial b_0} = 0; \quad \frac{\partial \Phi(b_0, b_1)}{\partial b_1} = 0. \quad (4.14)$$

Система нормальных уравнений (4.8) в этом случае примет вид

$$\begin{cases} \sum_{i=1}^n [y_i - (b_0 + b_1 x_i)] = 0; & n b_0 + b_1 \sum_{i=1}^n x_i = \sum_{i=1}^n y_i; \\ \sum_{i=1}^n [y_i - (b_0 + b_1 x_i)] \cdot x_i = 0; & b_0 \sum_{i=1}^n x_i + b_1 \sum_{i=1}^n x_i^2 = \sum_{i=1}^n x_i y_i. \end{cases} \quad (4.15)$$

Решение этой системы относительно  $b_0$  и  $b_1$  дает

$$b_0 = \frac{\sum_{i=1}^n y_i \sum_{i=1}^n (x_i)^2 - \sum_{i=1}^n (x_i y_i) \sum_{i=1}^n x_i}{n \sum_{i=1}^n (x_i)^2 - \left( \sum_{i=1}^n x_i \right)^2}; \quad (4.16a)$$

$$b_1 = \frac{\sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}, \quad (4.16б)$$

т.е. необходимо определить  $\sum x_i, \sum y_i, \sum x_i y_i, \sum (x_i)^2$ .

Коэффициент  $b_0$  (свободный член уравнения регрессии) геометрически представляет собой расстояние от начала координат до точки пересечения линии регрессии с ординатой, а коэффициент  $b_1$  характеризует тангенс угла наклона линии регрессии к оси ОХ.

Если же определяют уравнение регрессии в виде

$$\hat{y} = b_0 + b_1 x + b_{11} x^2,$$

то система уравнений для нахождения  $b_0, b_1, b_{11}$  будет иметь вид

$$\begin{cases} \sum_{i=1}^n y_i = b_0 n + b_1 \sum_{i=1}^n x_i + b_{11} \sum_{i=1}^n x_i^2; \\ \sum_{i=1}^n x_i y_i = b_0 \sum_{i=1}^n x_i + b_1 \sum_{i=1}^n x_i^2 + b_{11} \sum_{i=1}^n x_i^3; \\ \sum_{i=1}^n x_i^2 y_i = b_0 \sum_{i=1}^n x_i^2 + b_1 \sum_{i=1}^n x_i^3 + b_{11} \sum_{i=1}^n x_i^4. \end{cases} \quad (4.16в)$$

Из уравнений (4.15) и (4.16в) вытекает правило записи любых систем: **необходимо записать столько уравнений в системе, сколько неизвестных членов в искомом уравнении, всякий раз суммируя произведения членов исходного уравнения на переменную при искомом коэффициенте.**

Оценку силы линейной связи осуществляют по выборочному (эмпирическому) коэффициенту парной корреляции  $r_{xy}$ . Выборочный коэффициент корреляции может быть вычислен двумя способами.

1. Как частный случай корреляционного отношения для линейного уравнения регрессии.

С учетом того, что  $\bar{y} = b_0 + b_1 \bar{x}$ ,

$$S_y^{*2} = \frac{1}{n-1} \sum_{i=1}^n [b_0 + b_1 x_i - b_0 - b_1 \bar{x}]^2 = b_1^2 S_x^2, \quad (4.17)$$

величина отношения  $S_y^*/S_y$  будет равна

$$r_{xy} = b_1 S_x / S_y, \quad (4.18)$$

где  $S_x$  и  $S_y$  – выборочные средние квадратичные отклонения.

2. Как среднее значение произведения центрированных случайных величин, отнесенное к произведению их среднеквадратичных отклонений:

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{(n-1)S_x S_y} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}. \quad (4.19)$$

Покажем, что обе последние формулы эквивалентны. Для этого преобразуем выражение (4.19) к виду

$$\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = r_{xy}(n-1)S_x S_y.$$

Подставляя последнее выражение в формулу (4.16б), имеем

$$b_1 = \frac{r_{xy}(n-1)S_x S_y}{\sum_{i=1}^n (x_i - \bar{x})^2} = r_{xy} S_y / S_x, \text{ откуда } r_{xy} = \frac{b_1 S_x}{S_y}.$$

Как правило, по результатам экспериментов находят  $S_x$ ,  $S_y$ ,  $\bar{x}$ ,  $\bar{y}$  и рассчитывают  $r_{xy}$  по формуле (4.19), а затем, используя эти величины, определяют коэффициенты уравнения регрессии

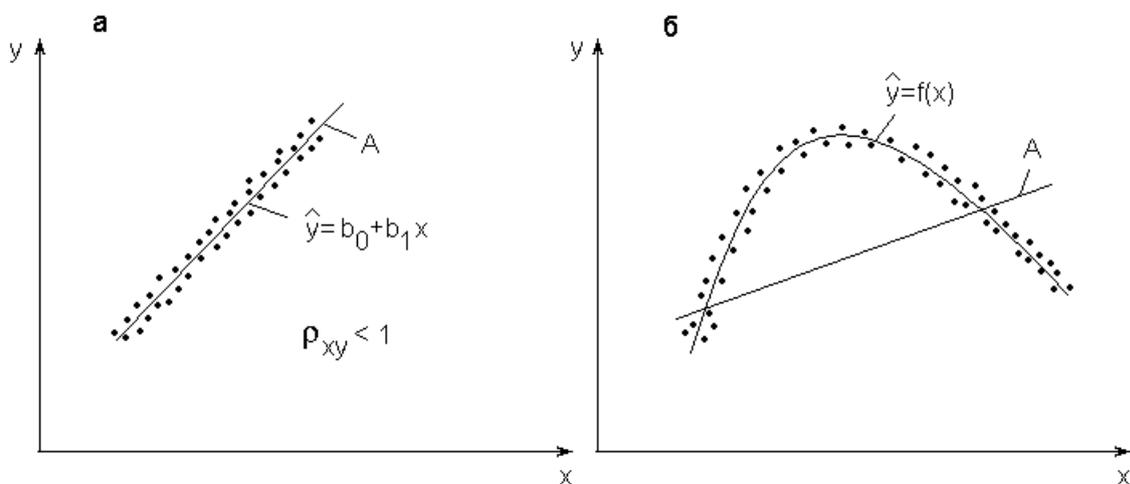
$$b_1 = r_{xy} S_y / S_x; \quad b_0 = \bar{y} - b_1 \bar{x}. \quad (4.20)$$

Коэффициент корреляции  $r_{xy}$  изменяется в пределах  $-1 \leq r_{xy} \leq +1$ .

Положительная корреляция между случайными величинами характеризует такую стохастическую зависимость между величинами, когда с возрастанием одной из них другая в среднем также будет возрастать. При отрицательной корреляции с возрастанием одной случайной величины другая в среднем будет уменьшаться. Чем ближе значение  $r_{xy}$  к единице, тем теснее статистическая связь.

Отметим еще раз область применимости выборочного коэффициента корреляции для оценки тесноты связи.

1. Коэффициент парной корреляции значений  $y$  и  $x$  применительно к однофакторной зависимости характеризует тесноту группирования данных лишь относительно прямой (например линия А на рис.4.8а). При более сложной зависимости (рис.4.8б) коэффициент корреляции  $r_{xy}$  будет оценивать тесноту экспериментальных точек относительно некоторой прямой, обозначенной буквой А, что, естественно, несет мало сведений о тесноте их группирования относительно искомой кривой  $\hat{y} = f(x)$ .



**Рис.4.8.** К понятию коэффициента парной корреляции

2. Коэффициент парной выборочной корреляции имеет четкий физический смысл только в случае двумерного нормального распределения параметров, т.е. когда для каждого значения  $X$  существует совокупность нормального распределения  $Y$  и наоборот, а дисперсия зависимой переменной при изменении значения аргумента остается постоянной (рис.4.9).

Даже при выполнении этих, вообще говоря достаточно жестких условий, не всякое значение выборочного коэффициента корреляции является достаточным для статистического обоснования выводов о наличии действительно надежной корреляционной связи между фактором и откликом. Надежность статистических характеристик ослабевает с уменьшением объема выборки ( $n$ ). Так, при  $n=2$  через две экспериментальные точки можно провести только одну прямую и зависимость будет функциональной, при этом выборочный коэффициент корреляции равен единице  $r_{xy}=1$ . Однако это не означает надежность полученных статистических характеристик в силу весьма и весьма ограниченного объема выборки. Значит, вычислять коэффициент корреляции по результатам двух наблюдений бессмысленно, так как он заведомо будет равен единице, и это будет обусловлено не свойствами переменных и их взаимным отношением, а только числом наблюдений.

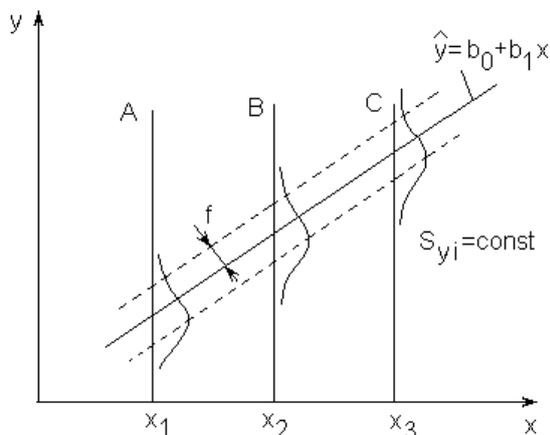


Рис.4.9. К понятию коэффициента парной корреляции в случае двумерного нормального распределения

В связи с этим требуется проверка того, насколько значимо отличается выборочный коэффициент корреляции  $r_{xy}$  от его действительного значения  $r_{xy}^*$ . При достаточно большом объеме выборки  $n \rightarrow N$ , где  $N$  – достаточно большое число,  $r_{xy}^* = r_{xy}$ . Таким образом, требуется проверка значимости выборочного коэффициента парной корреляции и оценка его доверительного интервала.

Для определения значимости  $r_{xy}$  сформулируем нуль-гипотезу  $H_0: r_{xy}=0$ , т.е. корреляция отсутствует. Для этого рассчитывается экспериментальное значение t-критерия Стьюдента

$$t_{\text{эксп}} = \left| r_{xy}^* \right| \frac{\sqrt{n-2}}{\sqrt{1-(r_{xy}^*)^2}} \quad (4.21)$$

и сравнивается с теоретическим при числе степеней свободы  $n-2$ .

Если  $t_{\text{эксп}} \geq t_{\alpha; n-2}$ , то нулевая гипотеза отклоняется, а гипотеза, что коэффициент корреляции существует, принимается.

Для определения доверительного интервала коэффициента корреляции Р.Фишер предложил такое нелинейное преобразование величины  $r_{xy}$ , при котором закон распределения этой оценки, вообще говоря, довольно сложный, практически приближается к нормальному. Это преобразование производится по формуле

$$Z^* = \frac{1}{2} \ln \frac{1 + |r_{xy}|}{1 - |r_{xy}|} \quad (4.22)$$

(см. также статистическую функцию ФИШЕР из электронных таблиц Microsoft Excel в п. 7.1).

Среднеквадратичное отклонение случайной величины  $Z^*$  зависит от числа опытов

$$S_{Z^*} = \frac{1}{\sqrt{n-3}}, \quad (4.23)$$

а математическое ожидание очень близко к числу, получающемуся после подстановки в формулу (4.22) вместо  $r_{xy}$  истинного значения коэффициента корреляции  $r_{xy}^*$ . Эти свойства величины  $Z^*$  позволяют просто оценить, в каких пределах может находиться истинное значение коэффициента корреляции, если по  $n$  опытам получены некоторые значения его выборочного значения (оценки)  $r_{xy}$ .

Если граничное значение  $r_{xy}$  имеет тот же знак, что и  $r_{xy}^*$ , то можно считать в первом приближении, что корреляционная связь между переменными достоверна.

**Пример.4.1.** При обработке 17 пар данных  $x$  и  $y$  получены следующие значения  $n=17$ ;  $\bar{x} = 26$ ;  $\bar{y} = 5,07$ ;  $S_y^2 = 7,8$ ;  $S_x^2 = 102$ . Выборочный коэффициент корреляции  $r_{xy} = -0,94$ , т.е. величина  $y$  связана с  $x$  достаточно сильной причинной связью, близкой к функциональной зависимости.

Определение значимости коэффициента  $r_{xy}$ .

$$t_{\text{эксп}} = \left| r_{xy}^* \right| \frac{\sqrt{n-2}}{\sqrt{1-(r_{xy}^*)^2}} = 0,94 \frac{\sqrt{17-2}}{\sqrt{1-(0,94)^2}} = 10,6;$$

$$t_{0,05;15} = 1,75.$$

Так как  $t_{\text{эксп}} > t_{\alpha;n-2}$ , то коэффициент корреляции существенен.

Определение доверительного интервала. По формулам (4.22) и (4.23) определим величину  $Z^*$

$$Z^* = \frac{1}{2} \ln \frac{1+0,94}{1-0,94} = 1,738$$

и ее среднеквадратичное отклонение

$$S_{Z^*} = \frac{1}{\sqrt{17-3}} = 0,267.$$

Зададимся вероятностью того, что истинное значение  $Z$  отличается от вычисленного на основании оценки коэффициента корреляции  $Z^*$  не более, чем на  $\delta_Z$ . Учитывая нормальный закон распределения  $Z$ , имеем при вероятности

$$95\%: \delta_Z = 1,96 \cdot S_{Z^*} = 1,96 \cdot 0,267 = 0,523;$$

$$90\%: \delta_Z = 1,67 \cdot 0,267 = 0,446;$$

$$99,7\%: \delta_Z = 3,00 \cdot 0,267 = 0,801.$$

Таким образом, истинное значение  $Z$  с вероятностью, например 95%, лежит в пределах  $Z_1 \leq Z \leq Z_2$ , где  $Z_1 = 1,738 - 0,523 = 1,215$  и  $Z_2 = 1,738 + 0,523 = 2,261$ .

Этим двум значениям  $Z$  соответствуют коэффициенты корреляции, полученные из формулы (4.22)  $-0,84 \leq r_{xy} \leq -0,98$ . Следовательно, доверительные интервалы подтверждают достаточно сильную причинную связь между анализируемыми параметрами.

Таким образом, корреляционный анализ устанавливает связь между исследуемыми случайными переменными и оценивает тесноту этой связи.

#### 4.5. Регрессионный анализ

Ниже излагаются основные положения регрессионного анализа, применение которого для обработки результатов наблюдений связано с меньшим числом ограничений, чем корреляционного анализа. Как и корреляционный анализ регрессионный анализ включает в себя построение уравнения регрессии, например методом наименьших квадратов, и статистическую оценку результатов. Если в регрессионном анализе расчет коэффициентов ведется теми же методами, например наименьших квадратов, то теоретические предпосылки регрессионного анализа требуют других способов статистической оценки результатов.

При проведении регрессионного анализа примем следующие допущения:

1. Входной параметр  $x$  измеряется с пренебрежимо малой ошибкой. Появление ошибки в определении  $y$  объясняется наличием в процессе невыявленных переменных и случайных воздействий, не вошедших в уравнение регрессии.

2. Результаты наблюдений  $y_1, y_2, \dots, y_i, \dots, y_n$  над выходной величиной представляют собой независимые нормально распределенные случайные величины.

3. При проведении эксперимента с объемом выборки  $n$  при условии, что каждый опыт повторен  $m^*$  раз, выборочные дисперсии  $S_1^2, \dots, S_i^2, \dots, S_n^2$  должны быть однородны. При выполнении измерений в различных условиях возникает задача сравнения точности измерений. При этом следует подчеркнуть, что экспериментальные данные можно сравнить только тогда, когда их дисперсии однородны. Это означает, как уже отмечалось (см. п. 3.5.2 и п. 3.5.3), принадлежность

экспериментальных данных к одной и той же генеральной совокупности. Напомним, что однородность дисперсий свидетельствует о том, что среди сравниваемых дисперсий нет таких, которые с заданной надежностью превышали бы все остальные, т.е. была бы большая ошибка. При одинаковом числе параллельных опытов однородность дисперсии, как мы уже показали, можно оценить по критерию Кохрена, а для сравнения двух дисперсий целесообразно воспользоваться F-критерием Фишера (см. примеры 3.7–3.9).

После того, как уравнение регрессии найдено, необходимо провести статистический анализ результатов. Этот анализ состоит в следующем: проверяется значимость всех коэффициентов и устанавливается адекватность уравнения.

#### 4.5.1. Проверка адекватности модели

При моделировании приходится формализовать связи исследуемого явления (процесса), из-за чего возможна потеря некоторой информации об объекте. Иногда некоторые связи не учитываются. В то же время основное требование к математической модели заключается в ее пригодности для решения поставленной задачи и адекватности процессу.

Так, построив модель в виде линейного уравнения регрессии, мы хотим в частности убедиться, что никакие другие модели не дадут значительного улучшения в описании предсказания значений  $Y$ .

Сформулируем нуль-гипотезу  $H_0$ : "Уравнение регрессии адекватно". Альтернативная гипотеза  $H_1$ : "Уравнение регрессии неадекватно". Для проверки этих гипотез принято использовать F-критерий Фишера.

При этом общую дисперсию (дисперсию выходного параметра)  $S_y^2$  сравнивают с остаточной дисперсией (дисперсией адекватности)  $S_{\text{ост}}^2$ .

Напомним, что

$$S_y^2 = \frac{\sum_{i=1}^n [y_i - \bar{y}]^2}{n-1}; \quad (4.24)$$

$$S_{\text{ост}}^2 = \frac{\sum_{i=1}^n [y_i - \hat{y}_i]^2}{n-l},$$

где  $l=k+1$  – число членов аппроксимирующего полинома, а  $k$  – число факторов. Так, например для линейной зависимости (4.5)  $k=1, l=2$ .

В дальнейшем определяется экспериментальное значение F-критерия Фишера

$$F_{\text{эксп}} = S_y^2 / S_{\text{ост}}^2, \quad (4.25)$$

который в данном случае показывает во сколько раз уравнение регрессии предсказывает результаты опытов лучше, чем среднее

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i = C = \text{const.}$$

Если  $F_{\text{эксп}} > F_{\alpha; m_1; m_2}$ , то уравнение регрессии адекватно. Чем больше значение  $F_{\text{эксп}}$  превышает  $F_{\alpha; m_1; m_2}$  для выбранного  $\alpha$  и числа степеней свободы  $m_1=n-1, m_2=n-l$ , тем эффективнее уравнение регрессии.

Рассмотрим так же случай, когда в каждой  $i$ -й точке  $x_i$  для повышения надежности и достоверности осуществляется не одно, а  $m^*$  параллельных измерений (примем для простоты, что  $m^*$  одинаково для каждого фактора). Тогда число экспериментальных значений величины  $Y$  составит  $n_{\Sigma} = n \cdot m^*$ .

В этом случае для оценки адекватности модели:

1. Определяется  $\bar{y}_i = \frac{\sum_{j=1}^{m^*} y_{ij}}{m^*}$  – среднее из серии параллельных опытов при  $x=x_i$ , где  $y_{ij}$  –

значение параметра  $Y$  при  $x=x_i$  в  $j$ -м случае.

2. Рассчитываются значения параметра  $\hat{y}_i$  по уравнению регрессии при  $x=x_i$ .

3. Рассчитывается дисперсия адекватности (аналог остаточной дисперсии)

$$S_{\text{ост}}^2 = \frac{m^* \sum_{i=1}^n [\bar{y}_i - \hat{y}_i]^2}{n-1}.$$

где  $n$  – число значений  $x_i$ ;  $l$  – число членов аппроксимирующего полинома (коэффициентов  $b_i$ ), для линейной зависимости  $l=2$ .

4. Определяется выборочная дисперсия  $Y$  при  $x=x_i$ :

$$S_i^2 = \frac{\sum_{j=1}^{m^*} [y_{ij} - \bar{y}_i]^2}{m^* - 1}.$$

5. Определяется дисперсия воспроизводимости (аналог общей дисперсии)

$$S_{\text{восп}}^2 = \sum_{i=1}^n S_i^2 / n.$$

Число степеней свободы этой дисперсии равно  $m_1 = n(m^* - 1)$ .

6. Определяется экспериментальное значение критерия Фишера

$$F_{\text{эксп}} = S_{\text{ад}}^2 / S_{\text{восп}}^2.$$

7. Теоретическое значение этого же критерия  $F_{\alpha; m_1; m_2}$ , где  $m_1 = m^*(n-1)$ ,  $m_2 = n-1$ .

8. Если  $F_{\text{эксп}} \leq F_{\alpha; m_1; m_2}$ , то уравнение регрессии адекватно, в противном случае – нет.

#### 4.5.2. Проверка значимости коэффициентов уравнения регрессии

Надежность оценок  $b_i$  уравнения регрессии можно охарактеризовать их доверительными интервалами  $\Delta b_i$ , в которых с заданной вероятностью находится истинное значение параметра.

Наиболее просто построить доверительные интервалы для параметров линейного уравнения регрессии, т.е. коэффициентов  $b_0$  и  $b_1$ . При этом предполагается, что для каждого значения случайной величины  $x=x_i$  имеется распределение со средним значением  $\hat{y}_i = b_0 + b_1 x_i$  и дисперсией

$S_{y_i}^2 = S_{\text{восп}}^2$ . Иными словами, делается допущение, что случайная величина  $Y$  распределена

нормально при каждом значении  $x_i$ , а дисперсия  $S_{y_i}^2$  во всем интервале изменения  $x$  постоянна

$S_{y_i}^2 = \text{const}$  (см.рис.4.9).

Для линейного уравнения среднеквадратичное отклонение  $i$ -го коэффициента уравнения регрессии  $S_{b_i}$  можно определить по закону накопления ошибок

$$S_{b_i} = \sqrt{\sum_{j=1}^n \left( \frac{\partial b_j}{\partial y_i} \right)^2 S_j^2}. \quad (4.26)$$

При условии, что  $S_{y_1}^2 = S_{y_2}^2 = \dots = S_{y_i}^2 = \dots = S_{y_n}^2 = S_{\text{восп}}^2$ , получим

$$S_{b_0} = \sqrt{\frac{S_{\text{восп}}^2 \sum_{i=1}^n x_i^2}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i\right)^2}}; \quad (4.27a)$$

$$S_{b_1} = \sqrt{\frac{S_{\text{восп}}^2 \cdot n}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i\right)^2}}. \quad (4.27b)$$

$S_{b_0}$  и  $S_{b_1}$  называются соответственно стандартной ошибкой свободного члена и стандартной ошибкой коэффициента регрессии.

Проверка значимости коэффициентов выполняется по критерию Стьюдента. При этом проверяется нуль-гипотеза  $H_0: b_i=0$ , т.е.  $i$ -й коэффициент генеральной совокупности при заданном уровне значимости  $\alpha$  отличен от нуля.

Построим доверительный интервал для коэффициентов уравнения регрессии

$$\Delta b_i = t_{\alpha; n-1} \cdot S_{b_i}, \quad (4.28)$$

где число степеней свободы в критерии Стьюдента определяется по соотношению  $n-1$ . Потеря  $l=k+1$  степеней свободы обусловлена тем, что все коэффициенты  $b_i$  рассчитываются зависимо друг от друга, что следует из уравнений (4.16a) и (4.16б).

Тогда доверительный интервал для  $\Delta b_i$  коэффициента уравнения регрессии составит  $(b_i - \Delta b_i; b_i + \Delta b_i)$ . Чем уже доверительный интервал, тем с большей уверенностью можно говорить о значимости этого коэффициента.

**Необходимо всегда помнить рабочее правило: "Если абсолютная величина коэффициента регрессии больше, чем его доверительный интервал, то этот коэффициент значим".**

Таким образом, если  $|b_i| > |\Delta b_i|$ , то  $b_i$  коэффициент значим, в противном случае – нет.

**Незначимые коэффициенты исключаются из уравнения регрессии, а оставшиеся коэффициенты пересчитываются заново, т.к. они зависимы и в формулы для их расчета (4.16a) и (4.16б) входят разноименные переменные.**

#### 4.6. Линейная множественная регрессия

При изучении множественной регрессии не существует графической интерпретации многофакторного пространства. При проведении экспериментов в такой ситуации исследователь записывает показания приборов о состоянии функции отклика  $y$  и всех факторов, от которых она зависит  $x_j$ . Результат исследований – это матрица наблюдений.

$$\begin{pmatrix} y_1 & x_{11} & x_{12} & \dots & x_{1j} & \dots & x_{1k} \\ y_2 & x_{21} & x_{22} & \dots & x_{2j} & \dots & x_{2k} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ y_i & x_{i1} & x_{i2} & \dots & x_{ij} & \dots & x_{ik} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ y_n & x_{n1} & x_{n2} & \dots & x_{nj} & \dots & x_{nk} \end{pmatrix}. \quad (4.29)$$

Здесь  $n$  – число опытов;  $k$  – число факторов;  $x_{ij}$  – значение  $j$ -го фактора в  $i$ -м опыте;  $y_i$  – значение выходного параметра для  $i$ -го опыта.

Задача линейной множественной регрессии состоит в построении плоскости в (k+1)-мерном пространстве, отклонения результатов наблюдений  $y_i$  от которой были бы минимальными. Или другими словами, следует определить значения коэффициентов  $b_0, \dots, b_j, \dots, b_k$  в линейном полиноме

$$\hat{y} = b_0 + \sum_{j=1}^k b_j x_j,$$

что равносильно минимизации выражения

$$\Phi = \sum_i (y_i - \hat{y}_i)^2 = \sum_i [y_i - (b_0 + b_1 x_{i1} + \dots + b_j x_{ij} + \dots + b_k x_{ik})]^2 \rightarrow \min_{b_j}. \quad (4.30)$$

Процедура определения коэффициентов  $b_0, \dots, b_j, \dots, b_k$  в принципе не отличается от одномерного случая, рассмотренного ранее и поэтому здесь не приводится.

Для оценки тесноты связи между функцией отклика  $\hat{y}$  и несколькими факторами  $x_1, x_2, \dots, x_i, \dots, x_k$  используют коэффициент множественной корреляции  $R$ , который всегда положителен и изменяется в пределах от 0 до 1. Чем больше  $R$ , тем лучше качество предсказаний данной моделью опытных данных с точки зрения близости ее к функциональной. При функциональной линейной зависимости  $R=1$ .

Расчеты обычно начинают с вычисления парных коэффициентов корреляции, при этом вычисляются два типа парных коэффициентов корреляции:

1)  $r_{yx_j}$  – коэффициенты, определяющие тесноту связи между функцией отклика  $\hat{y}$  и одним из факторов  $x_j$ ;

2)  $r_{x_j x_u}$  – коэффициенты, показывающие тесноту связи между одним из факторов  $x_j$  и фактором  $x_u$  ( $j, u = 1 \div k$ ).

После вычисления всех парных коэффициентов корреляции можно построить матрицу коэффициентов корреляции следующего вида

$$\begin{pmatrix} 1 & r_{yx_1} & r_{yx_2} & \dots & r_{yx_j} & \dots & r_{yx_k} \\ r_{x_1 y} & 1 & r_{x_1 x_2} & \dots & r_{x_1 x_j} & \dots & r_{x_1 x_k} \\ r_{x_2 y} & r_{x_2 x_1} & 1 & \dots & r_{x_2 x_j} & \dots & r_{x_2 x_k} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ r_{x_j y} & r_{x_j x_1} & r_{x_j x_2} & \dots & 1 & \dots & r_{x_j x_k} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ r_{x_k y} & r_{x_k x_1} & r_{x_k x_2} & \dots & r_{x_k x_j} & \dots & 1 \end{pmatrix}. \quad (4.31)$$

Используя матрицу (4.31), можно вычислить частные коэффициенты корреляции, которые показывают степень влияния одного из факторов  $x_j$  на функцию отклика  $\hat{y}$  при условии, что остальные факторы остаются на постоянном уровне. Формула для вычисления частных коэффициентов корреляции имеет вид

$$r_{yx_1, x_2, \dots, x_j, \dots, x_k} = D_{1j} / \sqrt{D_{11} \cdot D_{jj}}, \quad (4.32)$$

где  $D_{1j}$  – определитель матрицы, образованной из матрицы (4.32) вычеркиванием 1-й строки j-го столбца. Определитель  $D_{11}$  и  $D_{jj}$  вычисляются аналогично. Как и парные коэффициенты, частные коэффициенты корреляции изменяются от -1 до +1.

Значимость и доверительный интервал для коэффициентов частной корреляции определяются так же, как для коэффициентов парной корреляции, только число степеней свободы вычисляются по формуле

$$m = n - k^* - 2 \quad (4.33)$$

где  $k^* = k - 1$  – порядок частного коэффициента парной корреляции.

Для вычисления коэффициента множественной корреляции  $R_{yx_1, x_2, \dots, x_j, \dots, x_k}$  используют матрицу

$$R_{yx_1, x_2, \dots, x_j, \dots, x_k} = \sqrt{1 - D/D_{11}}, \quad (4.34)$$

где  $D$  – определитель матрицы (4.31).

Если число опытов  $n$  сравнимо с числом коэффициентов  $l=k+1$ , связи оказываются преувеличенными. Поэтому следует исключить систематическую погрешность, физический смысл которой состоит в следующем. Если разность  $n$  и  $l$  будет уменьшаться, то коэффициент множественной корреляции  $R$  будет возрастать и при  $n-l=0$  окажется равным  $R=+1$ , а уравнение регрессии превратится в функциональное уравнение гиперплоскости, которая пройдет через все  $n$  экспериментальных точек. Однако ясно, что случайный характер переменных процесса при этом не может измениться. В связи с этим требуется оценка значимости коэффициента множественной корреляции.

Значимость коэффициента множественной корреляции проверяется по критерию Стьюдента:

$$t_{\text{эксп}} = \frac{R}{\bar{S}_R} \geq t_{\alpha; m}; \quad m = n - k - 1,$$

где  $\bar{S}_R$  – среднеквадратичная погрешность коэффициента множественной корреляции, рассчитываемая по выражению

$$\bar{S}_R = (1 - R^2) \sqrt{n - k - 1}. \quad (4.35)$$

Значимость  $R$  можно проверить также по критерию Фишера

$$F_{\text{эксп}} = \frac{R(n - k - 1)}{(1 - R^2)n}. \quad (4.36)$$

Если расчетное значение  $F_{\text{эксп}}$  превышает теоретическое  $F_{\text{теор}}$ , то гипотезу о равенстве коэффициента множественной корреляции нулю отвергают и связь считают статистически значимой. Теоретическое (табличное) значение критерия Фишера определяется для выбранного уровня значимости  $\alpha$  и числа степеней свободы  $m_1=n-k-1$  и  $m_2=k$ .

Если коэффициент множественной корреляции оказался неожиданно малым, хотя априорно известно, что между выходом  $y$  и входами  $x_1, \dots, x_k$  должна существовать достаточно тесная корреляционная связь, то возможными причинами такого явления могут быть:

а) ряд существенных факторов не учтен и следует включить в рассмотрение дополнительно эти существенные входные параметры;

б) линейное уравнение плохо аппроксимирует в действительности нелинейную зависимость  $y = f(x_1, \dots, x_n)$  и следует определить коэффициенты уже нелинейного уравнения регрессии методами регрессионного анализа;

в) рабочий диапазон рассматриваемых факторов находится в районе экстремума функции отклика – в этом случае следует расширить диапазон изменения входных переменных и так же перейти к нелинейной математической модели объекта.

#### 4.7. Нелинейная регрессия

Используя подходы, изложенные ранее, можно построить практически любые формы нелинейной связи. С этой целью в инженерной практике очень часто используют линеаризующие преобразования.

Таблица 4.1

Функции и линеаризующие преобразования

№ п/п	Функция	Линеаризующие преобразования			
		Преобразование переменных		Выражения для величин $b_0$ и $b_1$	
		$y'$	$x'$	$b_0'$	$b_1'$

1	$y = b_0 + b_1/x$	$y$	$1/x$	$b_0$	$b_1$
2	$y = 1/(b_0 + b_1x)$	$1/y$	$x$	$b_0$	$b_1$
3	$y = x/(b_0 + b_1x)$	$x/y$	$x$	$b_0$	$b_1$
4	$y = b_0 b_1^x$	$\lg(y)$	$x$	$\lg(b_0)$	$\lg(b_1)$
5	$y = b_0 \cdot e^{b_1 x}$	$\ln(y)$	$x$	$\ln(b_0)$	$b_1$
6	$y = 1/(b_0 + b_1 e^{-x})$	$1/y$	$e^{-x}$	$b_0$	$b_1$
7	$y = b_0 x^{b_1}$	$\lg(y)$	$\lg(x)$	$\lg(b_0)$	$b_1$
8	$y = b_0 + b_1 \lg(x)$	$y$	$\lg(x)$	$b_0$	$b_1$
9	$y = b_0/(b_1 + x)$	$1/y$	$x$	$b_1/b_0$	$1/b_0$
10	$y = b_0 x/(b_1 + x)$	$1/y$	$1/x$	$b_1/b_0$	$1/b_0$
11	$y = b_0 e^{b_1/x}$	$\ln(y)$	$1/x$	$\ln(b_0)$	$b_1$
12	$y = b_0 + b_1 x^n$	$y$	$x^n$	$b_0$	$b_1$

В табл. 4.1. приведены часто встречающиеся парные зависимости и линеаризующие преобразования переменных. Качество преобразования результатов проверяют с помощью уравнения  $\hat{y} = b_0' + b_1' \cdot x'$ . После вычисления коэффициентов  $b_0'$  и  $b_1'$ , в частности по методу наименьших квадратов, как для линейной зависимости от одного фактора (см. п. 4.2) выполняют обратные преобразования, т.е. по  $b_0'$  и  $b_1'$  определяют  $b_0$  и  $b_1$ . Аналогичный подход обычно широко используют и при множественном нелинейном регрессионном анализе.

Вопросы к теме 4

1. Характеристика видов связей между рядами наблюдений
2. Определение коэффициентов уравнения регрессии
3. Определение тесноты связи между случайными величинами
4. Линейная регрессия от одного фактора
5. Регрессионный анализ
6. Проверка адекватности модели
7. Проверка значимости коэффициентов уравнения регрессии
8. Линейная множественная регрессия
9. Нелинейная регрессия

## Тема 5. ОЦЕНКА ПОГРЕШНОСТЕЙ РЕЗУЛЬТАТОВ НАБЛЮДЕНИЙ

### 5.1. Оценка погрешностей определения величин функций

Необходимость в определении погрешности величин-функций по известным значениям погрешностей их аргументов (факторов) возникает при оценке точности результатов математического эксперимента, а также результатов так называемых косвенных измерений. Под косвенным измерением понимают такое, в результате которого значение искомой величины у рассчитывают по известной зависимости ее от других величин  $x_1, x_2, \dots, x_k$ , измеренных другим способом, т.е.

$$\hat{y} = f(x_1, x_2, \dots, x_i, \dots, x_k), \quad (5.1)$$

где  $x_1, x_2, \dots, x_i, \dots, x_k$  – аргументы, определенные независимо друг от друга. В дальнейшем будем полагать, что погрешности определения величины  $y$  обусловлены лишь неточностью численных значений величин  $x_1, x_2, \dots, x_i, \dots, x_k$ , входящих под знак функции.

Обозначим истинное значение  $i$ -го параметра через  $x_i$ , среднее значение – через  $\bar{x}_i$ , а абсолютную погрешность его измерения – через  $\Delta x_i$ . Разложим функцию  $f(x_1, x_2, \dots, x_k)$  в ряд Тейлора, сохраняя члены с нулевым и первыми степенями погрешностей

$$f(x_1, \dots, x_i, \dots, x_k) = f(\bar{x}_1, \dots, \bar{x}_i, \dots, \bar{x}_k) + \sum_{i=1}^k \left( \frac{\partial f(x_1, \dots, x_i, \dots, x_k)}{\partial x_i} \right)_{x_i=\bar{x}_i} \cdot \Delta x_i,$$

где все производные  $\left( \frac{\partial f(x_1, \dots, x_i, \dots, x_k)}{\partial x_i} \right)_{x_i=\bar{x}_i}$  вычислены при значениях  $x_i = \bar{x}_i$ .

Тогда

$$\begin{aligned} \Delta y^2 &\approx \left[ f(x_1, \dots, x_i, \dots, x_k) - f(\bar{x}_1, \dots, \bar{x}_i, \dots, \bar{x}_k) \right]^2 = \left[ \sum_{i=1}^k \left( \frac{\partial f(x_1, \dots, x_i, \dots, x_k)}{\partial x_i} \right)_{x_i=\bar{x}_i} \Delta x_i \right]^2 \approx \\ &\approx \sum_{i=1}^k \left( \frac{\partial f(x_1, \dots, x_i, \dots, x_k)}{\partial x_i} \right)_{x_i=\bar{x}_i}^2 \Delta x_i^2 = \sum_{i=1}^k \Delta y_i^2, \end{aligned}$$

где  $\Delta y_i = \left( \frac{\partial f(x_1, \dots, x_k)}{\partial x_i} \right)_{x_i=\bar{x}_i} \cdot \Delta x_i.$  (5.2)

Следовательно  $\Delta y_i$  – это составляющие погрешности функции, обусловленные погрешностью  $i$ -го аргумента  $x_i$ .

Доверительная вероятность, соответствующая величине  $\Delta y_i$ , численно равна доверительной вероятности, с которой найдена погрешность  $\Delta x_i$ .

Для относительной погрешности вместо соотношения (5.2) используют выражение:

$$\Delta y_i^* = \frac{1}{f} \frac{\partial f}{\partial x_i} \Delta x_i = \frac{\partial \ln(f)}{\partial x_i} \Delta x_i. \quad (5.3)$$

Соотношения (5.2) и (5.3) применимы для расчета как случайных, так и систематических погрешностей.

Общая абсолютная ( $\Delta y_\Sigma$ ) и относительная ( $\Delta \Sigma^*$ ) погрешности определения функции могут быть найдены с помощью выражений

$$\Delta y_\Sigma = \pm \sqrt{\sum_{i=1}^k (\Delta y_i)^2}; \quad (5.4)$$

$$\Delta \Sigma^* = \pm \sqrt{\sum_{i=1}^k (\Delta y_i^*)^2}. \quad (5.5)$$

Предполагается, что все составляющие имеют нормальный закон распределения.

Частные производные, входящие в соотношения (5.2) и (5.3), не всегда могут быть взяты аналитически. Часто не удается разрешить искомую задачу в явном виде относительно искомой величины  $y$ . В этих случаях полезно использовать численные методы определения производных.

В качестве примера (пример 5.1) рассмотрим погрешность определения массового расхода газового потока стандартным сужающим устройством. При этом будем считать, что случайная составляющая погрешности отсутствует, а поправка на сжимаемость потока равна единице.

Тогда с учетом выражения для определения массового расхода вещества

$$G = \alpha F_0 \varepsilon \sqrt{2\rho(p_1 - p_2)} = \alpha F_0 \varepsilon \sqrt{2\rho h}. \quad (5.6)$$

где  $F_0$  – площадь сужающего устройства;  $\varepsilon$  – поправочный множитель на сжимаемость вещества, расход которого измеряется;  $\rho$  – плотность потока перед сужающим устройством;  $h$  – перепад статического давления на сужающем устройстве.

Используя соотношения (5.2) и (5.4), получим следующую формулу для расчета абсолютной и относительной погрешности определения расхода:

$$\Delta G = \pm \sqrt{\left(\frac{\partial G}{\partial \rho} \Delta \rho\right)^2 + \left(\frac{\partial G}{\partial h} \Delta h\right)^2}; \quad (5.7)$$

$$\Delta G^* = \pm \sqrt{\left(\frac{\partial G}{\partial \rho} \frac{\Delta \rho}{G}\right)^2 + \left(\frac{\partial G}{\partial h} \frac{\Delta h}{G}\right)^2} = \pm \frac{1}{2} \sqrt{\left(\frac{\Delta \rho}{\rho}\right)^2 + \left(\frac{\Delta h}{h}\right)^2}, \quad (5.8)$$

где  $\frac{\partial G}{\partial \rho} = \alpha F_0 \varepsilon \sqrt{\frac{h}{2\rho}}$ ;  $\frac{\partial G}{\partial h} = \alpha F_0 \varepsilon \sqrt{\frac{\rho}{2h}}$ .

Учтем далее погрешности определения плотности потоков. В соответствии с уравнением состояния газа  $\rho = p/RT$ , где  $p$  и  $T$  – соответственно абсолютное давление и температура газа перед сужающим устройством. Абсолютная погрешность определения плотности потока без учета погрешности газовой постоянной составит:

$$\Delta \rho = \pm \sqrt{\left(\frac{\partial \rho}{\partial p} \Delta p\right)^2 + \left(\frac{\partial \rho}{\partial T} \Delta T\right)^2}, \quad (5.9)$$

где  $\frac{\partial \rho}{\partial p} = \frac{1}{RT}$ ;  $\frac{\partial \rho}{\partial T} = -\frac{p}{RT^2}$ ;

относительная погрешность

$$\Delta \rho^* = \frac{\Delta \rho}{\rho} = \pm \sqrt{\left(\frac{\Delta p}{p}\right)^2 + \left(\frac{\Delta T}{T}\right)^2}. \quad (5.10)$$

Тогда относительная погрешность определения массового расхода газового потока будет равна:

$$\Delta G^* = \pm \frac{1}{2} \sqrt{\left(\frac{\Delta p}{p}\right)^2 + \left(\frac{\Delta T}{T}\right)^2 + \left(\frac{\Delta h}{h}\right)^2}. \quad (5.11)$$

Здесь  $p$ ,  $T$ ,  $h$  – значения измеренных параметров;  $\Delta p$ ,  $\Delta T$ ,  $\Delta h$  – их абсолютные погрешности. Численные значения  $\Delta p$ ,  $\Delta T$ ,  $\Delta h$  определяются в основном инструментальной погрешностью и могут быть вычислены с учетом класса точности используемых приборов для измерения  $p$  и  $h$ . Погрешность измерения  $T$  определяется с учетом вида измерительного устройства температуры.

Абсолютная погрешность определения массового расхода газового потока

$$\Delta G = G \Delta G^*, \quad (5.12)$$

где  $G$  – значение расхода, измеренное экспериментально.

$$G_{\text{ИСТ}} = G \pm \Delta G. \quad (5.13)$$

## 5.2. Обратная задача теории экспериментальных погрешностей

Целью обратной задачи является определение погрешностей величин-аргументов, если известны погрешности функций и вид функциональной зависимости. Необходимость в решении таких задач возникает при выборе того или иного комплекса измерительной аппаратуры или метода определения искомой величины, позволяющих найти значение этой величины с определенной погрешностью.

Обратная задача в общем случае является неопределенной, поскольку имеется одно уравнение с  $k$  неизвестными. Иначе говоря, удовлетворить условию задачи можно при различных комбинациях значений погрешностей аргументов.

**Очень часто удовлетворительное решение обратной задачи оказывается возможным при использовании так называемого принципа равных влияний. Он заключается в том, что при решении задачи накладывается дополнительное требование, чтобы все члены в правой части выражений (5.4) и (5.5) оказывали одинаковое влияние на погрешности функции.**

Применяя принцип равных влияний к относительной погрешности функции, определяемой соотношением (5.5), получим

$$\left(\Delta_{y_1}^*\right)^2 = \left(\Delta_{y_2}^*\right)^2 = \dots = \left(\Delta_{y_k}^*\right)^2 = \frac{\left(\Delta_{\Sigma}^*\right)^2}{k}, \quad (5.14)$$

$$\Delta_{y_i}^* = \frac{\Delta_{\Sigma}^*}{\sqrt{k}}. \quad (5.15)$$

С учетом (5.3) легко получить выражение для определения абсолютных  $\Delta x_i$  и относительных  $\Delta_{x_i}^*$  погрешностей всех аргументов:

$$\Delta x_i = \frac{\Delta_{\Sigma}}{\sqrt{k}} \frac{1}{\frac{\partial \ln [f(x_1, x_2, \dots, x_k)]}{\partial x_i}}; \quad (5.16)$$

$$\Delta_{x_i}^* = \frac{\Delta_{\Sigma}}{x_i \sqrt{k}} \frac{1}{\frac{\partial \ln [f(x_1, x_2, \dots, x_k)]}{\partial x_i}}. \quad (5.17)$$

В дальнейшем могут иметь место три возможных случая:

- значения погрешностей всех аргументов таковы, что лежат в пределах точности, доступной при измерениях с помощью имеющихся средств измерений;
- значения некоторых погрешностей настолько малы, что обеспечить соответствующую точность с помощью имеющихся средств измерений не представляется возможным;
- значения всех погрешностей малы и обеспечить такую точность невозможно.

В первом случае проблем не возникает и поставленная задача имеет решение. Во втором случае прежде всего следует попытаться решить задачу путем увеличения погрешности тех аргументов, у которых оказалось невозможным обеспечить требуемую первоначальную точность измерений при одновременном уменьшении погрешностей остальных аргументов.

Если этот путь не дает приемлемых результатов, то остается один выход, связанный с поиском другого метода определения величины  $x$ . Этот выход является единственно возможным и для случая, когда значения погрешностей всех аргументов настолько малы, что обеспечить требуемую их точность с помощью имеющихся средств измерений не представляется возможным. При выборе другого метода измерений меняется вид функции  $y=f(\mathbf{X})$ , а следовательно, меняются аргументы и значения их погрешностей.

Рассмотрим пример 5.2.

Пусть требуется определить объем цилиндра диаметром  $d=20$  мм и высотой  $h=50$  мм с относительной погрешностью  $\Delta_V^*=0,01$ , соответствующей доверительной вероятностью  $P=0,95$ . Найдем погрешности измерения величины  $d$  и  $h$ , соответствующие этому же значению доверительной вероятности, при которых исходная задача будет разрешена. Учитывая, что объем

цилиндра равен  $V = \frac{\pi d^2 h}{4}$  и приняв закон распределения нормальным, с помощью соотношения

(5.16) найдем

$$\Delta d = \pm \frac{\Delta_V^*}{\sqrt{2}} \frac{1}{\frac{\partial \ln(\pi d^2 h / 4)}{\partial d}} = \pm \frac{\Delta_V^*}{\sqrt{2}} \frac{d}{2} = \pm 0,07 \text{ мм};$$

$$\Delta h = \pm \frac{\Delta_V^*}{\sqrt{2}} \frac{1}{\frac{\partial \ln(\pi d^2 h / 4)}{\partial h}} = \pm \frac{\Delta_V^*}{\sqrt{2}} h = \pm 0,035 \text{ мм}.$$



Пассивный эксперимент предусматривает накопление информации “в режиме нормальной эксплуатации”, но это требует много времени и затрат. Поэтому предлагается “не ждать милостей от природы”, а активно вмешиваться в ход технологического процесса: разбалтывать (покачивать) его тихонько, но целенаправленно, и быстро накапливать при этом информацию. Программа покачивания как раз и задается планом. Сам метод планирования может изменяться в зависимости от вида задачи, но принцип покачивания остается.

Теория планирования эксперимента началась с работ знаменитого английского ученого Р.Фишера в 30-х годах XX столетия, которую он использовал для решения агробиологических задач. В дальнейшем это направление было развито в пятидесятых годах в США Дж.Боксом и его сотрудниками. Отечественные ученые внесли большой вклад в развитие теории эксперимента, предложив ряд новых методов, а инженеры-исследователи все шире применяют эти методы на практике.

Под математической теорией планирования эксперимента будем понимать науку о способах составления экономичных экспериментальных планов, которые одновременно позволяют извлекать наибольшее количество информации об объекте исследования, о способах проведения эксперимента, о способах обработки экспериментальных данных и их использования для оптимизации производственных процессов, а также инженерных расчетов.

Принятая терминология — это либо перевод терминов с английского, либо просто их перенос в оригинале, это необходимо иметь в виду при чтении литературы по теории планирования экспериментов.

Истинный вид функции отклика  $y=f(x_1, \dots, x_i, \dots, x_k)$  до эксперимента чаще всего неизвестен, в связи с чем для математического описания поверхности отклика используют уравнение

$$y = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i,j=1}^k \beta_{ij} x_i x_j + \sum_{i=1}^k \beta_{ii} x_i^2 + \dots \quad (6.1)$$

где  $x_i, x_j$  — переменные факторы при  $i=1, \dots, n; j=1, \dots, n; i \neq j$ ;

$$\beta_i = \left( \frac{\partial f}{\partial x_i} \right)_0; \beta_{ij} = \left( \frac{\partial^2 f}{\partial x_i \partial x_j} \right)_0; \beta_{ii} = \left( \frac{\partial^2 f}{2 \partial x_i^2} \right)_0 \quad \text{— коэффициенты.}$$

Это уравнение является разложением в ряд Тейлора неизвестных функций отклика в окрестности точки с  $x_i=x_{i0}$ .

На практике по результатам эксперимента производится обработка данных по методу наименьших квадратов. Этот метод позволяет найти оценку  $b$  коэффициентов  $\beta$ , и данный полином заменяется уравнением вида:

$$\hat{y} = b_0 + \sum_{i=1}^k b_i x_i + \sum_{i,j=1}^k b_{ij} x_i x_j + \sum_{i=1}^k b_{ii} x_i^2 + \dots, \quad (6.2)$$

которое является регрессионной моделью (моделью регрессионного анализа). В этом выражении  $\hat{y}$  означает модельное, т.е. рассчитываемое по уравнению модели, значение выхода. Коэффициенты регрессии определяются экспериментально и служат для статистической оценки теоретических коэффициентов, т.е.

$$b_0 \rightarrow \beta_0, b_i \rightarrow \beta_i, b_{ij} \rightarrow \beta_{ij}, b_{ii} \rightarrow \beta_{ii}.$$

В регрессионной модели члены второй степени  $x_i x_j, x_i^2$  характеризуют кривизну поверхности отклика. Чем больше кривизна этой поверхности, тем больше в модели регрессии членов высшей степени. На практике чаще всего стремятся ограничиться линейной моделью.

Последовательность активного эксперимента заключается в следующем.

1. Разрабатывается схема проведения исследований, т.е. выполняется планирование эксперимента. При планировании экспериментов обычно требуется с наименьшими затратами и с необходимой точностью либо:

- а) построить регрессионную модель процесса;
- б) определить его оптимальные условия.

2. Осуществляется реализация опыта по заранее составленному исследователем плану, т.е. осуществляется сам активный эксперимент.

3. Выполняется обработка результатов измерений, их анализ и принятие решений.

Таким образом, планирование эксперимента — это процедура выбора условий проведения опытов, их количества, необходимых и достаточных для решения задач с поставленной точностью.

Использование теории планирования эксперимента обеспечивает:

1. Минимизацию, т.е. предельное сокращение необходимого числа опытов.

2. Одновременное варьирование всех факторов.

3. Выбор четкой стратегии, что позволяет принимать обоснованные решения после каждой серии опытов.

4. Минимизацию ошибок эксперимента за счет использования специальных проверок.

Для иллюстрации некоторых из этих положений воспользуемся ставшим уже классическим примером из книги В.В.Налимова, Т.И.Голикова. Логические основания планирования эксперимента. М.: Металлургия, 1980. 152 с.

### 6.2. Пример хорошего и плохого эксперимента

Рассмотрим пример — взвешивание трех объектов А, В, С на аналитических весах. Первый, традиционный подход, предусматривает последовательное взвешивание каждого из образцов. Так поступает традиционно исследователь: вначале он делает холостое взвешивание для определения нулевой точки весов, а затем по очереди взвешивают каждый из образцов. Это пример традиционного использования однофакторного эксперимента, т.е. здесь исследователь изучает реакцию на поведение каждого из факторов в отдельности. Традиционная схема взвешивания трех объектов представлена в табл.6.1.

Таблица 6.1

Традиционное проведение эксперимента\*

Номер опыта	А	В	С	Результат взвешивания
1	-1	-1	-1	$y_0$
2	+1	-1	-1	$y_1$
3	-1	+1	-1	$y_2$
4	+1	+1	+1	$y_3$

\* Когда образец кладется на весы, в таблице ставится +1, когда он на весах отсутствует, то -1.

Масса каждого объекта оценивается только по результатам двух опытов: того опыта, в котором на весы был положен изучаемый объект, и холостым опытом. Например, масса объекта А равна  $m_A = y_1 - y_0$ . Как обычно, ошибка взвешивания предполагается независимой от взвешиваемой величины, аддитивной и имеющей одно и тоже распределение. Тогда дисперсия измерения веса образца равна

$$\sigma_A^2 = \sigma_{y_1}^2 + \sigma_{y_0}^2 = 2\sigma^2, \quad (6.3)$$

где  $\sigma^2$  — дисперсия любого взвешивания. Такими же будут и дисперсии весов образцов В и С.

Приведем теперь тот же эксперимент по несколько иной схеме, задаваемой матрицей планирования, приведенной в табл.6.2.

Таблица 6.2

Планирование эксперимента при взвешивании трех объектов

Номер опыта	А	В	С	Результат взвешивания
1	+1	-1	-1	$y_1$
2	-1	+1	-1	$y_2$
3	-1	-1	+1	$y_3$
4	+1	+1	+1	$y_4$

В первых трех опытах последовательно взвешивают объекты А, В, С, в последнем опыте взвешивают объекты А, В, С, т.е. все три объекта вместе, а "холостое" взвешивание не производится.

Легко заметить, что масса каждого объекта будет задаваться формулами

$$\begin{aligned} m_A &= \frac{1}{2}(y_1 - y_2 - y_3 + y_4); \\ m_B &= \frac{1}{2}(y_2 - y_1 - y_3 + y_4); \\ m_C &= \frac{1}{2}(y_3 - y_1 - y_2 + y_4). \end{aligned} \quad (6.4)$$

Масса объекта А, вычисленная по приведенной выше формуле, оказывается не искаженной массами весов объектов В и С, т.к. масса каждого из них входит в формулу для массы А дважды с разными знаками.

Найдем теперь дисперсию, связанную с ошибкой взвешивания по новой схеме постановки экспериментов

$$\sigma_A^2 = \frac{1}{4}(\sigma_{y_1}^2 + \sigma_{y_2}^2 + \sigma_{y_3}^2 + \sigma_{y_4}^2) = \sigma^2. \quad (6.5)$$

Аналогичным образом находим:

$$\sigma_B^2 = \sigma^2, \quad \sigma_C^2 = \sigma^2.$$

Мы видим, что при новой схеме дисперсия взвешивания получается вдвое меньше, чем при традиционной схеме, хотя в обоих случаях на взвешивание трех объектов затрачивалось по четыре опыта.

Зададимся вопросом: "В результате чего происходит увеличение точности экспериментов в два раза?".

В первом случае эксперимент был поставлен так, что каждую массу мы получали лишь из двух взвешиваний. При новой схеме взвешивания каждая масса вычислялась уже по результатам всех четырех взвешиваний. **Вторую схему можно назвать многофакторной, т.к. здесь оперируют всеми факторами так, что каждая масса вычислялась по результатам сразу всех опытов, проведенных в данной серии экспериментов — вот главная причина уменьшения дисперсии вдвое.**

Не подумайте, что мы зря потратили время на обсуждение такой тривиальной задачи. Точно такой же подход используется при изучении других, более сложных задач.

Таким образом, использование теории планирования эксперимента может явиться одним из путей существенного повышения эффективности многофакторных экспериментальных исследований.

В планировании экспериментов используются, в основном, планы первого и второго порядков. Планы более высоких порядков используются в инженерной практике редко. В связи с этим далее приводится краткое изложение методики составления планов эксперимента для моделей первого и второго порядков.

**Под планами первого порядка понимают такие планы, которые позволяют провести эксперимент для отыскания уравнения регрессии, содержащего только первые степени факторов и их произведения**

$$y = b_0 + \sum_{i=1}^k b_i x_i + \sum_{\substack{i,j=1 \\ i \neq j}}^k b_{ij} x_i x_j + \sum_{\substack{i,j,u=1 \\ i \neq j \neq u}}^k b_{iju} x_i x_j x_u + \dots \quad (6.6)$$

Планы второго порядка позволяют провести эксперимент для отыскания уравнения регрессии, содержащего и вторые степени факторов.

$$\hat{y} = b_0 + \sum_{i=1}^k b_i x_i + \sum_{i=1}^k b_{ii} x_i^2 + \sum_{\substack{i,j=1 \\ i \neq j}}^k b_{ij} x_i x_j + \dots \quad (6.7)$$

Нахождение уравнения регрессии методом планирования экспериментов состоит из следующих этапов:

- выбор основных факторов и их уровней;
- планирование и проведение собственно эксперимента;
- определение коэффициентов уравнения регрессии;
- статистический анализ результатов эксперимента.

### 6.3. Планирование первого порядка

На первой стадии исследования обычно принимают полином первой степени. Так, для трехфакторной задачи теоретическое уравнение регрессии имеет вид

$$y = \beta_0 + \sum_{i=1}^3 \beta_i x_i + \sum_{\substack{i,j=1 \\ i \neq j}}^3 \beta_{ij} x_i x_j + \beta_{123} x_1 x_2 x_3. \quad (6.8)$$

Уравнение регрессии, получаемое на основании результатов эксперимента, в отличие от приведенного теоретического уравнения, имеет вид

$$\hat{y} = b_0 + \sum_{i=1}^3 b_i x_i + \sum_{\substack{i,j=1 \\ i \neq j}}^3 b_{ij} x_i x_j + b_{123} x_1 x_2 x_3, \quad (6.9)$$

где коэффициенты регрессии  $b_0, b_1, \dots, b_3, \dots, b_{123}$  являются оценками для теоретических коэффициентов регрессии, т.е.

$$b_i \rightarrow \beta_i, \quad b_{ij} \rightarrow \beta_{ij}, \quad b_{123} \rightarrow \beta_{123}.$$

Члены, содержащие произведения  $x_1 x_2, x_2 x_3$  и т.д., называют членами, отражающими попарное взаимодействие факторов, члены вида  $x_1 x_2 x_3$  — членами тройного взаимодействия.

#### 6.3.1. Выбор основных факторов и их уровней

В качестве факторов можно выбирать только контролируемые и управляемые переменные, т.е. такие, которые исследователь может поддерживать постоянными в течение каждого опыта на заданном уровне. В число факторов должны быть включены параметры процесса, оказывающие наиболее сильное влияние на функцию отклика. Необходимо заметить, что несмотря на всю заманчивость и очевидные преимущества активного спланированного эксперимента перед пассивным, в его применении имеется целый ряд трудностей, связанных с определенными ограничениями на его реализацию. Важнейшим условием применимости этого подхода является управляемость процессов по каждому из выбранных факторов, т.е. возможность независимого изменения каждого из этих факторов и поддержания его на заданном уровне в период проведения опытов.

Для каждого фактора необходимо указать тот интервал изменения параметров, в пределах которого ставится исследование. Для этого на основе априорной информации устанавливаются ориентировочные значения факторов  $x_{10}, x_{20}, \dots, x_{i0}, \dots, x_{k0}$ . Этой комбинации значений факторов соответствует точка в многомерном факторном пространстве, которая принимается за исходную точку. Координаты этой точки принимаются за основной (нулевой) уровень.

Интервалом варьирования факторов называется некоторое число (каждое для соответствующего фактора), прибавление которого к основному уровню дает верхний, а вычитание — нижний пределы. Для упрощения записи условий эксперимента и обработки экспериментальных

данных масштабы по осям выбираются так, чтобы верхний уровень составлял +1, нижний -1, а основной — 0.

Для факторов с непрерывной областью определения это достигается с помощью преобразования (кодирования факторов)

$$X_i = \frac{x_i - x_{i0}}{\Delta x_i}. \quad (6.10)$$

**В теории планирования экспериментов показано, что минимально необходимое число уровней факторов на единицу больше порядка уравнения.**

### 6.3.2. Планирование эксперимента

Рассмотрим сначала частный случай, когда функция отклика линейно зависит от трех независимых факторов. Уравнение регрессии в этом случае имеет вид (6.9), а план эксперимента представлен в табл. 6.3.

Таблица 6.3

Таблица полного факторного эксперимента для трех факторов

Номер опыта	План								Результат $y_i$
	$x_0$	$x_1$	$x_2$	$x_3$	$x_1 x_2$	$x_1 x_3$	$x_2 x_3$	$x_1 x_2 x_3$	
1	+1	-1	-1	-1	+1	+1	+1	-1	$y_1$
2	+1	+1	-1	-1	-1	-1	+1	+1	$y_2$
3	+1	-1	+1	-1	-1	+1	-1	+1	$y_3$
4	+1	+1	+1	-1	+1	-1	-1	-1	$y_4$
5	+1	-1	-1	+1	+1	-1	-1	+1	$y_5$
6	+1	+1	-1	+1	-1	+1	-1	-1	$y_6$
7	+1	-1	+1	+1	-1	-1	+1	-1	$y_7$
8	+1	+1	+1	+1	+1	+1	+1	+1	$y_8$

Здесь добавлен столбец фиктивной переменной  $x_0$ , нужный для оценки свободного члена  $b_0$ . После реализации плана получают 4 уравнения с 4 неизвестными, их решение и даст оценку всех 8 коэффициентов регрессии  $b_0, b_1, \dots, b_3, b_{12}, \dots, b_{123}$ .

**План, в котором число опытов равно числу определяемых коэффициентов, называется насыщенным.**

Заметим, что мы использовали все точки с "крайними" координатами, т.е.  $\pm 1$ , или говоря другими словами, все возможные комбинации выбранных уровней. В самом деле, всех возможных комбинаций  $2^k=8$ , ( $k$  — число факторов) и мы все их использовали. Если эксперименты проводятся только на двух уровнях (при двух значениях факторов) и при этом в процессе эксперимента осуществляются все возможные неповторяющиеся комбинации из  $k$  факторов, то постановка опытов по такому плану носит название полного факторного эксперимента (ПФЭ) или  $2^k$ .

Иными словами, **полный факторный эксперимент (ПФЭ) — это эксперимент, реализующий все возможные неповторяющиеся комбинации уровней независимых факторов.**

Кодированный план геометрически может быть интерпретирован в виде куба, восемь вершин которого представляют собой восемь экспериментальных точек (рис.6.1).

При числе факторов  $k=2$  построение матрицы ПФЭ не вызывает затруднений, при увеличении же числа факторов возникает необходимость в некоторых специальных приемах построения матрицы.

Первый прием основан на чередовании знаков. В первом столбце (для  $x_1$ ) знаки чередуются поочередно. Во втором (для  $x_2$ ) — через 2, в третьем (для  $x_3$ ) — через 4 и т.д. по степеням двойки  $2^k$ . Этот подход и использован при составлении плана, представленного в табл.6.3.

Второй прием основан на последовательном достраивании матрицы. Для этого при добавлении нового фактора необходимо повторить комбинации уровней исходного плана — сначала при значениях нового фактора на верхнем уровне, а затем — на нижнем.

Матрица ПФЭ обладает следующими свойствами.

1. Свойство симметричности: алгебраическая сумма элементов вектор-столбца каждого фактора равна нулю (за исключением столбца соответствующего свободному члену)

$$\sum_{j=1}^n X_{ij} = 0, \quad (6.12)$$

где  $i$  — номер фактора;  $j$  — номер опыта.

2. Свойство нормирования: сумма квадратов элементов каждого столбца равна числу опытов

$$\sum_{j=1}^n X_{ij}^2 = n. \quad (6.13)$$

3. Свойство ортогональности: скалярное произведение всех вектор-столбцов (сумма почленных произведений любых двух вектор-столбцов матрицы) равно нулю.

$$\sum_{j=1}^n X_{ij}X_{uj} = 0, \quad i \neq u. \quad (6.14)$$

**Планы, для которых выполняется свойство 3, называют ортогональными.** Благодаря этому свойству резко уменьшаются трудности, связанные с расчетом коэффициентов уравнения регрессии.

Поскольку результаты наблюдений отклика носят случайный характер, приходится в каждой точке плана проводить не один, а  $m^*$  параллельных опытов (обычно  $m^*=2-4$ ), осреднение результатов которых, как уже отмечалось, дает возможность уменьшить погрешности оценки истинного значения отклика в  $\sqrt{m^*}$  раз.

В каждой серии экспериментов их последовательность рандомизируется, т.е. с помощью таблиц случайных чисел определяется случайная последовательность реализации экспериментов. Рандомизация дает возможность свести эффект некоторого случайного фактора к случайной погрешности. Это позволяет в определенной степени исключить предвзятость и субъективизм исследователя, а также систематические ошибки, связанные, например, с разогревом или охлаждением агрегатов и др. факторов при исследовании процесса.

### 6.3.3. Определение коэффициентов уравнения регрессии

Вспользуемся свойствами ПФЭ для определения коэффициентов уравнения регрессии методом наименьших квадратов  $\hat{y} = b_0 + b_1x_1 + b_2x_2$ .

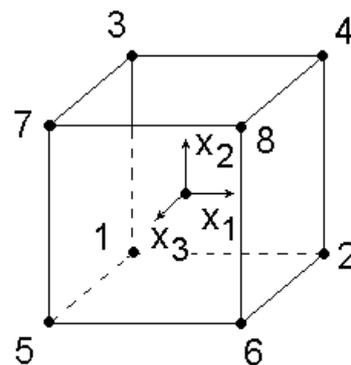


Рис.6.1. Геометрическое изображение ПФЭ

$$\Phi = \sum_{j=1}^n (y_j - \hat{y}_j)^2 \rightarrow \min_{b_i};$$

$$\frac{\partial \Phi}{\partial b_1} = 2 \sum_{j=1}^n (y_j - b_0 - b_1 X_{1j} - b_2 X_{2j}) X_{1j} = 0; \quad (6.15)$$

$$\sum_{j=1}^n y_j X_{1j} - b_0 \sum_{j=1}^n X_{1j} - b_1 \sum_{j=1}^n X_{1j}^2 - b_2 \sum_{j=1}^n X_{1j} X_{2j} = 0.$$

Воспользуемся свойствами ПФЭ:

- (симметричности)  $b_0 \sum X_{1j} = 0;$
- (нормирования)  $b_1 \sum X_{1j}^2 = nb_1;$
- (ортогональности)  $b_2 \sum X_{1j} X_{2j} = 0;$

$$b_1 = \frac{\sum_{j=1}^n y_j X_{1j}}{n}; \quad b_2 = \frac{\sum_{j=1}^n y_j X_{2j}}{n}; \quad b_0 = \frac{\sum_{j=1}^n y_j X_{0j}}{n}. \quad (6.16)$$

Следовательно, **любые коэффициенты уравнения регрессии определяются скалярным произведением столбца у на соответствующий столбец X.**

Можно показать, что аналогичным образом определяются и коэффициенты, если в уравнении регрессии (6.6) учитываются линейные взаимодействия (двойные, тройные)

$$b_{12} = \frac{\sum_{j=1}^n y_j (X_1 X_2)_j}{n}; \quad b_{123} = \frac{\sum_{j=1}^n y_j (X_1 X_2 X_3)_j}{n} \text{ и т.д.} \quad (6.17)$$

Следует обратить особое внимание на то, что **все линейные коэффициенты независимы**, так как в формулы для их расчета (6.16), (6.17) входят свои одноименные переменные. Поэтому **каждый коэффициент характеризует роль соответствующей переменной в процессе или силу влияния факторов.** Чем больше численная величина коэффициента, тем большее влияние оказывает этот фактор. Если коэффициент имеет знак плюс, то с увеличением значения фактора отклик увеличивается, а если минус — уменьшается.

В результате получения уравнения регрессии может получиться так, что один (или несколько) коэффициентов не очень большие и окажутся незначимыми. Факторы, имеющие коэффициенты, незначимо отличающиеся от нуля, могут быть выведены из состава уравнения, так как их влияние на параметры отклика будет отнесено к ошибке эксперимента. **Учитывая ортогональность плана оставшиеся коэффициенты уравнения регрессии можно не пересчитывать.** При отсутствии ортогональности плана эксперимента все коэффициенты необходимо пересчитывать заново.

#### 6.3.4. Статистический анализ результатов эксперимента

Планирование эксперимента исходит из статистического характера зависимостей, поэтому полученные уравнения подвергаются тщательному статистическому анализу с целью извлечь из результатов эксперимента максимум информации и убедиться в достоверности полученной зависимости и ее точности. Процедура проверки значимости коэффициентов уравнения регрессии и его адекватности принципиально не отличается от описания, данного в параграфах 4.5.1 и 4.5.2., поэтому остановимся только на отдельных моментах. Как уже отмечалось ранее, каждый эксперимент несет в себе какую-то погрешность, для повышения надежности результатов производятся повторения опытов при тех же условиях, т.е. повторяются опыты  $m^*$  раз для каждой строки таблицы планирования.

Построчные (выборочные) дисперсии подсчитываются по формуле

$$S_j^2 = \frac{\sum_{i=1}^{m^*} (y_{ji} - \bar{y}_j)^2}{m^* - 1}, \quad (6.18)$$

где  $m^*$  — число опытов в точках плана;  $\bar{y}_j = \frac{\sum_{i=1}^{m^*} y_{ji}}{m^*}$  — средний отклик по  $m^*$  опытам в точке с номером  $j$ .

Дисперсия воспроизводимости (отклика)  $S_{\text{восп}}^2$  есть среднеарифметическое дисперсий всех  $n$  различных вариантов опытов

$$S_{\text{восп}}^2 = \frac{\sum_{j=1}^n S_j^2}{n} = \frac{\sum_{j=1}^n \sum_{i=1}^{m^*} (y_{ji} - \bar{y}_j)^2}{n(m^* - 1)}, \quad (6.19)$$

которая представляет собой среднюю арифметическую величину из  $n$  дисперсий различных вариантов эксперимента.

Прежде чем производить объединение дисперсий, следует убедиться в их однородности. Проверка производится с помощью критерия Фишера или Кохрена (см. гл. 3). Для оценки значимости коэффициентов прежде всего находят дисперсию коэффициентов регрессии. Учитывая свойства 1-3 плана табл. 6.3 из выражений (4.27а) и (4.27б) при одинаковом дублировании опытов по точкам с числом повторных опытов  $m^*$  она определяется по формуле

$$S_b^2 = \frac{S_{\text{восп}}^2}{m^* n}, \quad (6.20)$$

а при отсутствии дублирования по соотношению

$$S_b^2 = \frac{S_{\text{восп}}^2}{n}. \quad (6.20a)$$

Следовательно, все коэффициенты уравнения регрессии ПФЭ имеют одинаковую точность (дисперсию). Это принципиальное отличие коэффициентов уравнения регрессии, полученных по плану табл.6.3, от коэффициентов уравнений, полученных пассивным экспериментом (см. параграф 4.5.2). **Планы, по результатам которых коэффициенты уравнения регрессии определяются с одинаковой дисперсией, называются ротатабельными.** В связи с этим план, представленный в табл.6.3, является не только ортогональным, но ротатабельным. В дальнейшем проверка значимости каждого коэффициента производится независимо с использованием  $t$ -критерия Стьюдента (см. гл.4). Статистически незначимые коэффициенты исключаются из уравнения, а остальные коэффициенты при этом не пересчитываются. После этого уравнение регрессии составляется в виде уравнения связи выходного параметра  $y$  и переменных  $X_i$ , включающего только значимые коэффициенты.

После вычисления коэффициентов уравнения следует прежде всего проверить его пригодность или адекватность. Для этого достаточно оценить отклонение выходной величины  $y$ , предсказанной уравнением регрессии, от результатов эксперимента  $y$  в различных точках.

Рассеяние результатов эксперимента относительно уравнения регрессии, аппроксимирующего искомую зависимость, можно, как уже было показано ранее, охарактеризовать с помощью остаточной дисперсии (дисперсии адекватности), оценка которой, справедливая при одинаковом числе дублирующих опытов, находится по формуле

$$S_{\text{ад}}^2 = \frac{m^* \sum_{j=1}^n (\bar{y}_j - \hat{y}_j)^2}{n - 1}. \quad (6.21)$$

Здесь  $n$  — число опытов (вариантов);  $l=k+1$ , где  $k$  — число членов в уравнении регрессии.

Проверка адекватности состоит в выяснении соотношения между дисперсией адекватности  $S_{ад}^2$  и дисперсии воспроизводимости  $S_{восп}^2$  и проводится с помощью F-критерия Фишера, который в данном случае рассчитывается как

$$F_{эксп} = \frac{S_{ад}^2}{S_{восп}^2}. \quad (6.22)$$

Если вычисленное значение критерия меньше теоретического  $F_{\alpha; m_1; m_2}$  для соответствующих степеней свободы  $m_1 = n(m^* - 1)$ ,  $m_2 = n - 1$  при заданном уровне значимости  $\alpha$ , то описание свойств объекта уравнением регрессии признается адекватным объекту. Адекватность модели может быть достигнута уменьшением интервала варьирования факторов, а если это не дает результата, то переходом к плану второго порядка.

### 6.3.5. Дробный факторный эксперимент

Во многих практических задачах взаимодействия второго и высших порядков отсутствуют или пренебрежимо малы. Кроме того, на первых этапах исследования часто необходимо получить в первом приближении лишь линейную аппроксимацию изучаемого уравнения связи при минимальном числе экспериментов. Так, для трех факторов вместо уравнения (6.9) достаточно рассмотреть уравнение вида

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + b_3x_3 \quad (6.23)$$

и достаточно определить только четыре коэффициента.

Поэтому использовать ПФЭ для определения коэффициентов только при линейных членах не эффективно из-за реализации большого числа опытов, особенно при большом числе факторов  $k$ .

Если при решении задачи можно ограничиться линейным приближением, то в ПФЭ оказывается много "лишних" опытов. Так, для трех факторов достаточно 4 опыта, а в ПФЭ их 8. Следовательно, есть четыре "лишних". Результаты этих "лишних" опытов могут быть использованы двояко: во-первых, с их помощью можно получить более точные оценки коэффициентов регрессии; во-вторых, их можно использовать для проверки адекватности модели. Однако при 7 факторах ПФЭ содержит  $2^7 = 128$  опытов, а для линейного уравнения требуется всего 8. Таким образом, остается 120 лишних, и конечно, нет необходимости их все реализовать, а достаточно лишь несколько из них использовать для проверки адекватности и уточнения оценок.

Другими словами, ПФЭ обладает большой избыточностью опытов. В связи с этим возникает вопрос: "Нельзя ли сократить число опытов необходимых для определения коэффициентов регрессии?"

Так, для определения коэффициентов уравнения (6.23) достаточно ограничиться четырьмя опытами, если в ПФЭ  $2^3$  использовать  $x_1x_2$  в качестве плана для  $x_3$  и матрица планирования эксперимента примет вид, представленный в табл.6.4.

Таблица 6.4

Дробный факторный эксперимент

Номер опыта	План				Результат $y$
	$x_0$	$x_1$	$x_2$	$x_3 = x_1x_2$	
1	+1	-1	-1	+1	$y_1$
2	+1	+1	-1	-1	$y_2$
3	+1	-1	+1	-1	$y_3$
4	+1	+1	+1	+1	$y_4$

Заметим, что мы использовали не все точки с "крайними" координатами, т.е.  $\pm 1$ , или говоря другими словами, не все возможные комбинации выбранных уровней. В самом деле, всех возможных комбинаций  $2^3 = 8$ , мы же использовали из них только 4.

**Такой сокращенный план — носит название дробного факторного эксперимента (ДФЭ).** Следует подчеркнуть, что формальное приравнивание произведения факторов фактору, не

входящему в это произведение, является основополагающей идеей метода ДФЭ. В данном случае используется только половина ПФЭ  $2^3$ , поэтому план, представленный в табл.6.4, называется полуреplikой от ПФЭ  $2^3$ . После реализации плана получают 4 уравнения с 4 неизвестными, их решение и даст оценку всех четырех коэффициентов регрессии  $b_i$ . Например, матрица из 8 опытов для четырехфакторного планирования будет полуреplikой от ПФЭ  $2^4$ , а для пятифакторного планирования — четвертьреplikой от  $2^5$ .

**Для того, чтобы дробная реплика представляла собой ортогональный план, в качестве реплика следует брать ближайший полный факторный эксперимент.** При этом число опытов должно быть больше числа коэффициентов.

Если коэффициенты регрессии при парных произведениях не равны нулю, то найденные коэффициенты  $b_i$  будут смешанными оценками их теоретических коэффициентов  $\beta_i$ . На практике обычно не удается априорно постулировать равенство нулю эффектов взаимодействия, однако часто имеются основания полагать, что некоторые из них малы по сравнению с линейными эффектами. Операцию смешивания оценок принято условно записывать в виде выражений

$$b_1 = \beta_1 + \beta_{12}; \quad b_2 = \beta_2 + \beta_{13}; \quad b_3 = \beta_3 + \beta_{23}, \quad (6.24)$$

где  $\beta$  — математическое ожидание для соответствующего коэффициента.

Эти генерирующие коэффициенты не могут быть отдельно оценены по плану, включающему всего четыре столбца, т.к. в этом случае неразличимы столбцы для линейных членов и парных произведений. Если, например, в дополнение к столбцам, вычислить еще столбцы для произведения  $x_1x_3$ , то увидим, что элементы этого столбца в точности равны элементам столбца  $x_2$ . Таким образом, сокращение числа опытов приводит к получению смешанных оценок для коэффициентов.

Для того, чтобы определить, какие коэффициенты смешаны, удобно пользоваться следующим приемом: подставив  $x_3$  на место  $x_1x_2$ , получим соотношение  $x_3 = x_1x_2$ , называемое генерирующим соотношением.

Умножив обе части генерирующего соотношения на  $x_3$ , получим

$$x_3^2 = x_1x_2x_3 = 1, \text{ т.е. } x_1x_2x_3 = 1. \quad (6.25)$$

Это произведение носит название определяющего контраста.

Умножив поочередно определяющий контраст на  $x_1, x_2, x_3$ , находим

$$x_1 = x_1^2x_2x_3 = x_2x_3; \quad x_2 = x_1x_3; \quad x_3 = x_1x_2. \quad (6.26)$$

Полученным соотношениям соответствует система смешанных оценок, т.е.  $\beta_1$  смешана с  $\beta_{23}$ ,  $\beta_2$  — с  $\beta_{13}$ , а  $\beta_3$  — с  $\beta_{12}$ .

Кроме соотношения  $x_3 = x_1x_2$ , возможны следующие случаи приравнивания одних факторов к взаимодействиям других факторов:  $x_2 = x_1x_3$ ,  $x_1 = x_2x_3$ . Соответствующий набор возможных операций смешанных оценок коэффициентов модели для них примет вид:  $b_2 \rightarrow \beta_2 + b_{13}$ ,  $b_1 \rightarrow \beta_1 + b_{23}$ .

Таким образом, при использовании ДФЭ необходимо иметь четкое представление о так называемой разрешающей способности дробных реплик, т.е. определить заранее, какие коэффициенты являются несмешанными оценками для соответствующих коэффициентов. Тогда в зависимости от постановки задачи подбирается дробная реплика, с помощью которой можно извлечь максимальную информацию из эксперимента.

Например, в задаче с четырьмя факторами  $k=4$  в качестве генерирующего соотношения можно взять  $x_4 = x_1x_2x_3$  или любой из эффектов двойного взаимодействия, например  $x_4 = x_1x_2$ . Таблица планирования эксперимента представлена в табл.6.5.

## Планирование ДФЭ

Номер опыта	План				Генерирующие соотношения	
	$x_0$	$x_1$	$x_2$	$x_3$	$x_4=x_1x_2x_3$	$x_4=x_1x_2$
1	+1	-1	-1	-1	-1	+1
2	+1	+1	-1	-1	+1	-1
3	+1	-1	+1	-1	+1	-1
4	+1	+1	+1	-1	-1	+1
5	+1	-1	-1	+1	+1	+1
6	+1	+1	-1	+1	-1	-1
7	+1	-1	+1	+1	-1	-1
8	+1	+1	+1	+1	+1	+1

В первом случае определяющий контраст  $X_4^2=X_1X_2X_3X_4=1$  и получим оценку совместных оценок

$$\begin{aligned} X_1 &= X_2X_3X_4 \rightarrow b_1 \rightarrow \beta_1 + \beta_{234}; \\ X_2 &= X_1X_3X_4 \rightarrow b_2 \rightarrow \beta_2 + \beta_{134}; \\ X_3 &= X_1X_2X_4 \rightarrow b_3 \rightarrow \beta_3 + \beta_{124}; \\ X_4 &= X_1X_2X_3 \rightarrow b_4 \rightarrow \beta_4 + \beta_{123}; \\ X_1X_4 &= X_2X_3 \rightarrow b_{14} \rightarrow \beta_{14} + \beta_{23}; \\ X_1X_2 &= X_3X_4 \rightarrow b_{12} \rightarrow \beta_{12} + \beta_{34}; \\ X_1X_3 &= X_2X_4 \rightarrow b_{13} \rightarrow \beta_{13} + \beta_{24}. \end{aligned}$$

В реальных задачах тройные взаимодействия бывают равным нулю значительно чаще, чем двойные. Значит, если по физическому смыслу задачи более всего интересуют оценки для линейный эффектов, следует брать генерирующее соотношение  $X_4 = X_1X_2X_3$ .

Во втором случае определяющий контраст выражается соотношением  $X_4^2=X_1X_2X_4=1$ ;  $X_1X_2X_4=1$ .

При этом получим следующую систему оценок.

$$\begin{aligned} X_1 &= X_2X_4; \quad b_1 \rightarrow \beta_1 + \beta_{24}; \\ X_2 &= X_1X_4; \quad b_2 \rightarrow \beta_2 + \beta_{14}; \\ X_3 &= X_1X_2X_3X_4; \quad b_3 \rightarrow \beta_3 + \beta_{1234}; \\ X_4 &= X_1X_2; \quad b_4 \rightarrow \beta_4 + \beta_{12}; \\ X_1X_2 &= X_2X_3X_4; \quad b_{12} \rightarrow \beta_{12} + \beta_{234}; \\ X_2X_3 &= X_1X_3X_4; \quad b_{23} \rightarrow \beta_{23} + \beta_{134}; \\ X_3X_4 &= X_1X_2X_3; \quad b_{34} \rightarrow \beta_{34} + \beta_{123}. \end{aligned}$$

Следовательно, дробную реплику с генерирующим соотношением  $X_4=X_1X_2$  имеет смысл использовать, если нас более всего интересуют коэффициенты  $\beta_{12}, \beta_{23}, \beta_{34}$ .

Дробную реплику, в которой  $P$  линейных эффектов приравнены к эффектам взаимодействия, обозначают  $2^{k-P}$ .

Таким образом, планы первого порядка, оптимальные двухуровневые планы ПФЭ  $2^k$  и ДФЭ  $2^{k-P}$  имеют следующие преимущества:

- 1 — планы ортогональны, поэтому все вычисления просты;
- 2 — все коэффициенты определяются независимо один от другого;
- 3 — каждый коэффициент определяется по результатам всех  $n$  опытов;
- 4 — все коэффициенты регрессии определяются с одинаковой дисперсией, т.е. эти планы обладают одновременно и свойством ротатабельности.

#### 6.4. Планы второго порядка

Часто для описания поверхности отклика полиномами первого порядка уже недостаточно. Во многих случаях удовлетворительная аппроксимация может быть достигнута, если воспользоваться полиномом второго порядка (6.7).

В этом случае требуется, чтобы каждый фактор варьировался не менее, чем на трех уровнях. В этом случае полный факторный эксперимент содержит слишком большое количество опытов, равное  $3^k$ . Так при  $k=3$  их 27, а число коэффициентов  $b$  — 10, при  $k=5$  число опытов 243, а коэффициентов 21. В связи с этим осуществить ПФЭ для планов второго порядка не только сложно, но и нецелесообразно.

Сократить число опытов можно воспользовавшись так называемым композиционным или последовательным планом, разработанным Боксом и Уилсоном. Так, при двух факторах модель функции отклика  $y=f(x_1, x_2)$  второго порядка представляет собой поверхность в виде цилиндра, конуса, эллипса и т.д., описываемую в общем виде уравнением

$$y = b_0 + b_1x_1 + b_2x_2 + b_{11}x_1^2 + b_{22}x_2^2 + b_{12}x_1x_2.$$

Для определения такой кривой необходимо располагать координатами не менее трех ее точек, т.е. факторы  $x_1$  и  $x_2$  должны варьироваться не менее, чем на трех уровнях. Поэтому план эксперимента в плоскости факторов  $x_1$  и  $x_2$  на рис.6,а не может состоять лишь из опытов 1, 2, 3, 4 ПФЭ  $2^2$ , располагающихся в вершинах квадрата как это было для модели первого порядка, а к ним должны быть добавлены опыты (звездные точки) 5, 6, 7, 8, расположенные на осях  $x_1$  и  $x_2$  с координатами  $(\pm\alpha; 0)$ ,  $(0; \pm\alpha)$  и обязательно опыт 9 в центре квадрата, чтобы по любому направлению (5-9-7), (1-9-4) и т.д. располагалось три точки, определяющие кривизну поверхности в этом направлении.

Таким образом, в общем случае ядро композиционного плана составляет при  $k < 5$  ПФЭ  $2^k$  при  $k \geq 5$  — дробная реплика от него. Если линейное уравнение регрессии оказалось неадекватным, необходимо:

1) добавить  $2 \cdot k$  звездных точек, расположенных на координатных осях факторного пространства  $(\pm\alpha, 0, 0, \dots, 0)$ ,  $(0, \pm\alpha, 0, \dots, 0)$ , ...,  $(0, 0, \dots, \pm\alpha)$ , где  $\alpha$  — звездное плечо, или расстояние до звездной точки;

2) провести  $n_0$  опытов при значениях факторов в центре плана.

При  $k$  факторах общее число опытов в матрице композиционного плана составит

$$n = 2^k + 2 \cdot k + n_0 \text{ при } k < 5,$$

$$n = 2^{k-1} + 2 \cdot k + n_0 \text{ при } k \geq 5.$$

При этом величина звездного плеча  $\alpha$  и число опытов в центре плана  $n_0$  зависит от выбранного вида композиционного плана.

Композиционный план для  $k=2$  и  $n_0=1$  представлен в табл.6.6.

Таблица 6.6

Композиционный план второго порядка

Номер опыта	Ф а к т о р ы						Результат $y_i$
	$x_0$	$x_1$	$x_2$	$x_1x_2$	$x_1^2$	$x_2^2$	
1	+1	-1	-1	+1	+1	+1	$y_1$
Ядро 2	+1	+1	-1	-1	+1	+1	$y_2$
плана 3	+1	-1	+1	-1	+1	+1	$y_3$
4	+1	+1	+1	+1	+1	+1	$y_4$
5	+1	$+\alpha$	0	0	$\alpha^2$	0	$y_5$
Звезд- 6	+1	$-\alpha$	0	0	$\alpha^2$	0	$y_6$
ные 7	+1	0	$+\alpha$	0	0	$\alpha^2$	$y_7$
точки 8	+1	0	$-\alpha$	0	0	$\alpha^2$	$y_8$

Центр 9 плана	+1	0	0	0	0	0	$y_9$
------------------	----	---	---	---	---	---	-------

Аналогичным образом составляются планы и для большего числа факторов.

#### 6.4.1. Ортогональные планы второго порядка

В общем виде данная матрица неортогональна, т.к.

$$\sum_{j=1}^n x_{0j} x_{uj}^2 \neq 0; \quad \sum_{j=1}^n x_{ij}^2 x_{uj}^2 \neq 0. \quad (6.27)$$

Приведем матрицу к ортогональному виду, для чего введем новые переменные (преобразования для квадратичных эффектов)

$$x'_{ij} = x_{ij}^2 - \frac{\sum_{j=1}^n x_{ij}^2}{n} = x_{ij}^2 - \bar{x}_i^2.$$

При этом  $\sum_{j=1}^n x_{0j} x'_{ij} = \sum_{j=1}^n (x_{ij}^2 - \bar{x}_i^2) = \sum_{j=1}^n x_{ij}^2 - n\bar{x}_i^2 = 0.$

Тогда уравнение регрессии (модель процесса) будет записана как

$$\hat{y} = b_0' + \sum_{i=1}^k b_i x_i + \sum_{i,j=1}^k b_{ij} x_i x_j + \sum_{i=1}^k b_{ii}' x_i'.$$

Композиционные планы легко привести к ортогональным, выбирая звездное плечо  $\alpha$ . В табл. 6.9 приведено значение  $\alpha$  для различного числа факторов  $k$  и числа опытов в центре плана  $n_0$ .

Таблица 6.7

Число опытов в центре плана, $n_0$	Звездное плечо $\alpha$ при различном числе факторов $k$		
	$k=2$	$k=3$	$k=4$
1	1,00	1,22	1,41
2	1,08	1,29	1,47
3	1,15	1,35	1,55
4	1,21	1,41	1,61

В частности, ортогональный план второго порядка для  $k=2$  и  $n_0=1$  представлен в табл. 6.8, а его геометрическая интерпретация на рис. 6.3а.

Представленный на рис.6.3а и табл. 6.10 прямоугольный (квадратный) план эксперимента для модели второго порядка работоспособен, хотя и несколько избыточен (9 опытов для определения 6 коэффициентов). Благодаря трем избыточным опытам он позволяет усреднить случайные погрешности и оценить их характер.

Таблица 6.8

Ортогональный план второго порядка

Номер опыта	Ф а к т о р ы						Результат $y_i$
	$x_0$	$x_1$	$x_2$	$x_1 x_2$	$x_1'$	$x_2'$	
1	+1	-1	-1	+1	+1/3	+1/3	$y_1$
Ядро 2	+1	+1	-1	-1	+1/3	+1/3	$y_2$
плана 3	+1	-1	+1	-1	+1/3	+1/3	$y_3$
4	+1	+1	+1	+1	+1/3	+1/3	$y_4$
5	+1	$\alpha=+1$	0	0	+1/3	-2/3	$y_5$

Звезд- 6 ные 7 точки 8	+1	$\alpha=-1$	0	0	+1/3	-2/3	$y_6$
Центр 9 плана	+1	0	0	0	-2/3	-2/3	$y_9$

В этой таблице  $x'_{uj} = x_{uj} - \frac{\sum_{j=1}^9 x_{uj}}{9} = x_{uj} - \frac{2}{3}$ .

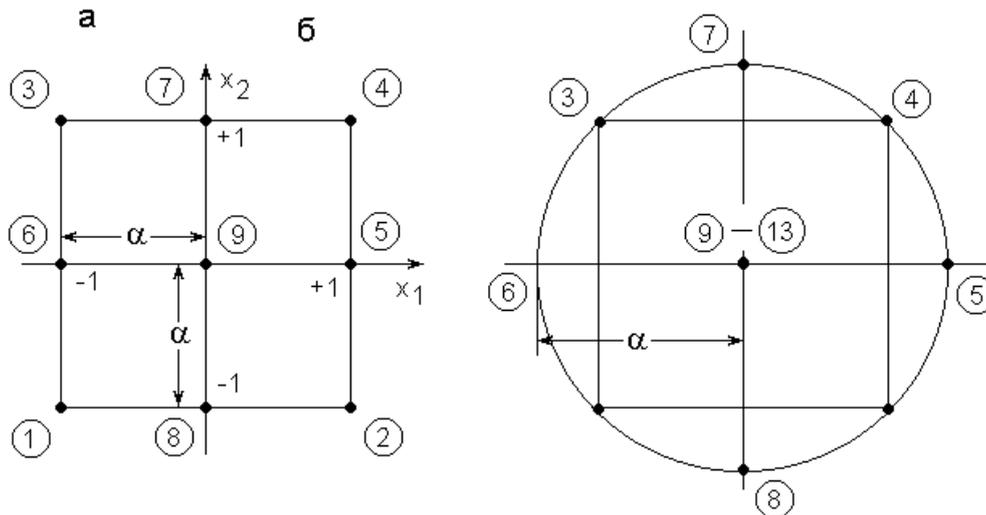


Рис. 6.3. Планы второго порядка при  $k=2$ :

*a* — ортогональный; *б* — ротатабельный

В силу ортогональности матрицы планирования все коэффициенты уравнения регрессии  $b$  определяются независимо один от другого по формуле:

$$b_i = \frac{\sum_{j=1}^n x_{ij} y_j}{\sum_{j=1}^n x_{ij}^2}; \quad b_{ii}' = \frac{\sum_{j=1}^n x_{ij}' y_j}{\sum_{j=1}^n x_{ij}'^2}. \quad (6.28)$$

Здесь  $i$  — номер столбца в матрице планирования;  $j$  — номер строки; сумма  $\sum_{j=1}^n x_{ij}^2$

различная для линейных, квадратичных эффектов и взаимодействий.

Дисперсии коэффициентов уравнения регрессии равны

$$S_{b,i}^2 = S_{\text{восп}} / \sum_{j=1}^n x_{ij}^2. \quad (6.29)$$

Следует особо отметить, что коэффициенты уравнения регрессии, получаемые с помощью ортогональных планов второго порядка, определяются с разной точностью (см. уравнение 6...), в то время как ортогональные планы первого порядка обеспечивают одинаковую точность коэффициентов, т.е. хотя план 6... и является ортогональным и обеспечивает независимость определения коэффициентов  $b$ , но не является ротатабельным !

В результате расчетов по матрице с преобразованными столбцами для квадратичных эффектов получим уравнение в виде

$$\hat{y} = b_0' + \sum_{i=1}^k b_i x_i + \sum_{i,j=1}^k b_{ij} x_i x_j + \sum_{i=1}^k b_{ii}' (x_i^2 - \bar{x}_i^2). \quad (6.30)$$

Для преобразования к обычной записи следует перейти от коэффициента  $b_0'$  к коэффициенту  $b_0$ , используя выражение

$$b_0 = b_0' - \sum_{i=1}^k b_{ii} \bar{x}_i^2. \quad (6.31)$$

При этом дисперсия этого коэффициента рассчитывается по следующему соотношению

$$S_{b_0}^2 = S_{b_0'}^2 + \sum_{i=1}^k \bar{x}_i^2 S_{b_{ii}}^2. \quad (6.32)$$

В дальнейшем, зная дисперсию воспроизводимости, проверяют значимость коэффициентов и адекватность уравнения:

$$\hat{y} = b_0 + \sum_{i=1}^k b_i x_i + \sum_{i,j=1}^k b_{ij} x_i x_j + \sum_{i=1}^k b_{ii} x_i^2. \quad (6.33)$$

Значимость коэффициентов проверяется по критерию Стьюдента  $t_i = |b_i| / S_{b_i}$ . Коэффициент значим, если  $t_i > t_{\alpha, m}$ , где  $m$  — число степеней свободы дисперсии воспроизводимости.

Адекватность уравнения проверяется по критерию Фишера  $F = S_{ад}^2 / S_{восп}^2$ . Уравнение адекватно, если составленное таким образом F-отношение меньше теоретического  $F < F_{\alpha; m_1; m_2}$ , где  $m_1 = n - 1$  — число степеней свободы дисперсии адекватности;  $m_2$  — число степеней свободы дисперсии воспроизводимости; 1 — число коэффициентов в уравнении регрессии второго порядка, равного числу сочетаний из  $n+2$  по 2, т.е.

$$1 = \frac{(n+2)(n+1)}{2}.$$

#### 6.4.2. Ротатабельные планы второго порядка

Как мы с Вами установили план второго порядка, представленный в табл.6... не обладает свойством ротатабельности. Действительно, удаление от центра точек 5,6,7,8 в  $\sqrt{2} = 1,414$  раза меньше, чем удаление точек 1,2,3,4 (см.рис.6.3а), и следовательно коэффициенты уравнения регрессии определяются с различной дисперсией. Бокс и Хантер предложили ротатабельные планы 2-го порядка. Для того, чтобы композиционный план был ротатабельным, величину звездного плеча  $\alpha$  выбирают из условия

$$\alpha = 2^{\frac{k-p}{4}},$$

где  $k$  — число факторов;  $p$  — дробная реплика (для ПФЭ  $p=0$ , для полуреплики  $p=1$ , для четвертьреплики  $p=2$  и т.д.). Число точек в центре плана  $n_0$  увеличивают. В таблице 6.9 приведены значения  $\alpha$  и  $n_0$  для различного числа независимых факторов.

Таблица 6.9

Параметр плана	Значения параметров при числе независимых факторов								
	2	3	4	5	6	6	6	7	7
Ядро плана	2 <sup>2</sup>	23	24	25	25-1	26	26-1	27	27-1
Звездное плечо	1,414	1,682	2,00	2,378	2,00	2,828	2,378	3,333	2,828

Число точек в центре плана, $n_0$	5	6	7	10	6	15	9	21	14
-----------------------------------	---	---	---	----	---	----	---	----	----

Поясним идею выбора значения звездного плеча  $\alpha$  на примере матрицы ротатабельного планирования второго порядка для  $k=2$ , представленной в табл. 6.12. Размещение точек этого плана показано на рис.6.3б. Для обеспечения ротатабельности точек 5,6,7,8 необходимо удалить их от центра плана на расстояние  $\alpha$  в  $\sqrt{2}=1,414$  раз большее, чем удаление точек 1,2,3,4 от осей  $x_2$  и  $x_1$ . В результате этого все точки плана б... оказываются лежащими на окружности. Учитывая существенно большее влияние на функцию отклика случайной ошибки в точке 9, рекомендуется ставить в этой точке плана не один, а несколько дублирующих опытов (в данном случае опыты с 9 до 13) для усреднения полученных результатов и для осуществления статистического анализа результатов всего эксперимента в целом.

Таблица 6.10

Номер опыта	Ф а к т о р ы						Результат $y_j$
	$x_0$	$x_1$	$x_2$	$x_1x_2$	$x_1^2$	$x_2^2$	
1	+1	-1	-1	+1	+1	+1	$y_1$
Ядро 2	+1	+1	-1	-1	+1	+1	$y_2$
плана 3	+1	-1	+1	-1	+1	+1	$y_3$
4	+1	+1	+1	+1	+1	+1	$y_4$
Звезд- 5	+1	+1,414	0	0	2	0	$y_5$
ные 6	+1	-1,414	0	0	2	0	$y_6$
точки 7	+1	0	+1,414	0	0	2	$y_7$
8	+1	0	-1,414	0	0	2	$y_8$
9	+1	0	0	0	0	0	$y_9$
Центр 10	+1	0	0	0	0	0	$y_{10}$
плана 11	+1	0	0	0	0	0	$y_{11}$
12	+1	0	0	0	0	0	$y_{12}$
13	+1	0	0	0	0	0	$y_{13}$

Однако, матрица ротатабельного планирования второго порядка неортогональна, т.к.

$$\sum_{j=1}^n x_{0j} \cdot x_{uj}^2 \neq 0; \quad \sum_{j=1}^n x_{ij}^2 \cdot x_{uj}^2 = 0; \quad i \neq u. \quad (6.34)$$

Следовательно, **если какой-либо из квадратичных эффектов оказался незначимым, то после его исключения коэффициенты уравнения регрессии необходимо пересчитать заново!**

При использовании ротатабельных планов второго порядка дисперсию воспроизводимости можно определить по опытам в центре плана. В связи с этим при проверке адекватности уравнения регрессии, полученного по ротатабельному плану второго порядка, поступают следующим образом.

Находят остаточную сумму квадратов

$$S_1^2 = \sum_{j=1}^n (y_j - \bar{y}_j)^2 \quad (6.35)$$

с числом степеней свободы

$$m_1 = n - 1 = n - \frac{(k+2)(k+1)}{2}.$$

По опытам в центре плана определяют сумму квадратов воспроизводимости

$$S_2^2 = \sum_{j=1}^{n_0} (y_{0j} - \bar{y}_{0j})^2 \quad (6.36)$$

с числом степеней свободы  $m_2 = n_0 - 1$ .

Далее находят сумму квадратов, характеризующих неадекватность  $S_3^2 = S_1^2 - S_2^2$ , число степеней свободы которой равно

$$m_3 = m_1 - m_2 = n - \frac{(k+2)(k+1)}{2} - (n_0 - 1).$$

Проверяют адекватность по F-критерию

$$F = \frac{S_3^2/m_3}{S_2^2/m_2}.$$

Уравнение адекватно, если  $F < F_{\alpha; m_3; m_2}$ .

Если модель второго порядка оказалась неадекватной, следует повторить эксперименты на меньшем интервале варьирования факторов или перенести центр плана в другую точку факторного пространства. В тех случаях, когда адекватность модели по-прежнему не достигается, рекомендуется перейти к планам третьего порядка.

### **6.5. Планирование экспериментов при поиске оптимальных условий**

Во многих случаях инженерной практики перед исследователем возникает задача не только выявления характера связи между двумя или несколькими рядами наблюдений, а требуется найти такие численные значения факторов, при которых отклик (выходной параметр) достигает своего экстремального значения (максимума или минимума). Эксперимент, решающий эту задачу, называется экстремальным. В этом случае задача сводится к оптимизационной и формулируется следующим образом: требуется определить такие координаты экстремальной точки  $(x_1^*, x_2^*, \dots, x_n^*)$  поверхности отклика  $y=f(x_1, x_2, \dots, x_n)$ , в которой она максимальна (минимальна):  $\max y(x_1, x_2, \dots, x_n)=y(x_1^*, x_2^*, \dots, x_n^*)$ .

Графическая интерпретация задачи оптимизации объекта  $y(x_1, x_2, \dots, x_n)$  при двух факторах  $y(x_1, x_2)$  представлена на рис.6.5.1а,б. Здесь точка А соответствует оптимальным значениям факторов  $x_1^*$  и  $x_2^*$ , обеспечивающим максимум функции отклика  $y_{\max}$ . Замкнутые линии на рис.6.5.1б характеризуют линии постоянного уровня и описываются уравнением  $y=f(x_1, x_2)=B=\text{const}$ .

Необходимость в экстремальных экспериментах довольно часто возникает в инженерной практике. Так, на модели шахтной печи с противоточно движущимся плотным продуваемым слоем, схема которой представлена на рис.6.5.2, требуется определить при заданных ее конструктивных и режимных параметрах расположение фурмы по высоте печи  $h$ , ее диаметр  $d$  и высов  $l$ , обеспечивающих максимальную степень использования теплового потенциала газового потока. В данном случае, факторами являются  $l, h, d$ , а в качестве функции отклика  $y(l, h, d)$  в первом приближении можно использовать температуру отходящих из печи газов. Заметим, что вид функции отклика в данном случае  $f(x_1, \dots, x_n)$  исследователю заранее неизвестен, т.е. отсутствует математическая модель, адекватно описывающая данный процесс. Требуется с наименьшими затратами (минимальном числе опытов) определить оптимальные значения  $l^*, h^*, d^*$ , при которых температура отходящих газов минимальна.

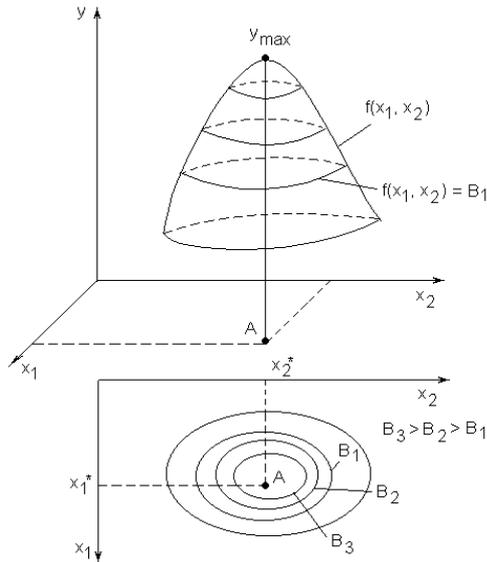


Рис.6.5.1. Поверхность отклика (а) и линии равного уровня (б):  $y=f(x_1, x_2)=B=const$  для  $n=2$

Известный из практики метод "проб" и "ошибок", в котором факторы изменяются на основании опыта, интуиции или наугад, при обычно имеющем место значительном числе факторов при исследовании процессов в металлургии зачастую оказывается малоэффективным вследствие весьма сложной зависимости функции отклика от факторов.

Требуют значительно меньшего числа опытов и быстрее приводят к цели те поисковые методы оптимизации, где шаговое варьирование факторами производится целенаправленно по определенному плану. Поисковые методы оптимизации относятся к классу итерационных процедур, при этом весь процесс разбивается на шаги, на каждом шаге делается ряд опытов и определяется каким образом изменить факторы, влияющие на процесс, чтобы получить улучшение результата. При этом на каждом очередном шаге получаемая информация используется для выбора последующего шага.

Разработано множество методов пошаговой оптимизации, которые подробно рассматриваются в разделе вычислительной математики — "Численные методы оптимизации". Мы же рассмотрим только некоторые из них, эффективность использования которых в промышленном и лабораторном эксперименте подтверждена практикой применительно к металлургическим процессам.

### 6.5.1. Метод покоординатной оптимизации (Гаусса - Зейделя)

Процесс поиска оптимума методом покоординатной оптимизации (Гаусса-Зейделя) в графическом виде для двумерного случая представлен на рис.6.5.3. По этому методу выбирается произвольная точка  $M_0$  и определяются ее координаты. Поиск оптимума осуществляется поочередным варьированием каждого из факторов. При этом сначала изменяют один фактор ( $x_1$ ) при фиксированных остальных ( $x_2=const$ ) до тех пор, пока не прекращается прирост функции отклика (точка  $M_1$ ). В дальнейшем изменяется другой фактор ( $x_2$ ) при фиксированных остальных ( $x_1=const$ ) и далее процедура повторяется.

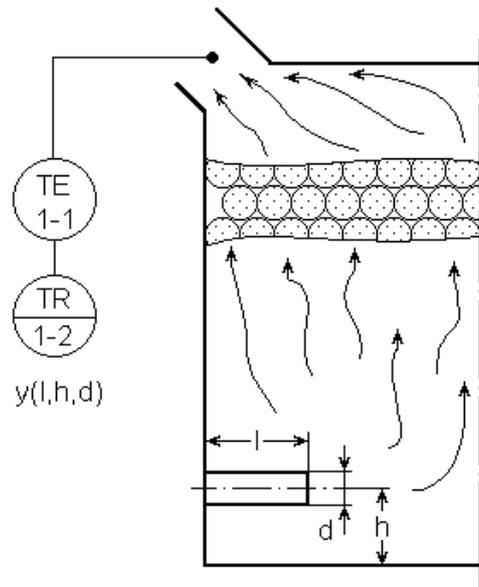


Рис.6.5.2. Схема шахтной печи

Данный метод весьма прост, однако при большом числе факторов требуется значительное число опытов, чтобы достичь координат оптимума. Более того, при некоторых зависимостях  $y=f(x_1, \dots, x_n)$  этот метод может привести к ложному результату. На рис.6.5.3 показан один из таких частных случаев, когда поочередное изменение каждого из факторов в любую сторону вдоль координатных осей  $x_1$  и  $x_2$  вызывает уменьшение  $y$ . В результате решения находится ложный экстремум, находящийся в точке  $A_1$  с координатами  $\tilde{x}_1; \tilde{x}_2$ , в то время как действительное значение максимума  $y_{\max}$  находится в точке  $A$  с координатами  $x_1^*$  и  $x_2^*$ .

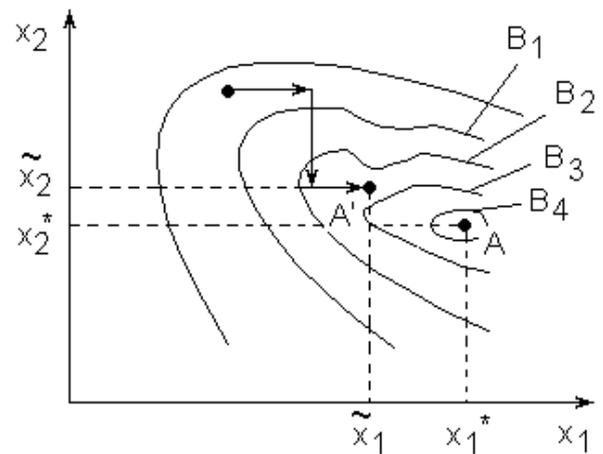


Рис.6.5.3. К методу Гаусса-Зейделя

В дальнейшем рассмотрим более совершенные методы.

### 6.5.2. Метод крутого восхождения (Бокса-Уилсона)

Известно, что кратчайший, наиболее короткий путь — это движение по градиенту, т.е. перпендикулярно линиям равного уровня, на которых функция отклика принимает постоянные значения  $y(x_1, x_2, \dots, x_n)=B$ . В связи с этим при оптимизации процесса рабочее движение целесообразно совершать в направлении наиболее быстрого возрастания функции отклика, т.е. в направлении градиента функции  $y$ .

Существуют различные модификации градиентного метода, одним из них является метод крутого восхождения (Бокса-Уилсона). Сущность этого метода также рассмотрим на примере двухфакторной задачи (рис.6.5.4). В этом случае шаговое движение осуществляется в направлении наискорейшего возрастания функции отклика, т.е.  $\text{grad } y(x_1, x_2)$ . Однако направление корректируют не после каждого следующего шага, а при достижении в некоторой точке на данном направлении частного экстремума функции отклика.

Пусть в окрестности точки  $M_0$  как центра плана поставлен ПФЭ  $2^2$ . Координаты отдельных опытов соответствуют точкам 1-4. По результатам ПФЭ можно рассчитать коэффициенты линейного уравнения регрессии.

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2.$$

Градиент функции отклика в этой точке определяется как

$$\text{grad } y = \frac{\partial y}{\partial x_1} \cdot \overset{P}{i} + \frac{\partial y}{\partial x_2} \cdot \overset{P}{j},$$

где  $\overset{P}{i}, \overset{P}{j}$  — единичные векторы в направлении координатных осей.

Следовательно, для движения по градиенту необходимо изменять факторы пропорционально их коэффициентам регрессии и в сторону, соответствующую знаку коэффициента. В процессе поиска двигаются в этом направлении до тех пор, пока не будет обнаружен локальный максимум (точка  $M_1$  на рис.6.5.4). В точке последнего находят новое направление градиента (направление  $M_1N$ ), осуществляя опять же ПФЭ, и далее процедура повторяется. Стрелками на рис.6.5.4. показано траектория движения к оптимуму.

Практически алгоритм сводится к следующей последовательности операций.

1. Планирование и постановка ПФЭ (или ДФЭ) в окрестности точки начального состояния  $x_{i0}$  (точка  $M_0$  на рис.6.5.4). Расчет коэффициентов  $b_i$  линейной математической модели с целью определения направления градиента.

2. Расчет произведений  $b_i \Delta x_i$ , где  $\Delta x_i$  — интервалы варьирования факторов при ПФЭ (ДФЭ).

3. Выбор базового фактора  $x_i = x_{i0}$ , у которого  $|b_i \Delta x_i| = a = \max$ .

4. Выбор шага крутого восхождения для базового фактора  $h_a$ .

Этот выбор производится на основании имеющейся априорной информации или с учетом опыта исследователя, технологических соображений или других критериев. Относительно выбора шага заметим, что слишком малый шаг потребует значительного числа опытов при движении к оптимуму, а большой шаг создает опасность проскакивания области оптимума.

5. Расчет шагов изменения других факторов по формуле

$$h_i = (b_i \Delta x_i) h_a / a.$$

Это соотношение между величинами шагов изменения отдельных факторов обеспечивает движение по градиенту в факторном пространстве.

6. Составление плана движения по градиенту. Для этого в соответствии с определенными значениями шагов изменения факторов и их последовательным алгебраическим суммированием с основным уровнем в точке

$$x_{ik} = x_{i0} + kh_i, \quad k = 1, 2, \dots$$

находят координаты опытов 5, 6, 7, 8, 9 (см. рис. 6.5.4). Часть этих опытов пролагают "мысленными". "Мысленный" опыт заключается в получении предсказанных (расчетных) значений функции отклика  $\bar{y}$  по линейному уравнению регрессии, что позволяет сократить объем реальных опытов, т.е. увеличить скорость продвижения к экстремуму. При "мысленном" эксперименте перевод координат в кодированную форму и подстановка их в уравнение модели объекта должна подтвердить действительное возрастание  $y$ . Обычно реальные опыты в начале движения из базовой точки вдоль направления градиента ставятся через 2-4 мысленных опытов. Другие опыты реализуют на практике, определяя последовательность значений  $y$  в направлении градиента. Из опытных данных находят положение локального экстремума (точка  $M_1$  на рис. 6.5.4).

7. В окрестности локального экстремума ставят новую серию опытов (ПФЭ или ДФЭ) для определения новых значений коэффициентов уравнения регрессии и нового направления градиента (направление  $M_1N$  на рис. 6.5.4). В дальнейшем процедура повторяется до достижения следующего локального экстремума и т.д. вплоть до определения окрестности координат максимума функции отклика, которая носит название — почти стационарной области.

Признаком достижения этой области является статистическая незначимость коэффициентов  $b_i$ . В почти стационарной области становятся значимы эффекты взаимодействия и квадратичные эффекты. Здесь требуется переходить от ДФЭ (если он использовался ранее) к ПФЭ, а если и этого окажется недостаточно, перейти от планов эксперимента первого порядка к планам второго порядка.

Очевидно, что в задачах, где требуется определить координаты не максимума, а минимума функции отклика, знаки коэффициентов  $b_i$  следует поменять на обратные. В этом случае осуществляется движение по антиградиенту в факторном пространстве.

### 6.5.3. Симплексный метод планирования

Метод симплексного планирования позволяет без предварительного изучения влияния факторов найти область оптимума. В этом методе не требуется вычисления градиента функции отклика, поэтому он относится к безградиентным методам поиска оптимума. Для этого используется специальный план эксперимента в виде симплекса.

Симплекс — это простейший выпуклый многогранник, образованный  $n+1$  вершинами в  $n$ -мерном пространстве, которые соединены между собой прямыми линиями. При этом координаты вершин симплекса являются значениями факторов в отдельных опытах. Так, например, в двухфакторном пространстве (на плоскости)  $n=2$  симплекс — любой треугольник, в трехфакторном (трехмерном) пространстве — тетраэдр и т.д.

Симплекс называется правильным или регулярным, если все расстояния между образующими его вершинами равны (равносторонний треугольник, правильный тетраэдр и др.).

После построения исходного симплекса и проведения опытов при значениях факторов, соответствующих координатам его вершин, анализируют результаты и выбирают вершину симплекса, в которой получено наименьшее (наихудшее) значение функции отклика. Для движения к оптимуму необходимо поставить опыт в новой точке, являющейся зеркальным отображением точки с наихудшим (минимальным) результатом относительно противоположной грани симплекса. На рис.6.5.5 представлено геометрическое изображение симплекс-метода для двумерного случая  $n=2$ . После проведения опытов 1, 2 и 3 худшим оказался опыт 3. Следующий опыт ставится в точке 4, которая образует с точками 1 и 2 новый правильный симплекс. Далее сопоставляются результаты опытов 1, 2 и 4. Наихудший результат получен в точке 1, поэтому она в новом симплексе заменяется зеркальным отображением (точкой 5) и т.д., пока не будет достигнута почти стационарная область. Следует заметить, что хотя этот путь и зигзагообразен, общее число опытов, необходимых для достижения области оптимума, может быть небольшим за счет того, что проводить  $n+1$  опыт приходится лишь в начале, а в дальнейшем каждый шаг сопровождается проведением только одного дополнительного опыта, условия которого выбираются на основе предшествующих результатов.

После изложения основных идей симплексного метода планирования оптимальных экспериментов остановимся на некоторых его деталях.

Выбор размеров симплекса и его начального положения в известной степени произволен. Для построения начального симплекса значения в каждом опыте исходного симплекса определяются по формуле

$$x_{ij} = x_{i0} + C_{ij}\Delta x_i, \quad (6.5.2)$$

где  $x_{i0}$  — координаты центра начального симплекса;  $\Delta x_i$  — интервал варьирования  $i$ -го фактора;  $C_{ij}$  — кодированное значение  $i$ -го фактора для  $j$ -го опыта, выбираемые из числовой матрицы для симплексного планирования, приведенные в табл. 6.5.1.

Таблица 6.5.1

Коэффициенты  $C_{ij}$  для выбора координат симплекса \*

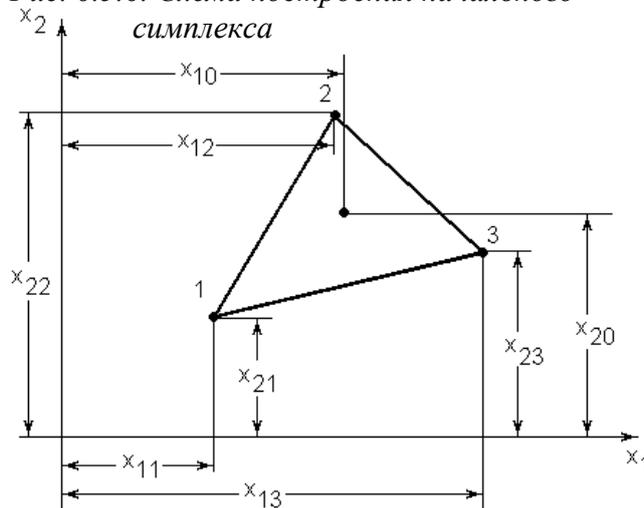
Номер опыта ( $\downarrow j$ )	факторы ( $\rightarrow i$ )					
	$x_1$	$x_2$	$x_3$	...	$x_{n-1}$	$x_n$
1	$k_1$	$k_2$	$k_3$	...	$k_{n-1}$	$k_n$
2	$-R_1$			...		$k_n$
3	0	$-R_2$		...		$k_n$
4	0	0	$-R_3$	...		$k_n$
...	...	...	...	...	...	$k_n$
$n$	0	0	0	0	$R_{n-1}$	$k_n$
$n+1$	0	0	0	0	0	$R_n$

\* Примечание:

$$k_i = \frac{1}{i+1} \sqrt{\frac{i+1}{2i}} = \sqrt{\frac{1}{2i(i+1)}}; \quad R_i = \sqrt{\frac{i}{2(i+1)}}; \quad i = 1, 2, \dots, n;$$

$n$  — число факторов

Рис. 6.5.6. Схема построения начального симплекса



Если, например, необходимо составить симплекс-план для двух факторов, то вначале ставят три опыта с координатами

1-й опыт

$$x_{11} = x_{10} + k_1 \Delta x_1;$$

$$x_{21} = x_{20} + k_2 \Delta x_2.$$

2-й опыт

$$x_{12} = x_{10} - R_1 \Delta x_1;$$

$$x_{22} = x_{20} + k_2 \Delta x_2.$$

3-й опыт

$$x_{13} = x_{10} + 0;$$

$$x_{23} = x_{20} - R_2 \Delta x_2.$$

Симплекс, рассчитанный по этим формулам, представлен на рис.6.5.7. Так, если  $x_{10}=0$  и  $x_{20}=0$ , а  $\Delta x_1=\Delta x_2=1$ , то координаты опытов будут равны: опыт 1 (0,5;0,289), опыт 2 (-0,5; 0,289) и опыт 3 (0;-0,577), что соответствует координатам вершин равностороннего треугольника с длиной стороны, равной 1. Начало координат в этом случае находится в точке пересечения медиан (биссектрис).

Для определения условий проведения опыта в отраженной точке (координат новой вершины симплекса) используется формула

$$x_i' = \frac{2}{n} \sum_{j=1}^n x_{ij} - x_{i'k}, \quad i \neq j, \quad (6.5.3)$$

где  $x_{i'n}$  — координата новой точки (новой вершины) симплекса для  $i$ -й переменной;  $x_{i'k}$  — координата заменяемой точки (координата вершины симплекса с наихудшим откликом перед ее отбрасыванием);

$\frac{1}{n} \sum_{j=1}^{n+1} x_{ij}$  — среднее значение из координат всех вершин симплекса кроме заменяемой.

Известны следующие критерии окончания процесса последовательного отражения наихудших вершин и постановки очередных опытов в новых вершинах:

1. Разность значений функции отклика в вершинах симплекса становится меньше наперед заданной величины. Это означает либо выход в "почти стационарную" область вблизи оптимума, либо достижение участка поверхности  $y = f(x_1; \dots; x_n) = \text{const}$  в виде "плато". В этом случае дополнительными опытами в стороне от симплекса следует удостовериться в отсутствии других участков с более существенной кривизной поверхности  $y = f(x_1; \dots; x_n)$  и принять величину с экстремальным значением функции отклика за точку оптимума.

2. Отражение любой из вершин симплекса после однократного качания приводит к его возврату в прежнее положение. При этом есть основания утверждать "накрытие" симплекса точки оптимума.

3. Циклическое движение симплекса вокруг одной из его вершин на протяжении более чем нескольких шагов. Подобная ситуация имеет место, когда искомый оптимум располагается внутри области, охватываемой циркулирующим симплексом.

В случаях 2 и 3 рекомендуется уменьшить размеры симплекса, т.е. расстояния между вершинами, и продолжить поиск до желаемого уточнения координат искомого оптимума.

Изложенный алгоритм симплексного метода сравнительно прост, он достаточно эффективен даже при незначительных ошибках в определении функции отклика, однако работает недостаточно

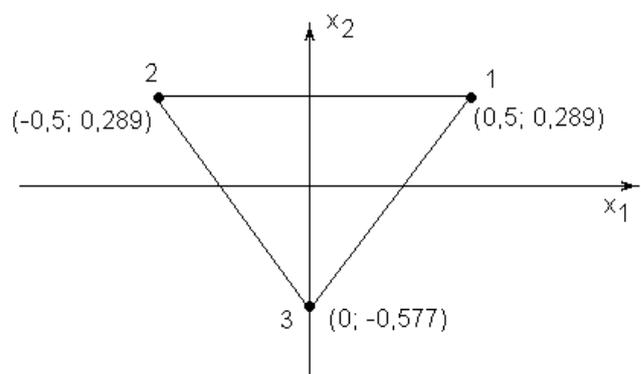


Рис.6.5.7. Координаты вершин симплекса при  $x_{i0}=0$  и  $\Delta x_i=1$  при  $n=2$

быстро. Существует его модификация, известная под названием "метод деформируемого симплекса", которая ускоряет процесс поиска оптимума за счет использования на данном шаге информации, накопленной на предыдущих шагах.

Сущность метода поиска по деформированному симплексу заключается в том, что при отражении наихудшей вершины относительно центра тяжести противоположной грани размер симплекса не остается постоянным, а осуществляется его деформация (растяжение или сжатие). Для пояснения существа метода введем координату центра тяжести  $\bar{x}_i$  остальных (за исключением наихудшей) вершин симплекса:

$$\bar{x}_i = \frac{\sum_{j=1}^{n+1} x_{ij}}{n}; \quad j \neq j_3.$$

Тогда формула (6.5.3) может быть преобразована к виду

$$x_{iH} = 2\bar{x}_i - x_{j_3}$$

или

$$\tilde{x}_{iH} = \bar{x}_i + \alpha(\bar{x}_i - x_{j_3}).$$

При  $\alpha=1$  получим выражение (6.5.3) и  $\tilde{x}_{iH} = x_{iH}$ .

Введем обозначения:

$y_3$  — наихудший (минимальный отклик) в симплексе;

$u_{max}$  — наилучший (максимальный) отклик;

$y_3'$  — отклик, следующий за наихудшим.

Следовательно  $y_3 < u_{max} < y_3'$ .

В зависимости от значения функции отклика в точке нормального отражения  $u_H$  при  $\alpha=1$  возможны следующие варианты.

1. Если  $y_3 < u_H < u_{max}$ , т.е.  $x_{iH}$  будет не худшей и не лучшей точкой в новом наборе точек, то  $x_{j_3}$  следует заменить на  $x_{iH}$  с  $\alpha=1$ . В этом случае осуществляется нормальное отражение.

2. Если  $u_H > u_{max}$ , то  $x_{iH}$  оказывается новой лучшей точкой в новом наборе точек. В этом случае направление растяжения признается "весьма удачным" и растягивается (по отношению к нормальному) симплекс в этом направлении. Для этого случая  $1 < \alpha < 2$  и  $\alpha$  называется коэффициентом растяжения.

3. Если  $y_3 < u_H < y_3'$ , то направление отражения признается правильным, но симплекс слишком велик и его следует сжать выбором коэффициента сжатия  $\alpha$  из диапазона  $0 < \alpha < 1$ .

4. Если  $u_H < y_3$ , то даже направление отражения выбрано неверно и следует осуществить отрицательное сжатие выбором отрицательного значения коэффициента  $\alpha$  ( $-1 < \alpha < 0$ ).

Таким образом, на каждом шаге следует вначале нормально отразить наихудшую вершину симплекса ( $\alpha=1$ ), поставить в этой точке опыт, определить  $u_H$ , а затем поставить следующий опыт в точке факторного пространства  $\tilde{x}_{iH}$ , координаты которой определяются по формуле (6.5.4) с учетом рассмотренных вариантов 1-4.

На рис.6.5.8 показаны точка 4 очередного опыта при нормальном отражении ( $\alpha=1$ ) наихудшей вершины 1, точки 5', 5'', 5''' последующих опытов для случаев соответственно растяжения ( $\alpha=2$ ), сжатия ( $\alpha=0,5$ ) и отрицательного сжатия ( $\alpha=-0,5$ )

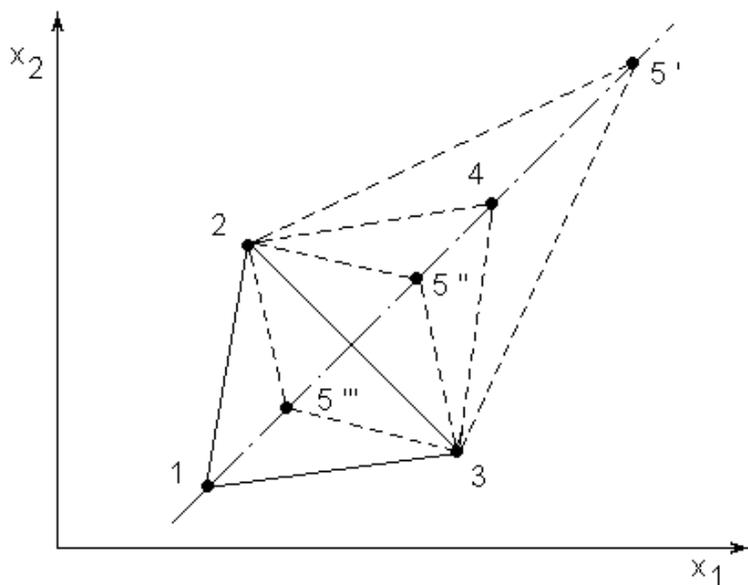


Рис.6.5.8. К методу деформированного симплекса

симплекса.

Таким образом, метод поиска по деформированному симплексу обладает повышенной гибкостью, позволяющей учитывать особенности поверхности отклика.

Вопросы к теме 6.

1. Основные определения и понятия
2. Пример хорошего и плохого эксперимента
3. Планирование первого порядка
4. Выбор основных факторов и их уровней
5. Планирование эксперимента
6. Определение коэффициентов уравнения регрессии
7. Статистический анализ результатов эксперимента
8. Дробный факторный эксперимент
9. Планы второго порядка
10. Ортогональные планы второго порядка
11. Ротатабельные планы второго порядка
12. Планирование экспериментов при поиске оптимальных условий
13. Метод покоординатной оптимизации (Гаусса - Зейделя)
14. Метод крутого восхождения (Бокса-Уилсона)
15. Симплексный метод планирования

Для заметок

Учебное издание

**КОНСПЕКТ ЛЕКЦИЙ**  
по дисциплине  
**«ТЕОРИЯ И ПРАКТИКА ЭКСПЕРИМЕНТА»**  
для магистров направления подготовки  
44.04.04 Профессиональное обучение (по отраслям)

**С о с т а в и т е л ь:**  
Сергей Николаевич Акинъшин

Печатается в авторской редакции.  
Компьютерная верстка и оригинал-макет автора.

Подписано в печать \_\_\_\_\_  
Формат 60x84<sup>1</sup>/<sub>16</sub>. Бумага типограф. Гарнитура Times  
Печать офсетная. Усл. печ. л. \_\_\_\_\_. Уч.-изд. л. \_\_\_\_\_  
Тираж 100 экз. Изд. № \_\_\_\_\_. Заказ № \_\_\_\_\_. Цена договорная.

Издательство Луганского государственного  
университета имени Владимира Даля

*Свидетельство о государственной регистрации издательства  
МИ-СРГ ИД 000003 от 20 ноября 2015г.*

**Адрес издательства:** 91034, г. Луганск, кв. Молодежный, 20а  
**Телефон:** 8 (0642) 41-34-12, **факс:** 8 (0642) 41-31-60  
**E-mail:** [uni@snu.edu.ua](mailto:uni@snu.edu.ua) **http:** [www.snu.edu.ua](http://www.snu.edu.ua)