

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО
ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
Механико-математический факультет

Д. П. КАСЫМОВ

**ПЛАНИРОВАНИЕ И
ОБРАБОТКА РЕЗУЛЬТАТОВ
ЭКСПЕРИМЕНТА**

Учебно-методическое пособие

Томск
Издательство Томского государственного университета
2021

РАССМОТРЕНО И РЕКОМЕНДОВАНО методической комиссией
Механико-математического факультета ТГУ
Протокол № 7 от «15» ноября 2021 года

Председатель комиссии: доцент кафедры теоретической
механики, канд. физ.-мат. наук Е.А. Тарасов

Касымов Д.П.

Планирование и обработка результатов эксперимента : учебно-методическое пособие. – Томск: Издательство Томского государственного университета, 2021. – 93 с.

В учебно-методическом пособии излагается современный курс планирования эксперимента. Освещаются основные понятия и определения в области измерений и измерительной техники, даются краткие сведения из теории вероятностей и математической статистики. Особое внимание уделено принципам организации предварительной обработки экспериментальных исследований.

Пособие необходимо для изучения теоретической и практической частей следующих учебных курсов: «Планирование эксперимента», «Лабораторные работы по гидромеханике», «Механика в современном естествознании».

Каждый раздел данного пособия содержит контрольные вопросы для закрепления и проверки знаний по изучаемым темам.

Пособие разработано для студентов направления «Механика и математическое моделирование» механико-математического факультета, а также для студентов других специальностей, занимающихся теоретико-экспериментальными исследованиями и обработкой результатов эксперимента.

СОДЕРЖАНИЕ

ПРЕДИСЛОВИЕ.....	5
1. ОСНОВНЫЕ ПОНЯТИЯ И ОПРЕДЕЛЕНИЯ В ОБЛАСТИ ИЗМЕРЕНИЙ И ИЗМЕРИТЕЛЬНОЙ ТЕХНИКИ.....	6
1.1. Основные понятия.....	6
1.2. Классификация измерений.....	10
1.3. Основные этапы измерений.....	12
1.4. Основные методы измерений.....	14
Контрольные вопросы.....	15
2. ЭКСПЕРИМЕНТ КАК ПРЕДМЕТ ИССЛЕДОВАНИЯ.....	16
2.1. Классификация видов экспериментальных исследований.....	16
2.2. Погрешности результатов исследований.....	17
Контрольные вопросы.....	19
3. КРАТКИЕ СВЕДЕНИЯ ИЗ ТЕОРИИ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ.....	20
3.1. Вероятность случайных событий, их характеристики.....	20
3.2. Нормальный закон распределения.....	24
Контрольные вопросы.....	28
4. ПРЕДВАРИТЕЛЬНАЯ ОБРАБОТКА ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ.....	29
4.1. Вычисление характеристик эмпирических распределений.....	29
4.2. Статистические гипотезы.....	31
4.3. Отсев грубых погрешностей.....	34
4.4. Определение доверительных интервалов для исследуемых величин.....	37
4.4.1. Оценка доверительного интервала для математического ожидания.....	38
4.4.2. Оценка доверительного интервала для дисперсии.....	40
Контрольные вопросы.....	49
5. АНАЛИЗ РЕЗУЛЬТАТОВ ПАССИВНОГО ЭКСПЕРИМЕНТА. ЭМПИРИЧЕСКИЕ ЗАВИСИМОСТИ.....	50
5.1. Характеристика видов связей между рядами наблюдений.....	50
5.2. Определение коэффициентов уравнения регрессии.....	52
5.3. Регрессионный анализ.....	55
5.3.1. Проверка адекватности модели.....	56
5.3.2. Проверка значимости коэффициентов уравнения регрессии.....	57
Контрольные вопросы.....	58
6. ОЦЕНКА ПОГРЕШНОСТЕЙ РЕЗУЛЬТАТОВ НАБЛЮДЕНИЙ.....	59
6.1. Оценка погрешностей определения величин функций.....	59
6.2. Обратная задача теории экспериментальных погрешностей.....	61
6.3. Определение наиболее выгодных условий эксперимента.....	62
Контрольные вопросы.....	63

7. МЕТОДА ПЛАНИРОВАНИЯ ЭКСПЕРИМЕНТОВ. ЛОГИЧЕСКИЕ ОСНОВЫ.....	64
7.1. Основные определения и понятия	64
7.2. Пример хорошего и плохого эксперимента.....	65
7.3. Планирование первого порядка.....	67
7.3.1. Выбор основных факторов и их уровней	67
7.3.2. Планирование эксперимента	68
7.3.3. Определение коэффициентов уравнения регрессии	70
7.3.4. Статистический анализ результатов эксперимента.....	71
7.3.5. Дробный факторный эксперимент	72
7.3.6. Разработка математической модели гидравлического режима методической печи.....	76
7.4. Планы второго порядка	79
7.4.1. Ортогональные планы второго порядка.....	80
7.4.2. Ротатабельные планы второго порядка	83
7.5. Планирование экспериментов при поиске оптимальных условий ...	85
7.5.1. Метод покоординатной оптимизации (Гаусса — Зейделя) .	87
7.5.2. Метод крутого восхождения (Бокса-Уилсона)	87
7.5.3. Симплексный метод планирования	89
Контрольные вопросы.	91
ОСНОВНАЯ ЛИТЕРАТУРА	92
ДОПОЛНИТЕЛЬНАЯ ЛИТЕРАТУРА.....	92
ЭЛЕКТРОННЫЕ РЕСУРСЫ.....	92
СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ	93

ПРЕДИСЛОВИЕ

Основной целью эксперимента являются выявление свойств исследуемых объектов, проверка справедливости гипотез и, на этой основе, широкое и глубокое изучение темы научного исследования. На начальной стадии эксперимента, наблюдая за поведением объекта или протеканием явления, исследователь делает предположения о наличии некоторых взаимосвязей и закономерностей. В заключительной стадии формируется цель исследования, определяются величины – факторы, влияющие на свойства объекта и вид их взаимосвязи, выдвигается гипотеза о модели исследуемого объекта. В соответствии с видом математической модели строится план эксперимента. От правильного выбора плана проведения целенаправленного эксперимента зависит успех дальнейших исследований – правильно выбранный план позволяет не только уменьшить объем исследований, но и минимизировать влияние на результат исследования неучтенных, неконтролируемых факторов. Принятие проектных решений в машиностроении и оценка их качества в основном осуществляются на основании данных эксперимента, а усложнение объектов испытания вызывает резкое повышение стоимости их исследования.

Поэтому задача извлечения наибольшего объема информации об изучаемых процессах или устройствах при ограничениях по затратам является достаточно актуальной. Решению указанной проблемы способствует широкое использование в прикладных исследованиях статистических методов планирования экспериментов, которые не только дают способ обработки экспериментальных данных, но и позволяют оптимально организовать эксперимент. Планирование эксперимента и математическая обработка его результатов все больше входят в круг вопросов, необходимых студентам старших курсов, аспирантам и инженерам. Усвоение и использование методов планирования эксперимента позволяют повысить эффективность принимаемых решений.

В настоящем пособии освещаются основные понятия и определения в области измерений и измерительной техники, даются краткие сведения из теории вероятностей и математической статистики. Особое внимание уделено принципам организации предварительной обработки экспериментальных исследований.

1. ОСНОВНЫЕ ПОНЯТИЯ И ОПРЕДЕЛЕНИЯ В ОБЛАСТИ ИЗМЕРЕНИЙ И ИЗМЕРИТЕЛЬНОЙ ТЕХНИКИ

1.1. Основные понятия

К общепринятым в метрологии определениям относятся понятия: *измерения, средства, принцип, метод и объект измерения, алгоритм измерения и шкалы измерений*, а также ряд других терминов.

Измерением называется процесс нахождения значения физической величины опытным путем с помощью специальных технических средств.

Зачастую информация об объекте измерения известна до проведения исследований, что является важнейшим фактором, обуславливающим эффективность измерения. Такую информацию об объекте измерения называют *априорной информацией*. При полном отсутствии этой информации измерение в принципе невозможно, так как неизвестно, что же необходимо измерить, а, следовательно, нельзя выбрать нужные средства измерений. При наличии априорной информации об объекте в полном объеме, т.е. при известном значении измеряемой величины, измерения попросту не нужны. Априорная информация определяет достижимую точность измерений и их эффективность.

Информация, получаемая в результате измерения, может содержаться в объекте измерения в двух формах: пассивной и активной.

Пассивная информация – это совокупность сведений, заключенных в том, как устроен объект; такой информацией является, например, информация о величине напряжения источника питания. С другой стороны, информация является *активной*, если она имеет форму энергетической характеристики какого-либо явления. Подобные энергетические явления называются сигналами. Их примерами являются электрические, оптические и акустические сигналы, используемые для передачи информации.

Имеются и другие определения, которые рассматривают измерения как процесс получения информации, заключающийся в сравнении опытным путем измеряемых и известных величин или сигналов, и представления ее в числовой форме.

Итак, измерение представляет собой специфический информационный процесс, результатом которого является получение количественной информации об измеряемых величинах – измерительной информации.

Метрологическая суть измерения сводится к основному уравнению измерения (основному уравнению метрологии):

$$A = k * A_0, \quad (1.1.1)$$

где A – значение измеряемой физической величины, A_0 – значение величины, принятой за образец, k – отношение измеряемой величины к образцу.

Наиболее удобен вид основного уравнения метрологии (1.1.1), если выбранная за образец величина равна единице. При этом параметр k представляет собой числовое значение измеренной величины, зависящее от принятого метода измерения и единицы измерения.

Основными характеристиками измерений являются результат, погрешность, точность, правильность, сходимость, воспроизводимость, достоверность.

Результат измерений физической величины – это значение физической величины, полученное путем ее измерений.

Часто в полученный результат вносят поправки, что находит отражение в терминологии:

1. Неисправленный результат измерения – значение физической величины, полученное при помощи средств измерений до внесения поправок;

2. Исправленный результат измерения – значение физической величины, полученное при помощи средств измерений и уточненное путем внесения в него необходимых поправок.

Погрешность измерения – отклонение результат измерения от истинного значения измеряемой величины.

Точность измерений – понятие, отражающее меру близости результатов измерений к истинному значению измеряемой физической величины. Точность и погрешность связаны обратной зависимостью.

Правильность измерений – это метрологическая характеристика, отражающая близость к нулю систематических погрешностей результатов измерений.

Сходимость результатов измерений характеризует качество измерений, отражающее близость друг к другу результатов измерений одной и той же величины, выполняемых повторно одними и теми же методами и средствами измерений и в одних и тех же условиях.

Воспроизводимость результатов измерений – характеристика качества измерений физической величины, отражающая близость друг к другу результатов измерений одной и той же величины, полученных в разных местах, разными методами и средствами измерений, разными операторами, но приведенных к одним и тем же условиям.

Достоверность измерений определяется степенью доверия к результату измерения и характеризуется вероятностью того, что истинное значение измеряемой величины находится в указанных пределах или в указанном интервале. Данный интервал в теории измерений называют доверительным, между его границами с заданной доверительной вероятностью находится истинное значение оцениваемого параметра.

Принцип измерений – совокупность физических явлений, на которых основаны измерения.

Метод измерений – совокупность приемов использования принципов и средств измерений. Это достаточно общее определение на практике часто конкретизируют, относя его только к применяемым средствам измерения, например метод измерения частоты частотомером, напряжения – вольтметром, силы тока – амперметром и т.д.

Понятие «метод измерения» следует отличать от методики измерения – общего или поэтапного плана проведения измерения – намеченного распорядка измерений, определяющего состав применяемых приборов, последовательность и правила проведения операций.

Объект измерения – это реальный физический объект, свойства которого характеризуются одной или несколькими измеряемыми физическими величинами. Он обладает многими свойствами и находится в сложных и многосторонних связях с другими объектами. Поэтому в теоретической метрологии введено понятие математической модели объекта.

Математическая модель объекта – совокупность математических символов (образов) и отношений между ними, которая адекватно описывает свойства объекта измерения.

В технической литературе и нормативной документации часто встречается термин алгоритм измерения, под которым следует понимать точное предписание о порядке выполнения операций, обеспечивающих измерение искомого значения физической величины.

На практике необходимо проводить измерения различных величин, характеризующих свойства веществ, тел, явлений и процессов. Некоторые свойства

проявляются только количественно, другие - качественно. Количественные или качественные проявления любого свойства отражаются множествами, которые образуют шкалы измерения.

Шкала измерений количественного свойства является шкалой физической величины.

Шкала физической величины – это упорядоченная последовательность значений физической величины, принятая по соглашению на основании результатов точных измерений.

Различают следующие основные типы:

1. Шкала наименований (шкала классификации). Она основана на приписывании объекту цифр (знаков), играющих роль простых имен: это приписывание служит для нумерации предметов только с целью их идентификации или для нумерации классов, причем, такой нумерации, что каждому из элементов соответствующего класса приписывается одна и та же цифра. Такое приписывание цифр выполняет на практике ту же функцию, что и наименование. Поэтому с цифрами, используемыми только как специфические имена, нельзя производить никаких арифметических действий, так, например, если один из резисторов обозначен в схеме R_6 , а другой R_{18} , то из этого нельзя сделать заключение, что значения их сопротивления отличаются втрое, а можно лишь установить, что оба они относятся к классу резисторов.

Примером такой шкалы может служить атлас цветов (шкала цветов). Процесс измерения заключается в визуальном сравнении окрашенного предмета с образцами цветов (эталонными образцами атласа цветов). Поскольку каждый цвет имеет немало вариантов, такое сравнение под силу опытному эксперту, который обладает не только практическим опытом, но и соответствующими особыми характеристиками зрительных возможностей.

2. Шкала порядка (шкала рангов) характеризует значение измеряемой величины в баллах (шкала землетрясений, силы ветра, твердости физических тел и т.п.).

Эта шкала предполагает упорядочение объектов относительно какого-то определенного их свойства, т.е. расположение их в порядке убывания или возрастания данного свойства. Полученный при этом упорядоченный ряд называют ранжированным рядом, а саму процедуру - ранжированием.

По шкале порядка сравниваются между собой однородные объекты, у которых значения интересующих свойств неизвестны. Поэтому ранжированный ряд может дать ответ на вопросы типа "что больше (меньше)" или, "что лучше (хуже)". Более подробную информацию - на сколько больше или меньше, во сколько раз лучше или хуже - шкала порядка дать не может.

Результаты оценивания по шкале порядка также не могут подвергаться никаким арифметическим действиям. Однако небольшое усовершенствование шкалы порядка позволило применить ее для числового оценивания величин в тех случаях, когда отсутствует единица величины. Для этого, расположив объекты в порядке возрастания (убывания) того или иного свойства, некоторые точки ранжированного ряда фиксируют в качестве отправных (реперных) точек. Совокупность реперных точек образует некую "лестницу" - шкалу возможных проявлений соответствующего свойства. Реперным точкам могут быть поставлены в соответствие цифры, называемые баллами, и, таким образом, появляется возможность оценивания, "измерения" данного свойства в баллах, по натуральной шкале. Так, для измерения скорости ветра в 1805 г. Бофортом была предложена натуральная шкала скорости ветра в "баллах Бофорта" (таблица 1.1.1).

Шкала для визуальной оценки силы ветра

Баллы	Наименование	Действие	Скорость, м/с
0	Штиль	Дым идет вертикально	0-0,9
1	Тихий	Дым идет слегка наклонно	0,9-2,4
2	Легкий	Ощущается лицом, шелестят листья	2,4-4,4
...
11	Жестокий шторм	Большие разрушения	30,5-34,8
12	Ураган	Опустошительное действие	34,8-39,2

С развитием методов и средств измерения физических величин условным баллам натуральной шкалы ставятся в соответствие числовые значения в принятых для данной величины единицах. Так, шкала Бофорта использовалась до 1964 г., когда международным соглашением был принят ее перевод в скорость ветра, измеряемую в метрах в секунду.

По натуральным шкалам до сих пор оценивают интенсивность землетрясений, морское волнение, твердость минералов и некоторые другие величины.

Основным недостатком натуральных шкал является полное отсутствие уверенности в том, что интервалы между выбранными реперными точками являются равновеликими, а следовательно, в такой шкале невозможно вычленить единицу и оценить погрешность полученной оценки.

3. Шкала интервалов (шкала разностей) является дальнейшим развитием шкал порядка.

Шкала интервалов состоит из одинаковых интервалов, имеет единицу измерения и произвольно выбранное начало - нулевую точку. К таким шкалам относится летоисчисление по различным календарям, в которых за начало отсчета принято либо сотворение мира, либо рождество Христово и т. д. Температурные шкалы Цельсия, Фаренгейта и Реомюра также являются шкалами интервалов.

На наиболее привычной для нас температурной шкале Цельсия за начало отсчета разности температур принята температура таяния льда. С ней сравниваются все другие температуры. Для удобства пользования шкалой Цельсия интервал между температурой таяния льда и температурой кипения воды разделен на 100 равных интервалов - градусов. Шкала Цельсия распространяется как в сторону отрицательных, так и положительных интервалов. Когда говорят, что температура воздуха равна 20°C, это означает, что она на 20 градусов выше температуры, принятой за нулевую отметку шкалы (выше нуля).

На температурной шкале Фаренгейта тот же интервал разбит на 180 градусов. Следовательно, градус Фаренгейта по размеру меньше, чем градус Цельсия. Кроме того, начало отсчета интервалов по шкале Фаренгейта сдвинуто на 32 градуса в сторону низких температур.

Деление шкалы интервалов на равные части - градации - устанавливает единицу физической величины, что позволяет не только выразить результат измерения в числовой мере, но и оценить погрешность измерения.

Результаты измерений по шкале интервалов можно складывать друг с другом и вычитать друг из друга, т.е. определять, на сколько одно значение физической величины больше или меньше другого. Определить по шкале интервалов, во сколько раз одно значение величины больше или меньше другого, невозможно, поскольку на шкале не определено начало отсчета физической величины. Но в то же время это может

быть сделано в отношении интервалов (разностей). Так, разность температур 25 градусов в 5 раз больше разности температур 5 градусов.

4. Шкала отношений является наиболее совершенной шкалой. Она отличается от шкалы интервалов тем, что в ней строго определено положение нулевой точки, благодаря чему шкала отношений не накладывает никаких ограничений на математический аппарат, используемый для обработки результатов наблюдений. При использовании шкалы отношений измерение какой-либо величины сводится к экспериментальному определению отношения этой величины к другой подобной, принятой за единицу. Шкала отношений охватывает интервал значений n от нуля до бесконечности и не содержит отрицательных значений.

По шкале отношений измеряют расстояние, массу, скорость и т.п. Другим примером шкалы отношений может служить температурная шкала Кельвина. В ней за начало отсчета принят абсолютный нуль температуры, при котором прекращается тепловое движение молекул, более низкой температуры быть не может. Второй реперной точкой служит температура таяния льда.

1.2. Классификация измерений

Измерения как экспериментальные процедуры определения значений измеряемых величин весьма разнообразны, что объясняется множеством измеряемых величин, различным характером их изменения во времени, различными требованиями к точности измерений т.д. В связи с этим измерения классифицируют по различным признакам.

Наибольшее распространение получила классификация по способу нахождения искомого значения измеряемой величины, согласно которой различают прямые, косвенные, совместные и совокупные измерения.

Прямым называется измерение, при котором искомое значение физической величины находят непосредственно по показаниям средства измерений. Например, измерение напряжения и силы тока соответственно - вольтметром и амперметром.

Косвенным называется измерение, при котором искомое значение величины находят на основании известной зависимости между ней и величинами, подвергаемыми прямым измерениям, которые проводились в одинаковых условиях. Так, если измерить силу тока амперметром, а напряжение вольтметром, то по известной функциональной взаимосвязи всех трех величин можно рассчитать мощность электрического тока.

Совокупные измерения осуществляются путем одновременного измерения нескольких одноименных величин, при которых искомое значение находят решением системы уравнений, получаемых в результате прямых измерений различных сочетаний этих величин.

Например, измеряя сопротивления R_{ab}, R_{ac}, R_{bc} между вершинами треугольника, в котором соединены сопротивления R_1, R_2, R_3 , можно определить искомые значения сопротивлений R_1, R_2, R_3 методом совокупных измерений:

$$R_{ab} = \frac{R_1(R_2 + R_3)}{R_1 + R_2 + R_3}, R_{ac} = \frac{R_2(R_1 + R_3)}{R_1 + R_2 + R_3}, R_{bc} = \frac{R_3(R_1 + R_2)}{R_1 + R_2 + R_3}. \quad (1.2.1)$$

Совместными измерениями называются проводимые одновременно измерения двух или нескольких неоднородных величин для установления зависимости между ними. Примером совместных измерений является ряд одновременных, прямых измерений электрического сопротивления проводника и его температуры для установления зависимости сопротивления от температуры.

Как видно из приведенных определений, эти два вида измерений весьма близки друг к другу. В обоих случаях искомые значения находятся в результате решения системы уравнений, коэффициенты в которых получены путем прямых измерений. Отличие состоит в том, что при совместных измерениях одновременно определяются несколько одноименных величин, а при совокупных - разноименных.

По характеристике точности измерения делятся на равноточные и неравноточные.

Равноточные измерения – ряд измерений какой-либо величины, выполненных одинаковыми по точности средствами измерений и в одних и тех же условиях.

Неравноточные измерения – ряд измерений какой-либо величины, выполненных несколькими различными по точности средствами измерений и (или) в различных условиях.

В зависимости от числа измерений, проводимых во время эксперимента, различают однократные и многократные измерения.

Однократное измерение – измерение, выполненное один раз.

Многократное измерение – измерение одного и того же размера физической величины, результат которого получен из нескольких следующих друг за другом измерений, т.е. состоящее из ряда однократных измерений.

По режиму работы средства измерения различают статические и динамические измерения. Любое средство измерений обладает инерцией (механической, тепловой, электрической) и, следовательно, не может мгновенно реагировать на изменение измеряемой величины. Поэтому при измерении переменной физической величины инерция средства измерения приведет к некоторому отставанию показаний средства измерений от истинного значения величины в каждый момент времени. Очевидно, что это отставание будет зависеть не только от инерционных (динамических) свойств средств измерений, но и от скорости изменения самой измеряемой величины. В том случае, когда показания средства измерения не зависят от его динамических свойств, или когда этой зависимостью можно пренебречь, говорят, что средство измерения работает в статическом режиме, а само измерение называют статическим. В противном случае измерение относят к динамическим.

В зависимости от метрологического назначения измерения делятся на технические и метрологические.

Технические измерения – измерения с помощью рабочих средств измерений.

Метрологические измерения - измерения при помощи эталонов и образцовых средств измерений с целью воспроизведения единиц физической величины для передачи их размера рабочим средствам измерений.

При метрологических измерениях в обязательном порядке учитываются погрешности, а при технических – принимается наперед заданная погрешность, достаточная для решения данной практической задачи. Поэтому при технических измерениях нет необходимости определять и анализировать погрешности получаемых результатов. Технические измерения являются наиболее массовым видом.

В зависимости от выражения результатов измерений последние подразделяются на абсолютные и относительные.

Абсолютное измерение – измерение, основанное на прямых измерениях одной или нескольких основных величин и (или) использовании значений физических констант. Так, в известной формуле Эйнштейна $E = mc^2$ масса (m) – основная физическая величина, которая может быть измерена прямым путем (взвешиванием), а скорость света (c) – физическая константа.

Относительное измерение - измерение отношения определяемой величины к одноименной величине, играющей роль единицы, или измерения величины по отношению к одноименной величине, принимаемой за исходную.

Характерные примеры относительных измерений: измерение отношения напряжений или мощностей, исследование различных частотных характеристик электрических цепей и т.д.

При относительных измерениях широко используется внесистемная безразмерная единица - децибел (дБ), определяемая при сравнении напряжений по формуле:

$$1\text{дБ} = 201g(U_2/U_1), \text{ при } U_2/U_1 = 10^{1/20} = 1,122,$$

а при сравнении мощностей P_2 и P_1 :

$$1\text{дБ} = 101g(P_2/P_1), \text{ при } P_2/P_1 = 10^{1/10} = 1,259.$$

Для перевода отношений мощностей и напряжений (токов) в децибелы и обратно применяют специальные таблицы, приведенные в справочниках.

По способу преобразования измеряемой величины и форме представления результата измерения делятся на аналоговые (непрерывные) и цифровые (дискретные).

При аналоговых измерениях измерительный прибор производит непрерывное преобразование измеряемой величины, результатом которого является перемещение указателя относительно шкалы. Заключение о численном значении величины делает оператор, отмечая положение указателя относительно отметок шкалы измерительного прибора. Точность измерения при этом ограничивается геометрическими особенностями указателя и шкалы и обычно не превышает 0,05 %.

При цифровых измерениях сравнение физической величины с рядом образцовых значений производится в измерительном приборе автоматически, оператор же получает численное значение измеренной величины в цифровой форме. Естественно, что здесь все зависит от точности сравнения в измерительном приборе и, к тому же, исключаются субъективные ошибки оператора. Современные цифровые приборы, как правило, обеспечивают более высокую точность, чем аналоговые. Роль оператора упрощается, так как он лишь считывает число.

1.3. Основные этапы измерений

Измерение - последовательность сложных и разнородных действий, состоящая из ряда этапов.

Первым этапом любого измерения является постановка измерительной задачи. Он включает в себя:

1. Сбор данных об условиях измерения и исследуемой физической величине, т.е. накопление априорной информации об объекте измерения с последующим ее анализом;

2. Формирование модели объекта и определение измеряемой величины, что является наиболее важным, особенно при решении сложных измерительных задач. Измеряемая величина определяется с помощью принятой модели как ее параметр или характеристика. В простых случаях, т.е. при измерениях невысокой точности, модель объекта в явном виде не выделяется, а пороговое несоответствие пренебрежимо мало;

3. Постановку измерительной задачи на основе принятой модели объекта измерения;

4. Выбор конкретных величин, посредством которых будет находиться значение измеряемой величины;

5. Формулирование уравнения измерения.

Вторым этапом процесса измерения является планирование измерения. В общем случае оно выполняется в следующей последовательности:

1. Выбор методов измерений непосредственно измеряемых величин и возможных типов средств измерений;

2. Априорная оценка погрешности измерения;

3. Определение требований к метрологическим характеристикам средств измерений и условиям измерений;

4. Выбор средств измерений в соответствии с указанными требованиями; выбор параметров измерительной процедуры (числа наблюдений для каждой измеряемой величины, моментов времени и точек выполнения наблюдений);

5. Подготовка средств измерений к выполнению экспериментальных операций;

6. Обеспечение требуемых условий измерений или создание возможности их контроля.

Эти первые два этапа, являющиеся подготовкой к измерениям, имеют принципиальную важность, поскольку определяют конкретное содержание следующих этапов измерения. Подготовка проводится на основе априорной информации. Качество подготовки зависит от того, в какой мере она была использована. Эффективная подготовка является необходимым, но недостаточным условием достижения цели измерения. Ошибки, допущенные при подготовке измерений, с трудом обнаруживаются и корректируются на последующих этапах.

Третий, главный этап измерения – измерительный эксперимент. В узком смысле он является отдельным измерением. В общем случае последовательность действий во время этого этапа следующая:

1. Взаимодействие средств и объекта измерений;
2. Преобразование сигнала измерительной информации;
3. Воспроизведение сигнала заданного размера;
4. Сравнение сигналов и регистрация результата.

Последний этап измерения – обработка экспериментальных данных. В общем случае она осуществляется в последовательности, которая отражает логику решения измерительной задачи:

1. Предварительный анализ информации, полученной на предыдущих этапах измерения;
2. Вычисление и внесение возможных поправок на систематические погрешности;
3. Анализ возможных алгоритмов обработки и выбор одного из них на основании известных свойств алгоритмов, априорных данных и предварительного анализа экспериментальных данных;
4. Проведение вычислений согласно принятому алгоритму, в итоге которых получают значения измеряемой величины и погрешностей измерений; анализ и интерпретация полученных результатов;
5. Запись результата измерений и показателей погрешности в соответствии с установленной формой представления.

Некоторые пункты данной последовательности могут отсутствовать при реализации конкретной процедуры обработки результатов измерений.

Задача обработки данных подчинена цели измерения и после выбора средства измерений однозначно вытекает из измерительной задачи и, следовательно, является вторичной.

Перечисленные выше этапы существенно различаются по выполняемым операциям и их трудоемкости. В конкретных случаях соотношение и значимость каждого из этапов заметно варьирует. Для многих технических измерений вся процедура измерения сводится к экспериментальному этапу, поскольку анализ и планирование, включая априорное оценивание погрешности, выбор нужных методов и средств измерений осуществляются предварительно, а обработка данных измерений, как правило, минимизируется.

Выделение этапов измерения имеет непосредственное практическое значение, а именно способствует своевременному осознанному выполнению всех действий и оптимальной реализации измерений. Это, в свою очередь, позволяет избежать

серьезных методических ошибок, связанных с переносом проблем одного этапа на другой.

1.4. Основные методы измерений

Метод измерений – совокупность приемов использования принципов и средств измерений.

Современные методы измерений принято делить на метод непосредственной оценки и метод сравнения (рис. 1.4.1).



Рис. 1.4.1. Классификация методов измерения

При методе непосредственной оценки численное значение измеряемой величины определяется непосредственно по показанию измерительного прибора (например, измерение напряжения с помощью вольтметра).

Метод сравнения – метод измерений, при котором измеряемую величину сравнивают с величиной, воспроизводимой мерой. Следовательно, отличительной особенностью методов сравнения является непосредственное участие меры в процессе измерения.

Различают 4 разновидности метода сравнений: дифференциальный, нулевой, метод замещения и метод совпадений.

При дифференциальном методе (рис. 1.4.2, б) полное уравнивание не производят, а разность между измеряемой величиной и величиной, воспроизводимой мерой, отсчитывают по шкале прибора.

Например, измерение массы на равноплечных весах, когда воздействие массы m_x на весы частично уравнивается массой гирь m_0 , а разность масс отсчитывается по шкале весов, градуированной в единицах массы. В этом случае значение измеряемой величины $m_x = m_0 + \Delta m$, где Δm - показания весов.

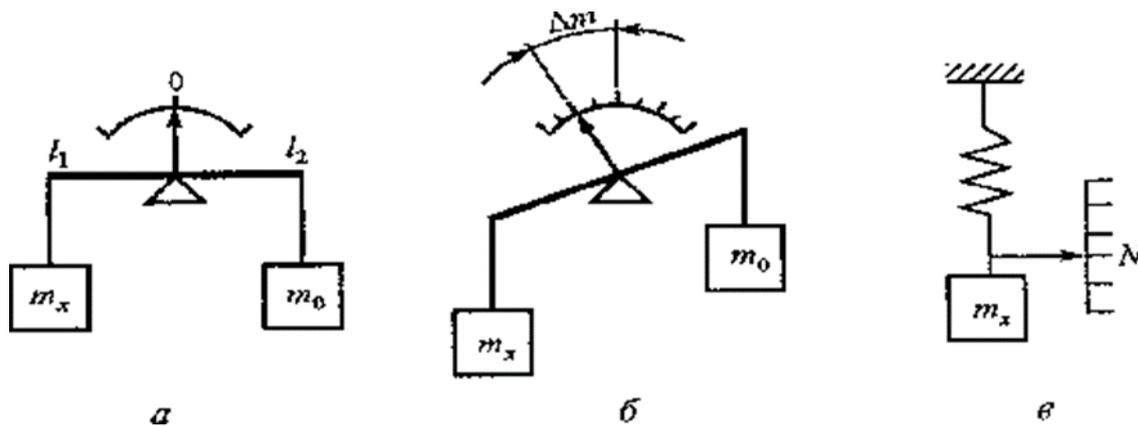


Рис. 1.4.2. Методы сравнений: а – нулевой метод; б – дифференциальный метод; в – метод замещения

Дифференциальный метод сравнения используют тогда, когда практическое значение имеет отклонение измеряемой величины от некоторого номинального значения.

Нулевой метод – метод, при котором действие измеряемой величины полностью уравновешивается образцовой (рис. 1.4.2, а).

Пример нулевого метода – взвешивание на весах, когда на одном плече находится взвешиваемый груз m_x , а на другом – набор эталонных грузов m_0 .

Метод замещения – метод, при котором измеряемую величину замещают известной величиной, воспроизводимой мерой (рис. 1.4.2, в).

Например, взвешивание на пружинных весах. Измерение производят в два приема. Вначале на чашу весов помещают взвешиваемую массу и отмечают положение указателя весов, затем массу m_x замещают массой гирь m_0 , подбирая ее так, чтобы указатель весов установился точно в том положении, что и в первом случае. При этом ясно, что $m_x = m_0$.

Метод совпадений – метод, когда разность между измеряемой величиной и величиной, воспроизводимой мерой, измеряют, используя совпадение отметок шкал или периодических сигналов.

Из всех перечисленных методов нулевой метод обеспечивает наибольшую точность измерений физической величины.

Контрольные вопросы

1. Назовите составляющие науки метрологии.
2. Дайте определение средства измерения.
3. Что обозначает термин «измерение»?
4. Что понимается под метрологической характеристикой?
5. Что такое система единиц физических величин?

2. 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:

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

2.1. 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 ни 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 следующие виды экспериментов:

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, температура 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.

2.2. 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 ни 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 Δx между результатом эксперимента x и истинным значением искомой величины x^* :

$$\Delta = |x - x^*|. \quad (2.2.1)$$

Относительная погрешность записывается в виде:

$$\Delta^* = \frac{|x - x^*|}{x^*} \cdot 100\% = \frac{\Delta}{x^*} \cdot 100\%. \quad (2.2.2)$$

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 Δx к нормирующему значению x_H , выраженному в процентах:

$$\gamma = \frac{\Delta x}{x_H} \cdot 100\%. \quad (2.2.3)$$

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. 2.2.1).

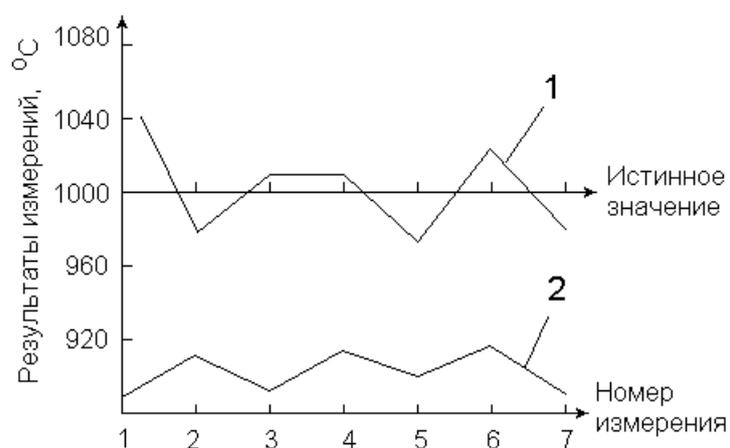


Рис. 2.2.1. Пример данных, иллюстрирующий различие между случайной и систематической погрешностями: 1 – измерения характеризуются наличием случайной погрешности; 2 – измерения характеризуются наличием систематической погрешности.

Можно выделить следующие источники систематических погрешностей:

1. Инструментальные (приборные или аппаратурные) погрешности средств измерений, 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.);

2. 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, зависящему не только от этой температуры, но и приведенной степени черноты $\varepsilon_{\text{пр}}$ ($q = \varepsilon_{\text{пр}} C_0 T^4$). (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.

3. 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 неизвестных экспериментатору и трудно поддающихся 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, при 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.

Контрольные вопросы

1. Что такое эксперимент? Какова его роль в инженерной практике?
2. Какие общие черты имеют научные методы исследований для изучения закономерностей различных процессов и явлений в промышленности?
3. Приведите классификации видов экспериментальных исследований, исходя из цели проведения эксперимента и формы представления результатов, а также в зависимости от условий его реализации.
4. В чем заключаются принципиальные отличия активного эксперимента от пассивного?
5. Поясните преимущества и недостатки лабораторного и промышленного эксперимента.
6. В чем отличие количественного и качественного экспериментов?

3. КРАТКИЕ СВЕДЕНИЯ ИЗ ТЕОРИИ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

3.1. 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

$$x_i \quad (i = 1, 2, \dots, n), \quad (3.1.1)$$

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 доменной печи, количество плавков в конвертере за 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:

$$P_i = \frac{m_i}{n} \quad n \rightarrow N, \quad (3.1.2)$$

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 3.1.1. Требуется определить вероятность числа остановок доменной печи, исходные данные представлены в табл. 3.1.1, а результаты расчета – на рис.3.1.1.

Таблица 3.1.1

Число остановок доменной 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

По графику, представленному на рис.3.1.1, 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?

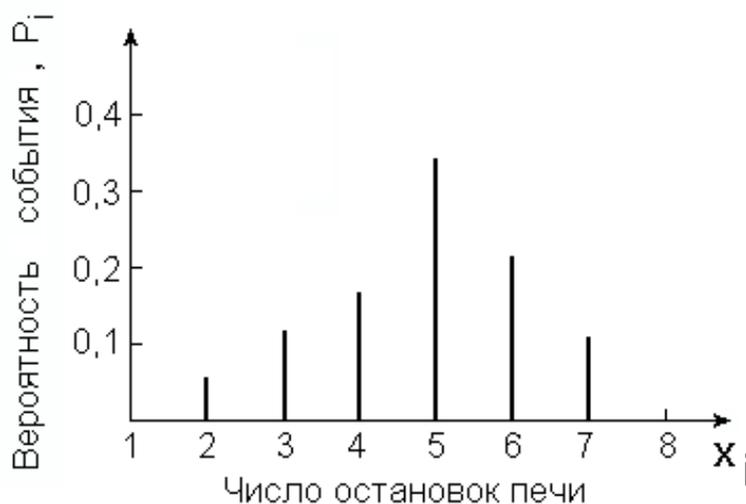


Рис. 3.1.1. Вероятность остановок доменной печи

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. $-\infty \leq X \leq x$

$$F(x) = P(X < x), \quad (3.1.3)$$

где x – текущая точка числовой оси.

Интегральная функция распределения имеет следующие свойства (рис.3.1.2):

$$1. 0 \leq F(x) \leq 1 \quad (3.1.4)$$

$$\lim_{x \rightarrow -\infty} F(x) = 0, \quad \lim_{x \rightarrow +\infty} F(x) = 1. \quad (3.1.5)$$

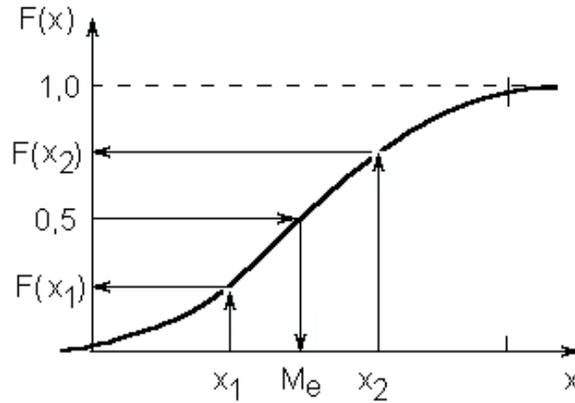


Рис. 3.1.2. Интегральная функция распределения: $F(x_2) - F(x_1) = P(x_1 \leq x \leq x_2)$

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

$$F(x_2) - F(x_1) = P(x_1 < x < x_2). \quad (3.1.6)$$

Дифференциальный закон распределения (плотность распределения) выражается соотношением (рис. 3.1.3)

$$f(x) = \frac{dF(x)}{dx}; F(x) = \int_{-\infty}^x f(Z)dZ, \quad (3.1.7)$$

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

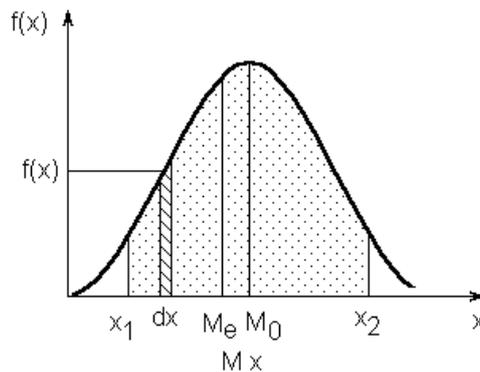


Рис. 3.1.3. Дифференциальный закон распределения (плотность распределения)

Напомним основные свойства $f(x)$:

1. $f(x) \geq 0$. (3.1.7, a)
2. Площадь, заключенная под кривой плотности распределения, согласно правилу нормировки, равна единице, т.е. отражает вероятность всех возможных событий.

$$\int_{-\infty}^{+\infty} f(Z)dZ = 1. \quad (3.1.7, б)$$

$$3. \quad \lim_{x \rightarrow \infty} f(x) = 0. \quad (3.1.7, в)$$

$$4. \quad F(x_2) - F(x_1) = P(x_1 < x < x_2) = \int_{x_1}^{x_2} f(Z)dZ. \quad (3.1.7, г)$$

Часто для характеристики случайной величины используют не сами функции распределения, а некоторые числовые параметры – параметры распределения (рис. 3.1.4). Важнейшими параметрами распределения, характеризующими случайную величину x , являются ее математическое ожидание M_x (центр рассеяния) и дисперсия σ_x^2 (степень рассеяния).

Рассмотрим характеристики положения центра распределения изучаемой случайной величины.

Для непрерывной случайной величины математическое ожидание определяется выражением

$$M_x = \int_{-\infty}^{+\infty} x \cdot f(x)dx, \quad (3.1.8)$$

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.

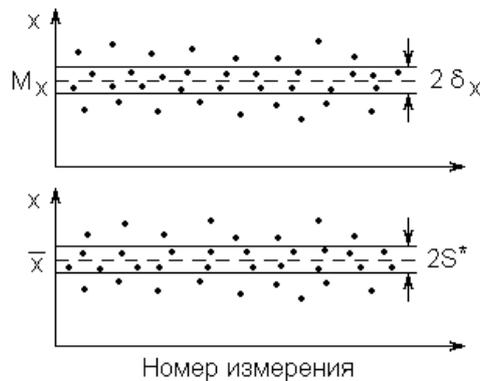


Рис. 3.1.4. Характеристика рассеяния случайной величины

Для дискретной случайной величины

$$M_x = \sum_{i=1}^n x_i \cdot p(x_i), n \rightarrow N, \quad (3.1.9)$$

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

Сказанное проиллюстрируем на рис. 3.1.3, 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 (3.1.8) даст абсциссу центра тяжести всей площади под кривой $f(x)$.

Для рассмотренного ранее примера 3.1.1 величина $M_x = 4,88$ и не равна ни одному из значений дискретной случайной величины (следовательно, оно может быть дробным).

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 3.1.1 $M_o = 5$).

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:

$$F(M_e) = \int_{-\infty}^{M_e} f(x)dx = 0,5, \quad (3.1.10)$$

или для дискретной случайной величины

$$\sum_{i=1}^{M_e} P(x_i) = 0,5; \quad x_i \leq M_e. \quad (3.1.11)$$

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 рассеяние случайных величин. Важнейшей из них является дисперсия (см. рис. 3.1.4).

Дисперсия для непрерывной случайной величины определяется соотношением

$$\sigma_x^2 = \int_{-\infty}^{+\infty} (x - M_x)^2 \cdot f(x)dx, \quad (3.1.12)$$

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

Для дискретной случайной величины

$$\sigma_x^2 = \sum_{i=1}^n (x_i - M_x)^2 \cdot P(x_i), \quad n \rightarrow N. \quad (3.1.13)$$

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

Положительное значение квадратного корня из дисперсии называется среднеквадратичным отклонением

$$\sigma_x = \sqrt{\sigma_x^2}. \quad (3.1.13a)$$

3.2. 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 бы 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. 3.2.1). 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.

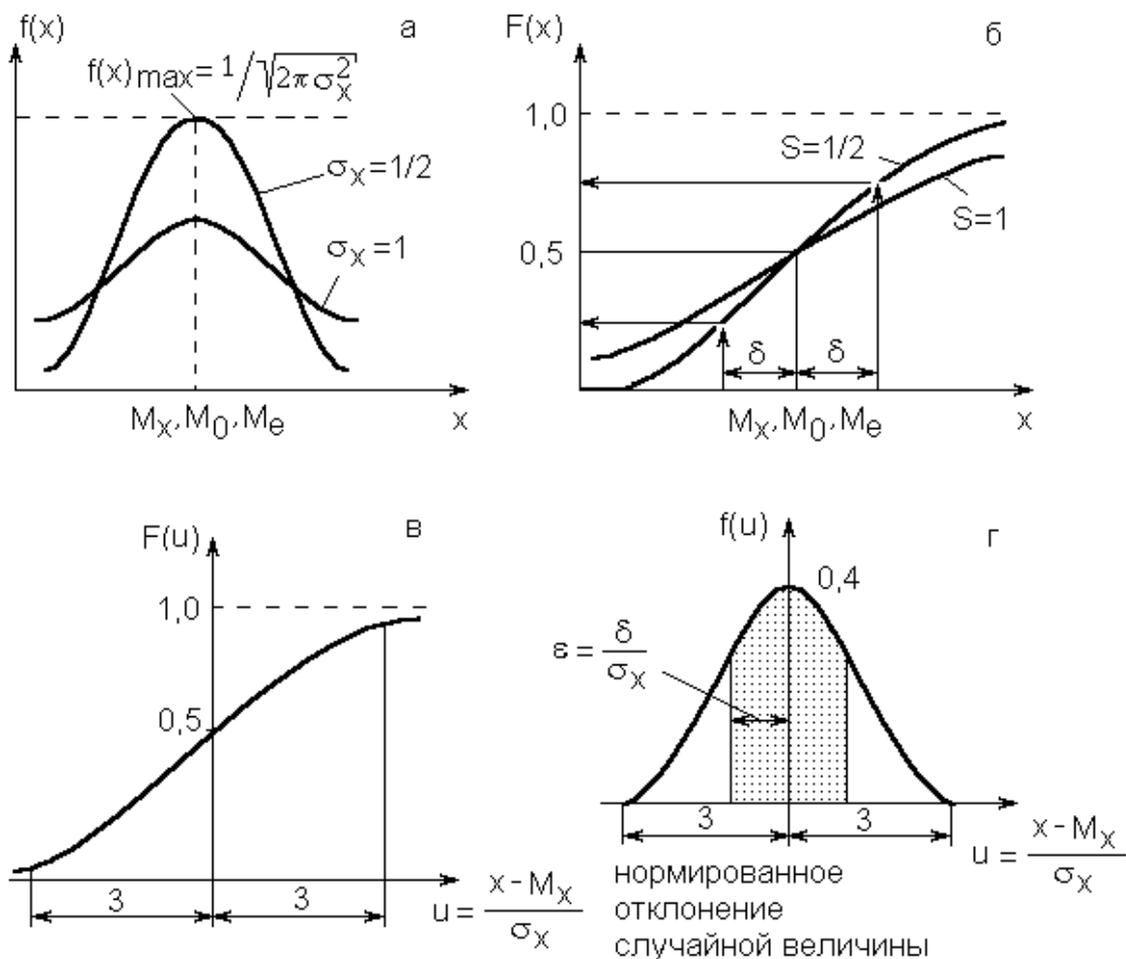


Рис. 3.2.1. Дифференциальная (а, г) и интегральная (б, в) функции при нормальном законе распределения случайных величин

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:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} \cdot e^{-\frac{[x-M_x]^2}{2\sigma_x^2}}. \quad (3.2.1)$$

Из уравнения (3.2.1) следует, что плотность распределения вероятностей для нормального распределения случайной величины определяется двумя параметрами M_x и σ_x^2 .

Интегральная функция нормально распределенной величины имеет вид:

$$F(x) = \int_{-\infty}^x f(Z) dZ = \frac{1}{\sqrt{2\pi\sigma_x^2}} \int_{-\infty}^x e^{-\frac{[Z-M_x]^2}{2\sigma_x^2}} dZ. \quad (3.2.2)$$

Отметим некоторые свойства нормального закона распределения.

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

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

3. Максимум ординаты кривой плотности распределения определяется выражением:

$$f_{max} = \frac{1}{\sqrt{2\pi\sigma_x^2}}, \quad (3.2.3)$$

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

4. Для нормального распределения среднее, мода и медиана совпадают

$$M_x = M_o = M_e. \quad (3.2.4)$$

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

$$u = \frac{x - M_x}{\sigma_x} = \frac{\Delta x}{\sigma_x}. \quad (3.2.5)$$

Для нормированного отклонения, распределенного по нормальному закону, получаем

$$f(u) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{u^2}{2}}; \quad (3.2.6)$$

$$F(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^u e^{-\frac{z^2}{2}} dz. \quad (3.2.7)$$

Графики этих функций показаны на рис. 3.2.1, в и рис. 3.2.1, 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 δ , т.е. нормально распределенная случайная величина примет значение на интервале $[a;b]$ (рис. 3.2.2). Эта вероятность называется доверительной вероятностью (P) и она показывает

вероятность того, что результат измерений отличается от истинного значения на величину, не большую δ :

$$F(M_x + \delta) - F(M_x - \delta) = P(M_x + \delta \leq x \leq M_x - \delta) = \frac{1}{\sqrt{2\pi\sigma_x^2}} \int_{M_x - \delta}^{M_x + \delta} e^{-\frac{(z-M_x)^2}{2\sigma_x^2}} dz. \quad (3.2.8)$$

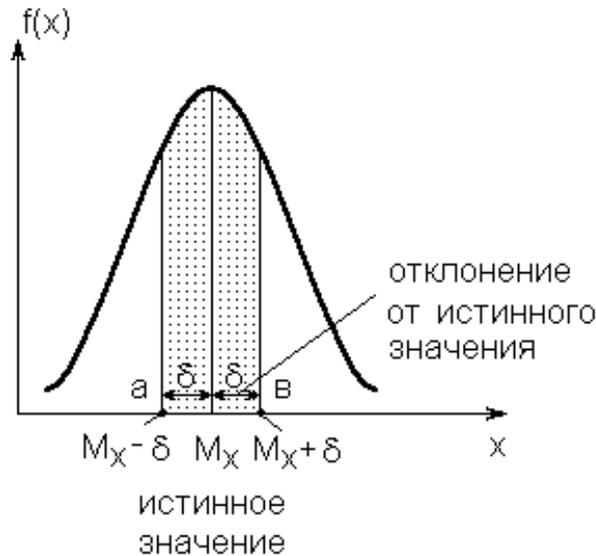


Рис. 3.2.2. Нормально распределенная случайная величина

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. 2.2.2. 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

$$\varepsilon = \delta/\sigma_x. \quad (3.2.9)$$

Тогда, в случае использования нормированного значения случайной величины, выражение (3.2.8) примет вид

$$P(M_x - \delta \leq x \leq M_x + \delta) = \frac{1}{\sqrt{2\pi}} \int_{-\varepsilon}^{+\varepsilon} e^{-\frac{z^2}{2}} dz. \quad (3.2.10)$$

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 Ox ($M_x = \pm\sigma_x$). 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. 3.2.1, в, 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 δ , тем с большей вероятностью величина x попадает в этот интервал.

Рассмотрим пример. 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:

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

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

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

$0,6 \pm 3,00 \times 0,15 = 0,6 \pm 0,45$ 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%.

Однако заметим, что мы предполагали нормальность закона распределения измерений, а также то, что изначально были известны математическое ожидание M_x и среднеквадратичное отклонение σ_x , т.е. выполнено большое (в пределе бесконечное) число измерений.

Контрольные вопросы

1. Что такое случайная величина? В чем заключаются отличия дискретной от непрерывной случайной величины? Приведите примеры.

2. Какие вероятностные характеристики используют для описания распределений случайных величин?

3. С какой целью используют законы распределения при обработке данных экспериментальных исследований?

4. Почему нормальный закон распределения наиболее применим в экспериментальной практике?

5. Какие параметры и свойства характерны для нормального закона распределения?

6. Какие задачи решают в ходе предварительной статистической обработки экспериментальных данных?

7. Что такое генеральная совокупность и выборка?

8. Что такое вариационный ряд выборки? Привести пример.

9. Что такое статистический ряд выборки? Привести пример.

10. Как определяется размах выборки?

11. В чем заключается сущность группированного статистического ряда?

12. Записать формулу для среднего арифметического выборки.

13. Дайте определение понятиям мода и медиана.

14. Что такое доверительный интервал и доверительная вероятность?

15. Как определяется доверительный интервал для среднего значения выборки?

4. 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.

4.1. 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) θ^* какого-либо параметра θ 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 θ .

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

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

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, i = 1, 2, \dots, n, \quad (4.1.1)$$

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.3.1.3):

$$\tilde{S}_x^2 = \frac{\sum_{i=1}^n (x_i - M_x)^2}{n}; \quad (4.1.2, а)$$

$$S_x^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}. \quad (4.1.2, б)$$

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

$$\tilde{S}_x = \sqrt{\tilde{S}_x^2}; \quad (4.1.3, а)$$

$$S_x = \sqrt{S_x^2} \quad (4.1.3, б)$$

Следует отметить, что в любой задаче, связанной с выполнением измерений, возможны два способа получения оценки значения σ_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 (4.1.3, а).

Второй способ получения оценки значения σ_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 (4.1.2, б). 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 ν , являющееся мерой относительной изменчивости случайной величины, вычисляют по формуле:

$$\nu = \frac{S_x}{\bar{x}} \quad (4.1.4а)$$

или в процентах

$$\nu = \frac{S_x}{\bar{x}} \cdot 100\% \quad (4.1.4б)$$

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 θ^* 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

$$\lim_{n \rightarrow \infty} S^2(n) = \sigma_x^2. \quad (4.1.5)$$

Оценка параметра θ^* 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 были 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 (4.1.2б), (4.1.3б), 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 θ_1^* называется эффективной, если среди прочих оценок того же параметра θ_2^*, θ_3^* она обладает наименьшей дисперсией.

$$M\{(\theta_1^* - \theta)^2\} = \min, \quad (4.1.6)$$

или

$$M\{(\theta_1^* - \theta)^2\} \leq M\{(\theta_i^* - \theta)^2\}, \quad (4.1.6a)$$

где θ_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:

$$\bar{x}_A = \frac{1}{n} \sum_{i=1}^n x_i, \quad \bar{x}_B = \frac{x_{\max}(n) + x_{\min}(n)}{2}, \quad (4.1.7)$$

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.

4.2. Статистические гипотезы

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

1. Отсев грубых ошибок (промахов) наблюдений;
2. Доверительная оценка измеряемых величин;
3. 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, σ_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 θ_A^* и θ_B^* , полученные для выборок A и B , принадлежат к одной генеральной совокупности, т.е. разница между θ_A^* и θ_B^* фактически равна нулю и возникла в силу случайности отбора элементов и ограниченности объема выборки. Противоречащие ей гипотезы H_i 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 Gr . Для этого строится случайная величина – степень рассогласования теоретического и экспериментального распределения. По величине рассогласования можно проверить, является ли это расхождение незначительным или существенным. Если $Gr_{\text{экс}}$ и $Gr_{\text{теор}}$ 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 α , который называется уровнем значимости. Произведение $\alpha \cdot 100\%$ показывает, в скольких случаях из 100 можно ошибаться, т.е. сколько процентов отвергается гипотезой. 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 α 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 случаях из 100 возможна ошибка. Конечно, никогда нельзя дать 100%-й гарантии. Подобное желание приводит к необходимости выполнять бесконечно 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 α . Область, отвечающая вероятности α , называется критической, а дополняющая ее область, вероятность попадания в которую $P(\theta_\alpha) = 1 - \alpha$, называется допустимой.

Вероятность ошибки второго рода обозначается β , 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 α показывает риск производства, отвергая правильную гипотезу, он бракует годную продукцию. Величина β характеризует риск потребителя, допускает ложную

гипотезу, он принимает и использует фактически непригодную продукцию. Ситуации, возникающие при статистической проверке гипотез, представлены в табл. 4.2.1.

Таблица 4.2.1

Возможные варианты принятия решения о правильности гипотез

Варианты принятия решения	Критерий Gr рекомендует допустить нуль-гипотезу H_0	Критерий Gr рекомендует отклонить нуль-гипотезу H_0 (т.е. допустить альтернативную H_1)
Фактически истинна 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

Общий подход решения задач при использовании статистических гипотез состоит в проверке нулевой гипотезы H_0 , т.е. 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 α .

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

4.3. 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. 3.1), что при нормальном законе распределения случайная величина не должна отличаться от своего математического ожидания с вероятностями: 90% на $1,64 \cdot \sigma_x$; 95% на $1,96 \cdot \sigma_x$ и т.п. Но это теоретические расчеты, в предположении, что известно математическое ожидание M_x , дисперсия σ_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

$$\left| \frac{x_{max} - M_x}{\sigma_x} \right| \geq \varepsilon(\alpha). \quad (4.3.1)$$

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 σ_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) σ_x . Поэтому при неизвестных действительных значениях M_x, σ_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 пивоварен 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 .

$$t = \frac{(x - \bar{x})}{S_x}. \quad (4.3.2)$$

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:

$$f(t) = \frac{\Gamma(n/2)}{\sqrt{\pi} \sqrt{n-1} \cdot \Gamma\left(\frac{n-1}{2}\right) \cdot \left(1 + \frac{t^2}{n-1}\right)^{n/2}}. \quad (4.3.3)$$

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

На рис. 4.3.1. 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 .

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 α).

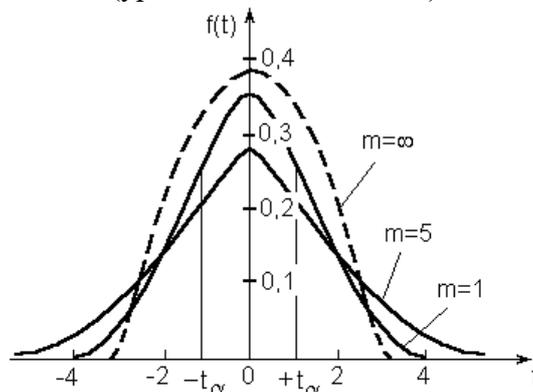


Рис. 4.3.1. Кривая t -распределения Стьюдента

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

$$t_{\text{эксп}} = \frac{|x_{\text{max}} - \bar{x}|}{S_x}, \quad (4.3.4)$$

где x_{max} – "выскакивающее" значение.

5. В качестве критерия проверки статистической нуль-гипотезы используется теоретическое значение критерия Стьюдента t , которое зависит от уровня значимости α 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 α представлены в справочной литературе. При наличии современных ПЭВМ можно воспользоваться пакетами прикладных программ или интегрированными средами, например, электронными таблицами Microsoft Excel.

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. 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, °C: 925, 950, 975, 1000, 1025, 1050 ($n = 6$).

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

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

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 °C и 1050 °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 $t_{\text{экср}} = 2,57$ было существенно выше 1,96, т.к. оно учитывало ограниченность экспериментальных данных.

4.4. Определение доверительных интервалов для исследуемых величин

Постановка задачи заключается в следующем: требуется оценить точность и достоверность M_x и σ_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. 4.4.1).

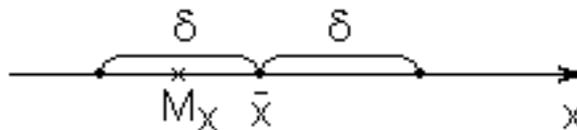


Рис. 4.4.1. К понятию доверительного интервала

Доверительный интервал (δ) – 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 θ 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 θ выполняется условие

$$\Theta_1^* < \Theta_2^*; P\{\Theta_1^*(x_1, x_2, \dots, x_n) \leq \Theta_2^*(x_1, x_2, \dots, x_n)\} = 1 - \alpha \quad (4.4.1)$$

Это означает, что действительное значение параметра θ находится в интервале значений (Θ_1^*, Θ_2^*) с вероятностью P .

Интервал (Θ_1^*, Θ_2^*) называют доверительным интервалом для неизвестного параметра θ , соответствующего доверительной вероятности (или надежности) $P = 1 - \alpha$.

Для симметричного доверительного интервала его ширина 2δ определяется условием

$$P\{|\theta - \bar{\theta}^*| \leq \delta\} = 1 - \alpha, \quad (4.4.2)$$

где $\bar{\theta}^*$ – оценка параметра θ ; δ – точность оценки. Чем меньше для данного α будет δ , тем точнее оценивается θ .

Вероятностное утверждение $P\{\Theta_1^* \leq \theta \leq \Theta_2^*\}$ следует понимать следующим образом: θ не есть случайная величина, лежащая с вероятностью P между Θ_1^* и Θ_2^* ; θ – есть константа нам неизвестная, но принимающая фиксированное значение. Пределы Θ_1^* и Θ_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 θ равна P . Таким образом, доверительный интервал – это интервал, который с заданной доверительной вероятностью P накрывает оцениваемый параметр.

4.4.1. 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

$$\sigma_y^2 = \sigma_{x_1}^2 + \sigma_{x_2}^2. \quad (4.4.1.1)$$

Дисперсия произведения случайной переменной X и постоянной величины C равна:

$$\delta^2(C \cdot x) = C \cdot \delta_x^2. \quad (4.4.1.2)$$

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 (4.4.1.1) и (4.4.1.2) имеем

$$\bar{S}_x^2 = S_x^2 \cdot \frac{1}{n}; \quad \bar{\sigma}_x^2 = \sigma_x^2 \cdot \frac{1}{n}; \quad (4.4.1.3)$$

$$\bar{S}_x = S_x \cdot \frac{1}{\sqrt{n}}, \quad \bar{\sigma}_x = \sigma_x \cdot \frac{1}{\sqrt{n}}, \quad (4.4.1.4)$$

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 (4.4.1.3, 4.4.1.4) 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 σ_x^2 (или другая, связанная с ней характеристика точности измерений), то доверительный интервал для оценки M_x , учитывая соотношение (4.4.1.4), рассчитывается очень просто:

$$\delta = \varepsilon(\alpha) \cdot \bar{\sigma}_x = \varepsilon(\alpha) \frac{\sigma_x}{\sqrt{n}}, \quad (4.4.1.5)$$

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

$$\bar{x} - 1,96 \frac{\sigma_x}{\sqrt{n}} \leq M_x \leq \bar{x} + 1,96 \frac{\sigma_x}{\sqrt{n}}, \quad (4.4.1.6)$$

а с вероятностью 0,90 в интервале

$$\bar{x} - 1,67 \frac{\sigma_x}{\sqrt{n}} \leq M_x \leq \bar{x} + 1,67 \frac{\sigma_x}{\sqrt{n}}. \quad (4.4.1.7)$$

Пример 4.4.1.1. 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 σ_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 σ_x^2 .

В этом случае следует заменить $\varepsilon(\alpha)$ на $t_{\alpha,m}$, а σ_x – на S_x и выражение (4.4.1.5) примет вид:

$$\delta = \frac{t_{\alpha,m} \cdot S_x}{\sqrt{n}} = t_{\alpha,m} \cdot \bar{S}_x, \quad (4.4.1.8)$$

где коэффициент $t_{\alpha,m}$ – теоретическое значение коэффициента Стьюдента при уровне значимости α 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.

Сравнивая выражения (4.4.1.5) и (4.4.1.8), 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 $(0,05) = 1,96$.

Пример 4.4.1.2 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 $0,59 \% \leq M_{Si} \leq 0,71 \%$.

4.4.2. Оценка доверительного интервала для дисперсии

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

$$\chi^2 = \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{\sigma_x} \right)^2 = \frac{n-1}{\sigma_x^2} \cdot S_x^2, \quad (4.4.2.1)$$

то величина χ^2 , связанная с дисперсией σ_x , будет иметь следующий закон распределения:

$$f(\chi^2) = \frac{1}{2^{m/2} \cdot \Gamma(m/2)} \cdot (\chi^2)^{m/2-1} \cdot e^{-\chi^2/2}, \quad 0 \leq \chi^2 \leq \infty, \quad (4.4.2.2)$$

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

На рис. 4.4.2.1 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 $\chi_{max}^2 = m - 2$. При больших значениях m χ^2 -распределение переходит в нормальное со средним значением $\sqrt{2m-1}$ и дисперсией $\sigma_{\chi^2} = 1$.

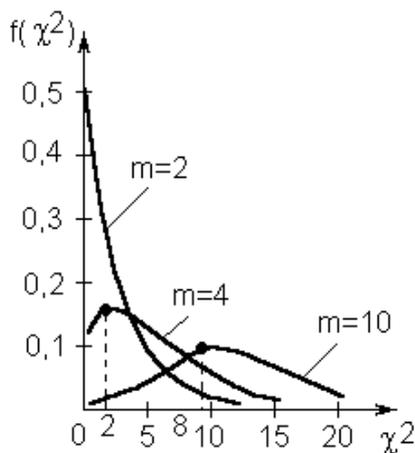


Рис. 4.4.2.1 Кривые χ^2 -распределения для различных степеней свободы m

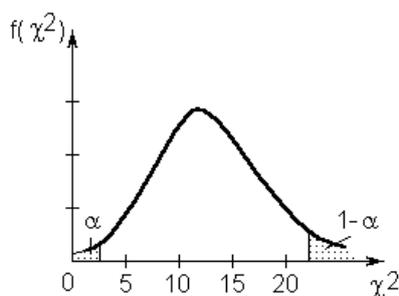


Рис. 4.4.2.2. Определение границ доверительного интервала для оценки дисперсии γ^2 при значении надежности $P = 1 - 2\alpha_1$

Применяя соотношение (4.4.2.1), 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 γ^2 остается неизвестной, т.к. значение χ^2 , входящее в это равенство, а следовательно, и значение коэффициента γ^2 остается неизвестным. Тем не менее с помощью χ^2 -распределения можно найти границы доверительных интервалов для σ_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. 4.4.2.2). Эти значения $\chi_2^2(m)$ и $\chi_1^2(m)$ определяют величины соответствующих им γ_1^2 и γ_2^2 . Таким образом, мы получили два неравенства, обуславливающих одинаковую надежность определения границ доверительного интервала для γ_x^2 с двух сторон (слева и справа):

$$\chi^2 = \frac{m \cdot S_x^2}{\gamma_x^2} \leq \chi_1^2 = \frac{m}{\gamma_1^2}. \quad (4.4.2.3)$$

$$\chi^2 = \frac{m \cdot S_x^2}{\gamma_x^2} \leq \chi_2^2 = \frac{m}{\gamma_2^2}. \quad (4.4.2.4)$$

Из неравенств (4.4.2.3) и (4.4.2.4) получим, что вероятность попадания χ^2 в интервал $(\chi_1^2; \chi_2^2)$, т.е.

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

равна разности вышеупомянутых площадей $\alpha_2 - \alpha_1 = P$. Значения γ_1^2 и γ_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

$$\bar{\gamma}_x^2 = \sigma_x^2/n, \bar{S}_x^2 = S_x^2/n, \text{ то } \gamma_1 \bar{S}_x \leq \bar{\sigma}_x \leq \gamma_2 \bar{S}_x.$$

Пример 4.4.2.1 В серии из десяти измерений ($n = 10$) было получено $\bar{S}_x^2 = 5,4 \cdot 10^{-3}$ мм². Требуется определить, в каком интервале находится фактическое значение дисперсии σ .

При $P=0,9$, $m = n - 1 = 9$, поэтому на основе (3.4.2.3) и (3.4.2.4) имеем

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

Значения χ_1^2 с надежностью $\alpha_1 = (1 - P)/2 = 0,05$ и χ_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. 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 $2,55 \cdot 10^{-3} \leq \sigma \leq 1,46 \cdot 10^{-2}$.

4.5. 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 правильность показаний приборов, 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.

4.5.1. Сравнение средних значений

Сравнение средних значений продемонстрируем на следующих задачах:

Задача 1. Сформулируем нуль-гипотезу $H_0: M_x = M_э$ т.е. математическое ожидание по выборке M_x равно математическому ожиданию эталона $M_э$.

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

$$t_{\text{эксп}} = \frac{\bar{x} - M_э}{S_x} \cdot \sqrt{n}. \quad (4.5.1.1)$$

В дальнейшем определяется теоретическое значение этого критерия при заданном уровне значимости α и числе степеней свободы $m = n - 1$. Для этого опять же можно воспользоваться таблицами из статистических справочников или пакетами прикладных программ для ЭВМ. 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 (4.5.1.1).
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.

Пример 4.5.1.1 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

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

В дальнейшем рассчитывается эмпирическое значение критерия

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

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.

Задача 2. Сравняются 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?

Воспользуемся 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 :

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

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

4. Рассчитывается экспериментальное значение критерия Стьюдента. Учитывая, что $M_x = 0$, выражение (4.5.1.1) примет вид:

$$t_{\text{эксп}} = \frac{\bar{x} - 0}{S_x} \cdot \sqrt{n} = \frac{\bar{x}}{S_x} \cdot \sqrt{n}. \quad (4.5.1.2)$$

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

4.5.2. 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 σ_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 (большой к меньшей), вычисленных или полученных различными способами:

$$F_{\text{эксп}} = \frac{S_{\text{больш.}}^2}{S_{\text{меньш.}}^2}. \quad (4.5.2.1)$$

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.

Иными словами, имеются значения критерия F_{α, 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 уровне значимости α , 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:

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

$$m_1 = n_1 - 1; m_2 = n_2 - 1. \quad (4.5.2.2)$$

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_{α, m_1, m_2} .

Для примера с бетоном теоретическое значение критерия Фишера мы определим с помощью статистической: $F_{0,05;7;16} = F_{\text{РАСПОБР}}(0,05; 7; 16) = 2,66$.

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 4.5.2.1. 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:

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

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 $F_{0,05;199;14} = 1649$ и $F_{\text{эксп}} > F_{\text{теор}}$, т.е. результаты измерений новым прибором лучше, чем старым! Следовательно, принимается альтернативная гипотеза H_1 .

4.5.3. 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 $S_{x_{max}}^2$ заметно больше всех остальных. Задача заключается в том, чтобы выяснить, можно ли считать отличие выделенной дисперсии $S_{x_{max}}^2$ 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

$$S_{x_i}^2 = \frac{\sum_{j=1}^{m^*} (x_{ij} - \bar{x}_i)^2}{m^* - 1}, \quad (4.5.3.1)$$

где $\bar{x}_i = \frac{\sum_{j=1}^{m^*} x_{ij}}{m^*}$ и выделим из них максимальную $S_{x_{max}}^2$ (рис. 4.5.3.1).

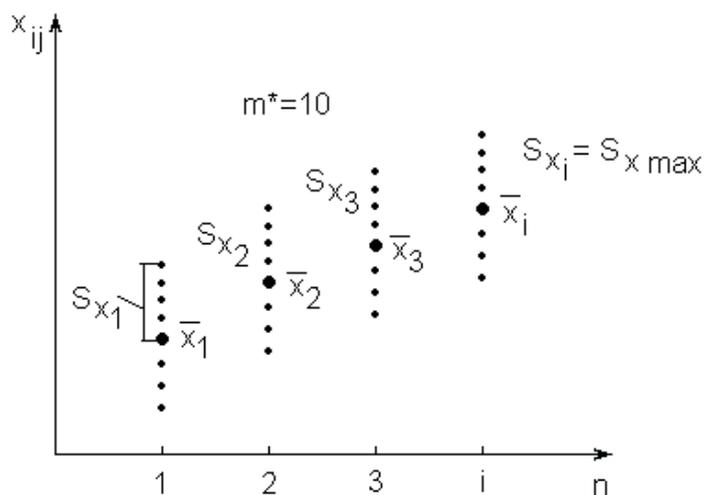


Рис. 4.5.3.1. К понятию дисперсий эксперимента

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

$$G_{\text{эксп}} = \frac{S_{x_{max}}^2}{\sum_{i=1}^n S_{x_i}^2}. \quad (4.5.3.2)$$

В дальнейшем определяется теоретическое значение этого критерия, которое зависит от уровня значимости α , 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":

$$S_{\text{восп}}^2 = \frac{\sum_{i=1}^n S_{x_i}^2}{n}, \quad (4.5.3.3)$$

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

$$S_{\text{восп}} = \sqrt{S_{\text{восп}}^2} = \sqrt{\frac{\sum_{i=1}^n S_{x_i}^2}{n}}. \quad (4.5.3.4)$$

Пример 4.5.3.1. Пусть шестью ($n = 6$) приборами произведено по семь измерений ($m^* = 7$) одного и того же параметра, которые дали следующие дисперсии, $S_{x_i}^2$: 3,82; 1,7; 1,3; 0,92; 0,78; 0,81.

Тогда экспериментальное значение критерия Кохрена в соответствии с уравнением (4.5.3.2) 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 $S_{x_{\max}}^2 = 3,82$ от остальных с вероятностью 95% несущественно, а дисперсии однородны. Дисперсия воспроизводимости (ее оценка) равна

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

а среднее квадратичное отклонение измерений составит

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

Попытайтесь самостоятельно оценить при каком значении $S_{x_i}^2$ 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 Бартлетта. 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 4.5.3.2. После 100 испытаний материала на прочность было установлено, что $\bar{x}^{(1)} = 150$ кг/см², $S_x^{(1)} = 20$ кг/см². 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 4.5.3.1. 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?

Результаты испытаний на прочность различных образцов материала

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

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

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

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:

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

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. $H_0: M_{x_1} = M_{x_2}$. Альтернативная гипотеза "Математическое ожидание прочности металла по новой технологии возросло", т.е. $H_1: M_{x_1} > M_{x_2}$. Определим средневзвешенную оценку дисперсии (суммарную выборочную дисперсию):

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

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

Рассчитаем эмпирическое значение критерия Стьюдента:

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

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.

4.6. Определение необходимого количества измерений

Увеличение количества измерений (числа проб, образцов и т.п.), как видно из выражений (4.4.1.5), (4.4.1.8) 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) δ .

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

Действительно, при этих допущениях число измерений можно определить из выражения (4.4.1.5)

$$n \geq \left(\frac{\varepsilon(\alpha) \cdot \sigma_x}{\delta} \right)^2 = \varepsilon^2(\alpha) \cdot \left(\frac{\sigma_x}{\delta} \right)^2. \quad (4.6.1)$$

Таким образом, число измерений n определяется требуемой надежностью P (доверительной вероятностью α) и относительным (по отношению к среднеквадратичному отклонению) значением доверительного интервала (σ_x/δ), 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 (σ_x) неизвестно, а имеется только ее оценка (S_x). В этом случае следует воспользоваться соотношением (4.4.1.8), т.е. критерием Стьюдента. Необходимое число измерений следует определять из следующего соотношения

$$n \geq \frac{t_{\alpha,m}^2 \cdot S_x^2}{\delta^2} = t_{\alpha,m}^2 \cdot \left(\frac{S_x}{\delta} \right)^2. \quad (4.6.2)$$

При расчетах по этому уравнению следует иметь в виду, что теоретическое значение критерия Стьюдента зависит не только от доверительной вероятности α , но и числа степеней свободы m , последние же определяются числом измерений. В связи с этим уравнение (4.6.2) следует решать методом последовательных приближений. В качестве начального приближения можно задать, в частности, число измерений, рассчитанных по формуле (4.6.1). 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 (4.6.1) и (4.6.2) 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 при различных значениях величины доверительного интервала δ/S_x и δ/σ_x (в скобках) приведено в табл. 4.6.1.

Таблица 4.6.1

Необходимое количество измерений

δ/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

Контрольные вопросы.

1. В чем заключается сущность статистических гипотез? Что такое нулевая и альтернативная статистические гипотезы?
2. С помощью каких критериев производится отсев грубых погрешностей?
3. Какие задачи возникают при сравнении двух рядов наблюдений экспериментальных данных? С помощью каких критериев они решаются?
4. Что такое критерий согласия? Какова основная идея его использования при проверке гипотез о виде функции распределения?
5. В чем заключается алгоритм использования критерия Пирсона для проверки гипотезы нормального распределения экспериментальных данных?
6. Какова процедура использования критерия Колмогорова-Смирнова для проверки гипотезы нормального распределения?
7. Сущность ошибок первого и второго рода, возникающих при проверке статистических гипотез.
8. В чем заключается основная идея оценивания с помощью доверительного интервала? С помощью каких распределений происходит построение доверительных интервалов для математического ожидания и дисперсии?

5. АНАЛИЗ РЕЗУЛЬТАТОВ ПАССИВНОГО ЭКСПЕРИМЕНТА. ЭМПИРИЧЕСКИЕ ЗАВИСИМОСТИ

5.1. 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 $\hat{y} = f(X)$. Для этого должны одновременно определяться как значения X , так и соответствующие им значения \hat{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

$$\hat{y} = f(x), \quad (5.1.1)$$

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

$$\hat{y} = f(X), \quad (5.1.1a)$$

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. 5.1.1а). 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. 5.1.1. 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. 5.1.1, б). 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 $\hat{y} = f(x)$ может быть найдена лишь в результате совместной обработки полученных значений x и y .

На рис. 5.1.1, б это кривая зависимости, проходящая по центру полосы экспериментальных точек (математическому ожиданию), которые могут и не лежать на искомой кривой $\hat{Y} = f(x)$, 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 и 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 $\hat{y} = f(X)$ 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

$$\hat{y} = b_0 + \sum_{j=1}^k b_j x_j, \quad (5.1.2)$$

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. 5.1.2). Для этого весь диапазон изменения x на поле корреляции разбивают на равные интервалы Δx . Все точки, попавшие в данный интервал Δx_j , относят к его середине \bar{x}_j . Для этого подсчитывают частные средние для каждого интервала

$$\bar{y}_j = \frac{\sum_{i=1}^{n_j} y_{ji}}{n_j}. \quad (5.1.3)$$

Здесь n_j – число точек в интервале Δx_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. 5.1.1б). В дальнейшем уточним это положение.

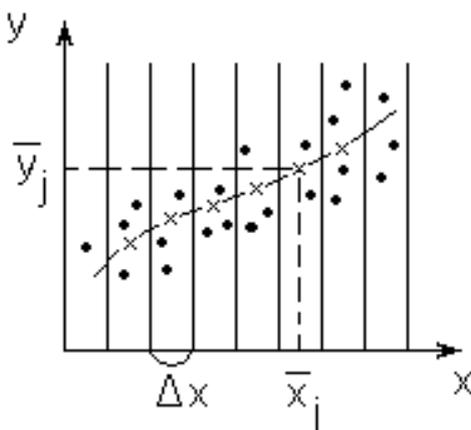


Рис. 5.1.2. Корреляционное поле

5.2. 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 $\hat{y} = (X, \mathbf{b})$ совпадала с экспериментальными значениями в некоторых точках, выбранных в качестве опорных (основных, главных) u_i .

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

$$f(x_i, b_0, \dots, b_j, \dots, b_k) = y_i, 1 \leq i \leq n \quad (5.2.1)$$

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.4.2.1).

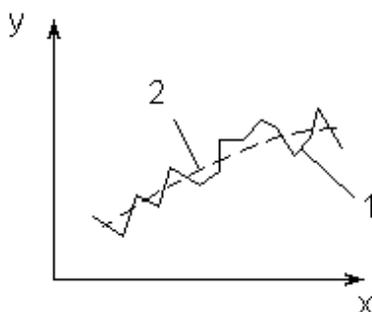


Рис. 5.2.1. Аппроксимация функции с большим (1) и небольшим (2) числом коэффициентов b_i

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 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:

1. Метод избранных точек (рис. 5.2.2). На основании анализа данных выдвигают гипотезу о виде (форме) зависимости $f(X)$. Предположим, что она линейная, т.е. статистическая связь – это линейная одномерная регрессия

$$\hat{y} = b_0 + b_1x \quad (5.2.2)$$

Выбирают две наиболее характерные по мнению исследователя точки, через которые и проходит линия регрессии (см. рис. 5.2.2). 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.

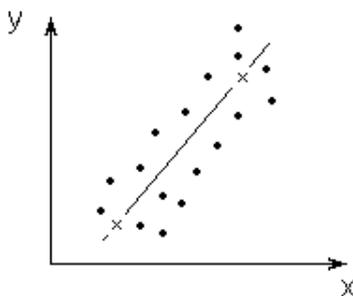


Рис. 5.2.2. Метод избранных точек (× – избранные точки)

2. Метод медианных центров. Сущность этого метода поясняет рис. 5.2.3. 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 (5.2.3) 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:

$$\bar{y}_I = b_0 + b_1\bar{x}_I, \bar{y}_{II} = b_0 + b_1\bar{x}_{II} \quad (5.2.2a)$$

Если при выборе вида уравнения регрессии число его коэффициентов b_j окажется больше числа уравнений (имеющихся результатов измерений) $k + 1 > n$, система (5.2.1) 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.

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:

$$\Phi(b_0, b_1, \dots, b_j, \dots, b_k) = \sum_{i=1}^n [f(x_i, b_0, b_1, \dots, b_j, \dots, b_k) - y_i]^2 \rightarrow \min b_j, \quad (5.2.3)$$

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

$$\partial\Phi/\partial b_j = 0, 0 \leq j \leq k \quad (5.2.4)$$

или

$$\sum_{i=1}^n [f(x_i, b_0, b_1, \dots, b_j, \dots, b_k) - y_i] \frac{\partial f(x_i)}{\partial b_j} = 0, 0 \leq j \leq k \quad (5.2.4, a)$$

После преобразований получим

$$\sum_{i=1}^n f(x_i, b_0, b_1, \dots, b_j, \dots, b_k) \frac{\partial f(x_i)}{\partial b_j} - \sum_{i=1}^n y_i \frac{\partial f(x_i)}{\partial b_j} = 0 \quad (5.2.5)$$

Система уравнений (5.2.5) 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 Φ 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 (5.2.5) 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.

5.3. 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. 4.5.2 и п. 4.5.3), 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 4.5.1.3, 4.5.2.1, 4.5.3.1).

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

5.3.1. 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) $S_{y_{\text{ост}}}^2$.

Напомним, что

$$S_y^2 = \frac{\sum_{i=1}^n [y_i - \bar{y}]^2}{n - l}; S_{y_{\text{ост}}}^2 = \frac{\sum_{i=1}^n [y_i - \hat{y}_i]^2}{n - l}, \quad (5.3.1.1)$$

где $l = k + 1$ – число членов аппроксимирующего полинома, а k – число факторов. Так, например для линейной зависимости (5.2.2) $k = 1, l = 2$.

В дальнейшем определяется экспериментальное значение F -критерия Фишера

$$F_{\text{эксп}} = S_y^2 / S_{y_{\text{ост}}}^2, \quad (5.3.1.2)$$

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_{α, m_1, m_2} для выбранного α 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 $\bar{y}_i = \sum_{j=1}^{m^*} y_{ij} / m^*$ – 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)

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

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_{α, m_1, m_2} , где $m_1 = m^*(n - 1)$, $m_2 = n - 1$.

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

5.3.2. Проверка значимости коэффициентов уравнения регрессии

Надежность оценок b_i уравнения регрессии можно охарактеризовать их доверительными интервалами Δb_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 $S_{y_i}^2 = \text{const}$.

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

$$S_{b_i} = \sqrt{\sum_{j=1}^n \left(\frac{\partial b_j}{\partial y_i} \right)^2 S_j^2}. \quad (5.3.2.1)$$

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

$$S_{b_0} = \sqrt{\frac{S_{\text{восп}}^2 \sum_{i=1}^n x_i^2}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}} \quad (5.3.2.2a)$$

$$S_{b_1} = \sqrt{\frac{S_{\text{восп}}^2 \cdot n}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}}. \quad (5.3.2.2б)$$

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 α отличен от нуля.

Построим доверительный интервал для коэффициентов уравнения регрессии

$$\Delta b_i = t_{\alpha; n-l} \cdot S_{b_i}, \quad (5.3.2.3)$$

где число степеней свободы в критерии Стьюдента определяется по соотношению $n - l$. Потеря $l = k + 1$ степеней свободы обусловлена тем, что все коэффициенты b_i рассчитываются зависимо друг от друга.

Тогда доверительный интервал для Δb_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.

Контрольные вопросы.

1. В чем заключается сущность статистических гипотез? Что такое нулевая и альтернативная статистические гипотезы?
2. С помощью каких критериев производится отсев грубых погрешностей?
3. Какие задачи возникают при сравнении двух рядов наблюдений экспериментальных данных? С помощью каких критериев они решаются?
4. Что такое критерий согласия? Какова основная идея его использования при проверке гипотез о виде функции распределения?
5. В чем заключается алгоритм использования критерия Пирсона для проверки гипотезы нормального распределения экспериментальных данных?
6. В чем заключаются сущность и основные задачи корреляционного, регрессионного и дисперсионного анализа?
7. Какие подходы используют при нахождении коэффициентов уравнения регрессии?
8. Сформулируйте исходные положения метода наименьших квадратов.
9. С помощью какого параметра оценивается теснота связи между случайными величинами? Поясните физическую суть этого параметра.
10. Как оценивается адекватность статистической модели?
11. Что такое частный коэффициент корреляции?
12. Дайте определение множественного коэффициента корреляции?
13. Какими свойствами обладают коэффициенты корреляции?
14. Каким образом производится проверка значимости коэффициентов уравнения регрессии?
15. В чем заключается постановка задачи линейной множественной регрессии?

6. ОЦЕНКА ПОГРЕШНОСТЕЙ РЕЗУЛЬТАТОВ НАБЛЮДЕНИЙ

6.1. 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 понимаются такие, 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.

$$\hat{y} = f(x_1, x_2, \dots, x_i, \dots, x_k), \quad (6.1.1)$$

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 у 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 Δx_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

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

где

$$\Delta y_i = \left(\frac{\partial f(x_1, \dots, x_k)}{\partial x_i} \right)_{x_i=\bar{x}_i} \cdot \Delta x_i. \quad (6.1.2)$$

Следовательно, Δy_i – это составляющие погрешности функции, обусловленные погрешностью i -го аргумента x_i .

Доверительная вероятность, соответствующая величине Δy_i , численно равна доверительной вероятности, с которой найдена погрешность Δx_i .

Для относительной погрешности вместо соотношения (6.1.2) используют выражение:

$$\Delta y_i^* = \frac{1}{f} \frac{\partial f}{\partial x_i} \Delta x_i = \frac{\partial \ln(f)}{\partial x_i} \Delta x_i. \quad (6.1.3)$$

Соотношения (6.1.2) и (6.1.3) применимы для расчета как случайных, так и систематических погрешностей.

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

$$\Delta y_\Sigma = \pm \sqrt{\sum_{i=1}^k (\Delta y_i)^2}; \quad (6.1.4)$$

$$\Delta_{\Sigma}^* = \pm \sqrt{\sum_{i=1}^k (\Delta y_i^*)^2}. \quad (6.1.5)$$

Предполагается, что все составляющие имеют нормальный закон распределения.

Частные производные, входящие в соотношения (6.1.2) и (6.1.3), 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.

Пример 6.1.1. 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

$$G = \alpha F_0 \varepsilon \sqrt{2\rho(p_1 - p_2)} = \alpha F_0 \varepsilon \sqrt{2\rho h}, \quad (6.1.6)$$

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

Используя соотношения (6.1.2) и (6.1.4), получим следующую формулу для расчета абсолютной и относительной погрешности определения расхода:

$$\Delta G = \pm \sqrt{\left(\frac{\partial G}{\partial \rho} \Delta \rho\right)^2 + \left(\frac{\partial G}{\partial h} \Delta h\right)^2}; \quad (6.1.7)$$

$$\Delta_G^* = \pm \sqrt{\left(\frac{\partial G}{\partial \rho} \frac{\Delta \rho}{G}\right)^2 + \left(\frac{\partial G}{\partial h} \frac{\Delta h}{G}\right)^2} = \pm \frac{1}{2} \sqrt{\left(\frac{\Delta \rho}{\rho}\right)^2 + \left(\frac{\Delta h}{h}\right)^2}, \quad (6.1.8)$$

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:

$$\Delta \rho = \pm \sqrt{\left(\frac{\partial \rho}{\partial p} \Delta p\right)^2 + \left(\frac{\partial \rho}{\partial T} \Delta T\right)^2}, \quad (6.1.9)$$

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

относительная погрешность

$$\Delta_{\rho}^* = \frac{\Delta \rho}{\rho} = \pm \sqrt{\left(\frac{\Delta p}{p}\right)^2 + \left(\frac{\Delta T}{T}\right)^2}. \quad (6.1.10)$$

Тогда относительная погрешность определения массового расхода газового потока будет равна:

$$\Delta_G^* = \pm \frac{1}{2} \sqrt{\left(\frac{\Delta p}{p}\right)^2 + \left(\frac{\Delta T}{T}\right)^2 + \left(\frac{\Delta h}{h}\right)^2}. \quad (6.1.11)$$

Здесь p , T , h – значения измеренных параметров; Δp , ΔT , Δh – их абсолютные погрешности. Численные значения Δp , ΔT , Δh 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

$$\Delta G = G \Delta G^*, \quad (6.1.12)$$

где G – значение расхода, измеренное экспериментально.

$$G_{\text{ист}} = G \pm \Delta G. \quad (6.1.13)$$

6.2. 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 (6.1.4) и (6.1.5) оказывали одинаковое влияние на погрешности функции.

Применяя принцип равных влияний к относительной погрешности функции, определяемой соотношением (6.1.5), получим

$$(\Delta_{y_1}^*)^2 = (\Delta_{y_2}^*)^2 = \dots = (\Delta_{y_r}^*)^2 = \frac{(\Delta_{\Sigma}^*)^2}{k}, \quad (6.2.1)$$

$$\Delta_{y_i} = \frac{\Delta_{\Sigma}^*}{\sqrt{k}}. \quad (6.2.2)$$

С учетом (6.1.3) легко получить выражение для определения абсолютных Δx_i и относительных $\Delta_{x_i}^*$ погрешностей всех аргументов:

$$\Delta x_i = \frac{\Delta_{\Sigma}}{\sqrt{k}} \frac{1}{\frac{\partial \ln[f(x_1, x_2, \dots, x_k)]}{\partial x_i}}; \quad (6.2.3)$$

$$\Delta_{x_i}^* = \frac{\Delta_{\Sigma}}{x_i \sqrt{k}} \frac{1}{\frac{\partial \ln[f(x_1, x_2, \dots, x_k)]}{\partial x_i}}. \quad (6.2.4)$$

В дальнейшем могут иметь место три возможных случая:

1. Значения погрешностей всех аргументов таковы, что лежат в пределах точности, доступной при измерениях с помощью имеющихся средств измерений.

2. Значения некоторых погрешностей настолько малы, что обеспечить соответствующую точность с помощью имеющихся средств измерений не представляется возможным.

3. 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 вторые производные окажутся положительными, то это соответствует минимуму величины Δ_{Σ}^* .

Контрольные вопросы.

1. Что такое погрешность определения величин функций?
2. С какой целью рассчитывают погрешность?
3. Какие виды погрешностей существуют? Как они определяются?
4. В чем заключается цель решения обратной задачи теории экспериментальных погрешностей?
5. Что понимают под выражением «наивыгоднейшие условия проведения эксперимента»?
6. Какова основная идея математического решения задачи поиска наивыгоднейших условий проведения эксперимента?
7. Из каких этапов состоит последовательность проведения активного эксперимента?

7. МЕТОДА ПЛАНИРОВАНИЯ ЭКСПЕРИМЕНТОВ. ЛОГИЧЕСКИЕ ОСНОВЫ

7.1. Основные определения и понятия

Ранее мы рассматривали пассивный эксперимент и математическая статистика использовалась, в частности, при обработке экспериментальных данных. На стадии постановки эксперимента она не использовалась. При активном же эксперименте математическая статистика используется уже на стадии постановки и планирования эксперимента.

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.

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

$$y = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i,j=1}^k \beta_{ij} x_i x_j + \sum_{i=1}^k \beta_{ii} x_i^2 + \dots, \quad (7.1.1)$$

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

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

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 β , и данный полином заменяется уравнением вида:

$$\hat{y} = b_0 + \sum_{i=1}^k b_i x_i + \sum_{i,j=1}^k b_{ij} x_i x_j + \sum_{i=1}^k b_{ii} x_i^2 + \dots, \quad (7.1.2)$$

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.

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.

7.2. 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 таблице 7.2.1.

Таблица 7.2.1

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 то же распределение. Тогда дисперсия измерения веса образца равна

$$\sigma_A^2 = \sigma_{y_1}^2 + \sigma_{y_0}^2 = 2\sigma^2, \quad (7.2.1)$$

где σ^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 таблице 7.2.2.

Планирование эксперимента при взвешивании трех объектов

Номер опыта	A	B	C	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 A, B, C, в последнем опыте взвешивают объекты A, B, C, т.е. все три объекта вместе, а «холостое» взвешивание не производится.

Легко заметить, что масса каждого объекта будет задаваться формулами:

$$\begin{aligned} m_A &= \frac{1}{2}(y_1 - y_2 - y_3 + y_4); \\ m_B &= \frac{1}{2}(y_2 - y_1 - y_3 + y_4); \\ m_C &= \frac{1}{2}(y_3 - y_1 - y_2 + y_4); \end{aligned} \quad (7.2.2)$$

Масса объекта A, вычисленная по приведенной выше формуле, оказывается не искаженной массами весов объектов B и C, т.к. масса каждого из них входит в формулу для массы A дважды с разными знаками.

Найдем теперь дисперсию, связанную с ошибкой взвешивания по новой схеме постановки экспериментов

$$\sigma_A^2 = \frac{1}{4}(\sigma_{y_1}^2 + \sigma_{y_2}^2 + \sigma_{y_3}^2 + \sigma_{y_4}^2) = \sigma^2. \quad (7.2.3)$$

Аналогичным образом находим:

$$\sigma_B^2 = \sigma^2, \sigma_C^2 = \sigma^2.$$

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.

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.

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

$$\hat{y} = b_0 + \sum_{i=1}^k b_i x_i + \sum_{\substack{i,j=1 \\ i \neq j}}^k b_{ij} x_i x_j + \sum_{\substack{i,j,u=1 \\ i \neq j \neq u}}^k b_{iju} x_i x_j x_u + \dots \quad (7.2.4)$$

Планы второго порядка позволяют провести эксперимент для отыскания уравнения регрессии, содержащего и вторые степени факторов.

$$\hat{y} = b_0 + \sum_{i=1}^k b_i x_i + \sum_{i=1}^k b_{ii} x_i^2 + \sum_{\substack{i,j=1 \\ i \neq j}}^k b_{ij} x_i x_j + \dots \quad (7.2.5)$$

Нахождение уравнения регрессии методом планирования экспериментов состоит из следующих этапов:

1. Выбор основных факторов и их уровней.
2. Планирование и проведение собственно эксперимента.
3. Определение коэффициентов уравнения регрессии.
4. Статистический анализ результатов эксперимента.

7.3. Планирование первого порядка

На первой стадии исследования обычно принимают полином первой степени. Так, для трехфакторной задачи теоретическое уравнение регрессии имеет вид

$$\hat{y} = b_0 + \sum_{i=1}^3 b_i x_i + \sum_{\substack{i,j=1 \\ i \neq j}}^3 b_{ij} x_i x_j + \beta_{123} x_1 x_2 x_3. \quad (7.3.1)$$

Уравнение регрессии, получаемое на основании результатов эксперимента, в отличие от приведенного теоретического уравнения, имеет вид

$$\hat{y} = b_0 + \sum_{i=1}^3 b_i x_i + \sum_{\substack{i,j=1 \\ i \neq j}}^3 b_{ij} x_i x_j + b_{123} x_1 x_2 x_3, \quad (7.3.2)$$

где коэффициенты регрессии $b_0, b_1, \dots, b_3, \dots, b_{123}$ являются оценками для теоретических коэффициентов регрессии, т.е.

$$b_i \rightarrow \beta_i, b_{ij} \rightarrow \beta_{ij}, b_{123} \rightarrow \beta_{123}.$$

Члены, содержащие произведения $x_1 x_2; x_2 x_3$ и т.д., называют членами, отражающими попарное взаимодействие факторов.

7.3.1. 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)

$$X_i = \frac{x_i - x_{i0}}{\Delta x_i}. \quad (7.3.1.1)$$

В теории планирования экспериментов показано, что минимально необходимое число уровней факторов на единицу больше порядка уравнения.

7.3.2. Планирование эксперимента

Рассмотрим сначала частный случай, когда функция отклика линейно зависит от трех независимых факторов. Уравнение регрессии в этом случае имеет вид (7.3.2), а план эксперимента представлен в таблице 7.3.2.1.

Таблица 7.3.2.1

Таблица полного факторного эксперимента для трех факторов

Номер опыта	План								Результат y_i
	x_0	x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3	$x_1x_2x_3$	
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 , необходимый 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. ± 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. 7.3.2.1).

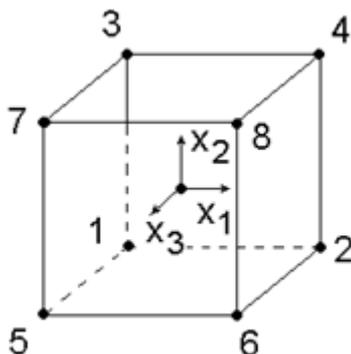


Рис. 7.3.2.1. Геометрическое изображение ПФЭ

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 таблице 7.3.2.1

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):

$$\sum_{j=1}^n X_{ij} = 0, \quad (7.3.2.1)$$

где i — номер фактора; j — номер опыта.

2. Свойство нормирования: сумма квадратов элементов каждого столбца равна числу опытов:

$$\sum_{j=1}^n X_{ij}^2 = n. \quad (7.3.2.2)$$

3. Свойство ортогональности: скалярное произведение всех вектор-столбцов (сумма почленных произведений любых двух вектор-столбцов матрицы) равно нулю:

$$\sum_{j=1}^n X_{ij}X_{uj} = 0, i \neq u. \quad (7.3.2.3)$$

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 $m^* = 2 \div 4$), 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.

7.3.3. Определение коэффициентов уравнения регрессии

Воспользуемся свойствами ПФЭ для определения коэффициентов уравнения регрессии методом наименьших квадратов $\hat{y} = b_0 + b_1x_1 + b_2x_2$.

$$\Phi = \sum_{j=1}^n (y_j - \hat{y}_j)^2 \rightarrow \min_{b_i};$$

$$\frac{\partial \Phi}{\partial b_1} = 2 \sum_{j=1}^n (y_j - b_0 - b_1X_{1j} - b_2X_{2j}) X_{1j} = 0; \quad (7.3.3.1)$$

$$\sum_{j=1}^n y_j X_{1j} - b_0 \sum_{j=1}^n X_{1j} - b_1 \sum_{j=1}^n X_{1j}^2 - b_2 \sum_{j=1}^n X_{1j}X_{2j} = 0.$$

Воспользуемся свойствами ПФЭ:

1. (симметричности) $b_0 \sum X_{1j} = 0$.

2. (нормирования) $b_1 \sum X_{1j}^2 = nb_1$.

3. (ортогональности) $b_2 \sum X_{1j}X_{2j} = 0$.

$$b_1 = \frac{\sum_{j=1}^n y_j X_{1j}}{n}; \quad b_2 = \frac{\sum_{j=1}^n y_j X_{2j}}{n}; \quad b_0 = \frac{\sum_{j=1}^n y_j X_{0j}}{n}. \quad (7.3.3.2)$$

Следовательно, любые коэффициенты уравнения регрессии определяются скалярным произведением столбца y на соответствующий столбец X .

Можно показать, что аналогичным образом определяются и коэффициенты, если в уравнении регрессии (7.2.4) учитываются линейные взаимодействия (двойные, тройные)

$$b_{12} = \frac{\sum_{j=1}^n y_j (X_1 X_2)_j}{n}; \quad b_{123} = \frac{\sum_{j=1}^n y_j (X_1 X_2 X_3)_j}{n}. \quad (7.3.3.3)$$

Следует обратить особое внимание на то, что все линейные коэффициенты независимы, так как в формулы для их расчета (7.3.3.2), (7.3.3.3) 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.

7.3.4. 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 5.3.1 и 5.3.2., 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:

$$S_j^2 = \frac{\sum_{i=1}^{m^*} (y_{ji} - \bar{y}_j)^2}{m^* - 1}, \quad (7.3.4.1)$$

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

$$S_{\text{восп}}^2 = \frac{\sum_{j=1}^n S_j^2}{n} = \frac{\sum_{j=1}^n \sum_{i=1}^{m^*} (y_{ji} - \bar{y}_j)^2}{n(m^* - 1)}, \quad (7.3.4.2)$$

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. 4). Для оценки значимости коэффициентов прежде всего находят дисперсию коэффициентов регрессии. Учитывая свойства 1-3 плана таблицы 7.3.2.1 из выражений (5.3.2.2а) и (5.3.2.2б) при одинаковом дублировании опытов по точкам с числом повторных опытов m^* она определяется по формуле:

$$S_b^2 = \frac{S_{\text{восп}}^2}{m^* n}, \quad (7.3.4.3)$$

а при отсутствии дублирования по соотношению

$$S_b^2 = \frac{S_{\text{восп}}^2}{n}. \quad (7.3.4.3a)$$

Следовательно, все коэффициенты уравнения регрессии ПФЭ имеют одинаковую точность (дисперсию). Это принципиальное отличие коэффициентов уравнения регрессии, полученных по плану таблицы 7.3.2.1, от коэффициентов уравнений, полученных пассивным экспериментом (см. параграф 5.3.2). Планы, по результатам которых коэффициенты уравнения регрессии определяются с одинаковой дисперсией, называются ротатабельными. В связи с этим план, представленный в таблице 7.3.2.1, является не только ортогональным, но ротатабельным. В дальнейшем проверка значимости каждого коэффициента производится независимо с использованием t -критерия Стьюдента (см. гл. 5). 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 \hat{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

$$S_{\text{ад}}^2 = \frac{m^* \sum_{j=1}^n (\bar{y}_j - \hat{y}_j)^2}{n - l}. \quad (7.3.4.4)$$

Здесь n — число опытов (вариантов); $l = k + 1$, где k — число членов в уравнении регрессии.

Проверка адекватности состоит в выяснении соотношения между дисперсией адекватности $S_{\text{ад}}^2$ и дисперсии воспроизводимости $S_{\text{восп}}^2$ и проводится с помощью F -критерия Фишера, который в данном случае рассчитывается как

$$F_{\text{эксп}} = \frac{S_{\text{ад}}^2}{S_{\text{восп}}^2}. \quad (7.3.4.5)$$

Если вычисленное значение критерия меньше теоретического F_{α, m_1, m_2} для соответствующих степеней свободы $m_1 = n(m^* - 1)$, $m_2 = n - l$ при заданном уровне значимости α , 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.

7.3.5. 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 (7.3.2) достаточно рассмотреть уравнение вида

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3, \quad (7.3.5.1)$$

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 (7.3.5.1) 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 таблице 7.3.5.1.

Таблица 7.3.5.1

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. ± 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 приравнение произведения факторов к 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.6.3.5.1, называется полурепликой 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 полурепликой от ПФЭ 2^4 , а для пятифакторного планирования — четвертьрепликой от 2^5 .

Для того чтобы дробная реплика представляла собой ортогональный план, в качестве реплики 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 β_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

$$b_1 = \beta_1 + \beta_{12}; \quad b_2 = \beta_2 + \beta_{13}; \quad b_3 = \beta_3 + \beta_{23}, \quad (7.3.5.2)$$

где β — 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

$$X_3^2 = X_1X_2X_3 = 1, \text{ т. е. } X_1X_2X_3 = 1. \quad (7.3.5.3)$$

Это произведение носит название определяющего контраста.

Умножив поочередно определяющий контраст на x_1, x_2, x_3 , находим

$$X_1 = X_1^2X_2X_3 = X_2X_3; \quad X_2 = X_1X_3, \quad X_3 = X_1X_2. \quad (7.3.5.4)$$

Полученным соотношениям соответствует система смешанных оценок, т.е. β_1 смешана с β_{23} , β_2 — с β_{13} , а β_3 — с β_{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: $b_2 \rightarrow \beta_2 + b_{13}; \quad b_1 \rightarrow \beta_1 + b_{23}$.

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 таблице 7.3.5.2.

В первом случае определяющий контраст $X_4^2 = X_1X_2X_3X_4$ и получим систему 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 $X_4 = X_1X_2X_3$.

Таблица 7.3.5.2

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

Во втором случае определяющий контраст выражается соотношением $X_4^2 = X_1X_2X_3$; $X_1X_2X_3 = 1$.

При этом получим следующую систему оценок.

$$\begin{aligned}
 X_1 &= X_2X_3; & b_1 &\rightarrow \beta_1 + \beta_{234}; \\
 X_2 &= X_1X_3; & b_2 &\rightarrow \beta_2 + \beta_{134}; \\
 X_3 &= X_1X_2X_3; & b_3 &\rightarrow \beta_3 + \beta_{1234}; \\
 X_4 &= X_1X_2; & b_4 &\rightarrow \beta_4 + \beta_{12}; \\
 X_1X_2 &= X_2X_3X_4; & b_{12} &\rightarrow \beta_{12} + \beta_{234}; \\
 X_2X_3 &= X_1X_3X_4; & b_{23} &\rightarrow \beta_{23} + \beta_{134}; \\
 X_3X_4 &= X_1X_2X_3; & b_{34} &\rightarrow \beta_{34} + \beta_{123}.
 \end{aligned}$$

Следовательно, дробную реплику с генерирующим соотношением $X_4 = X_1X_2$ имеет смысл использовать, если нас более всего интересуют коэффициенты β_{12} , β_{23} , β_{34} .

Дробную реплику, в которой Р 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.

7.3.6. Разработка математической модели гидравлического режима методической печи

В качестве примера рассмотрим разработку математической модели гидравлического режима четырехзонной методической печи с использованием теории планирования эксперимента. При планировании опытов использовалась методика проведения дробного факторного эксперимента (ДФЭ) первого порядка с двухуровневым варьированием факторов.

Перед разработкой плана эксперимента на основе априорной информации были выявлены факторы, влияющие на величину давления в томильной зоне печи. К числу таких факторов относятся: расходы топлива на каждую зону нагрева и угол поворота дымового клапана. Расходы воздуха на каждую зону в качестве факторов не фигурировали, поскольку схема управления горением топлива автоматически меняет расход воздуха при изменении расхода газа.

Обозначим факторы: X_1 — расход газа на томильную зону, м³/ч; X_2 — расход газа во вторую сварочную зону, м³/ч; X_3 — расход газа в первую сварочную зону, м³/ч; X_4 — расход газа в нижнюю сварочную зону, м³/ч; X_5 — положение дымового клапана, % хода ИМ (рис. 7.3.6.1).

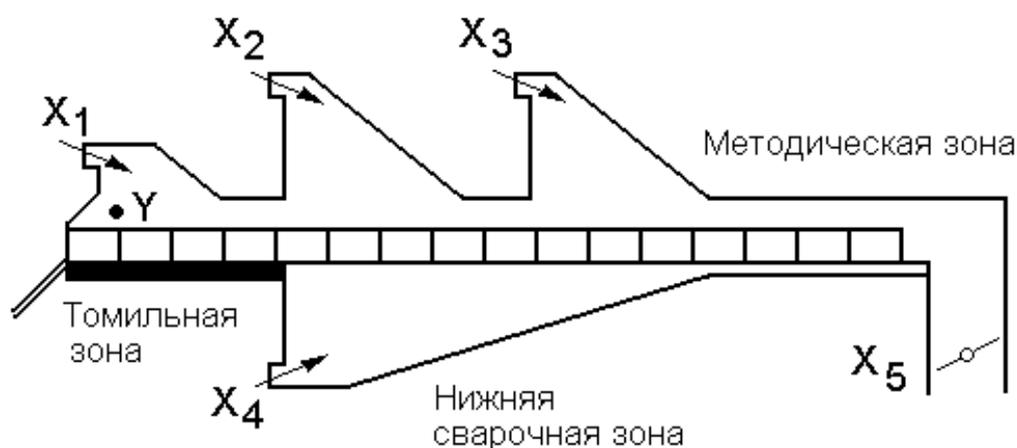


Рис. 7.3.6.1 Положение факторов (X_1, \dots, X_5) и переменной состояния (Y) при проведении исследования на методической печи

Реализация ПФЭ в этом случае при варьировании всех факторов на двух уровнях потребовала бы постановки $2^5 = 32$ опыта.

Будем предполагать, что эффекты взаимодействия факторов в исследуемом объекте маловероятны и пренебрежительно малы. Воспользуемся 1/4 репликой ПФЭ, т.е. ДФЭ типа 25-2, где формально 2 фактора заменены соответствующими произведениями остальных факторов. Это позволит сократить число опытов до $2^3 = 8$. Уровни варьирования факторов представлены в таблице 7.3.6.1.

Уровни варьирования факторов

Уровни факторов	Факторы				
	X_1 , м ³ /ч	X_2 , м ³ /ч	X_3 , м ³ /ч	X_4 , м ³ /ч	X_5 , % хода ИМ
Основной (нулевой)	5250	3900	2650	1100	74
Нижний	4000	3100	1750	700	50
Верхний	6500	4700	3500	1500	98
Интервал варьирования	1250	800	900	400	24

В таблице 7.3.6.2 приведены: матрица планирования ДФЭ 2^3 и результаты эксперимента — значения выходных переменных (давлений в томительной зоне методической печи).

Для обработки результатов эксперимента используем методику, изложенную ранее в параграфе 5.3:

1. Расчет построчных средних:

$$\bar{y}_j = \frac{y_{j1} + y_{j2} + \dots + y_{jm^*}}{m^*},$$

где m^* — число повторных опытов ($m^* = 12$). Например,

$$\bar{y}_1 = \frac{(-0,6) + (-0,5)}{2} = -0,55.$$

Результаты расчета представлены в таблице 7.3.6.2.

Матрица ДФЭ 2^3 с двумя параллельными опытами

Факторы (кодированные значения)						Переменная состояния			
						Опыт 1	Опыт 2	Среднее	Модель
X_0	X_1	X_2	X_3	X_4	X_5	y_1	y_2	\bar{y}_i	\hat{y}_i
1	1	1	1	1	1	-0,6	-0,5	-0,55	-0,3750
1	-1	1	1	-1	-1	0,1	0,5	0,30	0,2625
1	1	-1	1	-1	-1	0,6	0,4	0,50	0,5375
1	-1	-1	1	1	1	-0,1	0,2	0,05	-0,1250
1	1	1	-1	1	-1	0,6	0,2	0,40	0,2250
1	-1	1	-1	-1	1	-0,2	-0,2	-0,20	-0,1625
1	1	-1	-1	-1	1	0,1	0,2	0,15	0,1125
1	-1	-1	-1	1	-1	0,3	0,3	0,30	0,4750

2. Определение построчных (выборочных) дисперсий:

$$S_j^2 = \frac{\sum_{i=1}^{m^*} (y_{ji} - \bar{y}_j)^2}{m^* - 1}; S_1^2 = \frac{(-0,6 - (-0,55))^2 + (-0,5 - (-0,55))^2}{2 - 1} = 0,005.$$

Аналогично, $S_2^2 = 0,08$; $S_3^2 = 0,02$; $S_4^2 = 0,045$; $S_5^2 = 0,08$; $S_6^2 = 0$; $S_7^2 = 0,005$; $S_8^2 = 0$. Сумма построчных (выборочных) дисперсий

$$S_{\Sigma}^2 = 0,005 + 0,08 + 0,02 + 0,045 + 0,08 + 0,005 = 0,235.$$

3. Определение однородности дисперсий по критерию Кохрена

$$G_{\text{эксп}} = \frac{S_{j\text{max}}^2}{S_{\Sigma}^2} = \frac{0,08}{0,235} = 0,34.$$

Далее по справочнику находим $G_{\alpha, m, n}$. Для $\alpha = 0,05$, $m = m^* - 1 = 2 - 1$ и $n = 8$ значение $G_{0,05; 1; 8} = 0,6798$. Поскольку $G_{\text{эксп}} < G_{\text{теор}}$, то дисперсии однородны.

4. Определение коэффициентов в уравнении регрессии

$$b_0 = \frac{\sum_{j=1}^n \bar{y}_j \cdot x_{0j}}{n} = \frac{-0,55 + 0,3 + 0,5 + 0,05 + 0,4 - 0,2 + 0,15 + 0,3}{8} = 0,119;$$

$$b_1 = \frac{\sum_{j=1}^n \bar{y}_j \cdot x_{1j}}{n} = \frac{-0,55 - 0,3 + 0,5 - 0,05 + 0,4 + 0,2 + 0,15 - 0,3}{8} = 0,006;$$

$$b_2 = \frac{\sum_{j=1}^n \bar{y}_j \cdot x_{2j}}{n} = \frac{-0,55 + 0,3 - 0,5 - 0,05 + 0,4 - 0,2 - 0,15 - 0,3}{8} = -0,131;$$

$$b_3 = \frac{\sum_{j=1}^n \bar{y}_j \cdot x_{3j}}{n} = \frac{-0,55 + 0,3 + 0,5 + 0,05 - 0,4 + 0,2 - 0,15 - 0,3}{8} = -0,044;$$

$$b_4 = \frac{\sum_{j=1}^n \bar{y}_j \cdot x_{4j}}{n} = \frac{-0,55 - 0,3 - 0,5 + 0,05 + 0,4 + 0,2 - 0,15 + 0,3}{8} = -0,069$$

$$b_5 = \frac{\sum_{j=1}^n \bar{y}_j \cdot x_{5j}}{n} = \frac{-0,55 - 0,3 - 0,5 + 0,05 - 0,4 - 0,2 + 0,15 - 0,3}{8} = -0,256.$$

5. Проверка значимости коэффициентов регрессии. Предварительно определим дисперсию воспроизводимости (дисперсию отклика)

$$S_{\text{восп}}^2 = \frac{\sum_{j=1}^n S_j^2}{n} = \frac{S_{\Sigma}^2}{n} = \frac{0,235}{8} = 0,029.$$

Дисперсия коэффициентов уравнения регрессии

$$S_b^2 = \frac{S_{\text{восп}}^2}{n \cdot m^*} = \frac{0,029}{8 \cdot 2} = 0,0018; S_b = \sqrt{S_b^2} = 0,0429.$$

Находим значение доверительного интервала для коэффициентов регрессии

$$\Delta b_j = t_{\alpha; m} \cdot S_b.$$

Здесь $m = n(m^* - 1) = 8(2 - 1) = 8$, тогда теоретическое значение критерия Стьюдента $t_{0,05; 8} = 2,31$, откуда $\Delta b_j = 2,31 \cdot 0,0429 = 0,099$. Из сопоставления доверительного интервала Δb_j с абсолютными значениями коэффициентов модели следует, что $|b_1| = 0,006 < 0,099$; $|b_3| = 0,044 < 0,099$ и $|b_4| = 0,069 < 0,099$. Эти коэффициенты оказались незначимы, а остальные значимы. Таким образом, окончательное уравнение регрессии запишется в виде

$$\hat{y} = 0,119 - 0,131X_2 - 0,256X_5.$$

Результаты расчета выходных параметров по уравнению полученной модели \hat{y}_i занесены в табл. 7.3.6.2.

6. Проверка адекватности полученной модели. Предварительно определим дисперсию адекватности:

$$S_{\text{ад}}^2 = \frac{m^* \sum_{i=1}^n (\bar{y}_i - \hat{y}_i)^2}{n - 1}.$$

В нашем случае $m^* = 2$; $n = 8$; $l = k + 1 = 5 + 1 = 6$ и в результате имеем:

$$S_{ад}^2 = \frac{2}{8-6} [(-0,55 + 0,375)^2 + (0,3 - 0,2625)^2 + (0,5 - 0,5375)^2 + (0,05 + 0,125)^2 + (0,4 - 0,225)^2 + (-0,2 + 0,1625)^2 + (0,15 - 0,1125)^2 + (0,3 - 0,475)^2] = 0,1281.$$

С учетом ранее найденной выборочной дисперсии $S_{\Sigma}^2 = 0,235$ определяем дисперсию воспроизводимости

$$S_{восп}^2 = \frac{S_{\Sigma}^2}{n} = \frac{0,235}{8} = 0,0294.$$

Экспериментальное значение критерия Фишера будет равно

$$F_{эксп} = \frac{S_{ад}^2}{S_{восп}^2} = \frac{0,1281}{0,0294} = 4,3571.$$

Теоретическое значение критерия Фишера F_{α, m_1, m_2} при $\alpha = 0,05$ определим по справочнику. Для $m_1 = m * (n - 1) = 2(8 - 1) = 14$ и $m_2 = n - (k + 1) = 8 - (5 + 1) = 2$ значение $F_{0,05;14;2} = 19,4243$. Поскольку $F_{эксп} < F_{теор}$, то полученная модель адекватна.

7.4. 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.

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

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + b_{11}x_1^2 + b_{22}x_2^2 + b_{12}x_1x_2.$$

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. 7.3.2.1, 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. располагались 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^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 α - 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 α 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 таблице 7.4.1.

Таблица 7.4.1

Композиционный план второго порядка

Номер опыта	Факторы						Результат y_i	
	x_0	x_1	x_2	x_1x_2	x_1^2	x_2^2		
Ядро плана	1	+1	-1	-1	+1	+1	+1	y_1
	2	+1	+1	-1	-1	+1	+1	y_2
	3	+1	-1	+1	-1	+1	+1	y_3
	4	+1	+1	+1	+1	+1	+1	y_4
Звездные точки	5	+1	$+\alpha$	0	0	α^2	0	y_5
	6	+1	$-\alpha$	0	0	α^2	0	y_6
	7	+1	0	$+\alpha$	0	0	α^2	y_7
	8	+1	0	$-\alpha$	0	0	α^2	y_8
Центр плана	9	+1	0	0	0	0	0	y_9

Аналогичным образом составляются планы и для большего числа факторов.

7.4.1. Ортогональные планы второго порядка

В общем виде данная матрица неортогональна, т.к.

$$\sum_{j=1}^n x_{0j}x_{ii}^2 \neq 0; \sum_{j=1}^n x_{ij}^2x_{ii}^2 \neq 0. \quad (7.4.1.1)$$

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

$$\hat{y} = b'_0 + \sum_{i=1}^k b_i x_i + \sum_{i,j=1}^k b_{ij} x_{ij} + \sum_{i,j=1}^k b'_{ii} x'_{ii}.$$

Композиционные планы легко привести к ортогональным, выбирая звездное плечо α . В таблице 7.4.1.1 приведено значение α для различного числа факторов k и числа опытов в центре плана n_0 .

Таблица 7.4.1.1

Таблица значений звездного плеча α при различном числе факторов k

Число опытов в центре плана, n_0	Звездное плечо α 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 таблице 7.4.1.2, а его геометрическая интерпретация на рис. 7.4.1.1а.

Таблица 7.4.1.2

Ортогональный план второго порядка

Номер опыта	Факторы						Результат y_i	
	x_0	x_1	x_2	x_1x_2	x'_1	x'_2		
Ядро плана	1	+1	-1	-1	+1	+1/3	+1/3	y_1
	2	+1	+1	-1	-1	+1/3	+1/3	y_2
	3	+1	-1	+1	-1	+1/3	+1/3	y_3
	4	+1	+1	+1	+1	+1/3	+1/3	y_4
Звездные точки	5	+1	$\alpha = +1$	0	0	+1/3	-2/3	y_5
	6	+1	$\alpha = -1$	0	0	+1/3	-2/3	y_6
	7	+1	0	$\alpha = +1$	0	-2/3	+1/3	y_7
	8	+1	0	$\alpha = -1$	0	-2/3	+1/3	y_8
Центр плана	9	+1	0	0	0	-2/3	-2/3	y_9

В этой таблице $x'_{uj} = x_{uj}^2 - \frac{\sum_{j=1}^2 x_{uj}^2}{9} = x_{uj}^2 - \frac{2}{3}$.

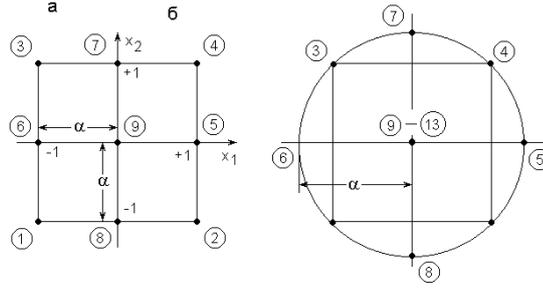


Рис. 7.4.1.1. Планы второго порядка при $k=2$: а — ортогональный; б — ротатабельный

Представленный на рис. 7.4.1.1а и в таблице 7.4.1.2 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.

В силу ортогональности матрицы планирования все коэффициенты уравнения регрессии b определяются независимо один от другого по формуле:

$$b_i = \frac{\sum_{j=1}^n x_{ij} y_j}{\sum_{j=1}^n x_{ij}^2}; b'_{ii} = \frac{\sum_{j=1}^n x'_{ij} y_j}{\sum_{j=1}^n x_{ij}^2}. \quad (7.4.1.2)$$

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

$$S_{b,i}^2 = S_{\text{восп}} / \sum_{j=1}^n x_{ij}^2. \quad (7.4.13)$$

Следует особо отметить, что коэффициенты уравнения регрессии, получаемые с помощью ортогональных планов второго порядка, определяются с разной точностью, в то время как ортогональные планы первого порядка обеспечивают одинаковую точность коэффициентов.

В результате расчетов по матрице с преобразованными столбцами для квадратичных эффектов получим уравнение в виде

$$\hat{y} = b'_0 + \sum_{i=1}^k b_i x_i + \sum_{i,j=1}^k b_{ij} x_i x_j + \sum_{i=1}^k b_{ii} (x_i^2 - \bar{x}_i^2). \quad (7.4.1.4)$$

Для преобразования к обычной записи следует перейти от коэффициента b'_0 к коэффициенту b_0 , используя выражение

$$b_0 = b'_0 - \sum_{i=1}^k b_{ii} \bar{x}_i^2. \quad (7.4.1.5)$$

При этом дисперсия этого коэффициента рассчитывается по следующему соотношению

$$S_{b,0}^2 = S_{b',0}^2 + \sum_{i=1}^k \bar{x}_i^2 S_{b',ii}^2. \quad (7.4.1.6)$$

В дальнейшем, зная дисперсию воспроизводимости, проверяют значимость коэффициентов и адекватность уравнения:

$$\hat{y} = b'_0 + \sum_{i=1}^k b_i x_i + \sum_{i,j=1}^k b_{ij} x_i x_j + \sum_{i=1}^k b_{ii} x_i^2. \quad (7.4.1.7)$$

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 -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.

$$1 = \frac{(n + 2)(n - 1)}{2}.$$

7.4.2. Ротатабельные планы второго порядка

Как мы с Вами установили, план второго порядка, представленный в таблице 7.4.1.1, 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. 7.4.1.1а), 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 α 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 7.4.2.1 приведены значения α и n_0 для различного числа независимых факторов.

Таблица 7.4.2.1

Значения α и n_0 для различного числа независимых факторов

Параметр плана	Значения параметров при числе независимых факторов									
	2	3	4	5	6	6	6	7	7	
Ядро плана	2 ²	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 α на примере матрицы ротатабельного планирования второго порядка для $k = 2$, представленной в таблице 7.4.2.2. Размещение точек этого плана показано на рис. 7.4.1.1б. Для обеспечения ротатабельности точек 5, 6, 7, 8 необходимо удалить их от центра плана на расстояние

α 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 (рис. 7.3.2.1) 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 7.4.2.2

Матрица ротатабельного планирования второго порядка для $k=2$

Номер опыта	Факторы						Результат y_i
	x_0	x_1	x_2	x_1x_2	x_1^2	x_2^2	
Ядро плана 1	+1	-1	-1	+1	+1	+1	y_1
2	+1	+1	-1	-1	+1	+1	y_2
3	+1	-1	+1	-1	+1	+1	y_3
4	+1	+1	+1	+1	+1	+1	y_4
Звездные точки 5	+1	+1,414	0	0	2	0	y_5
6	+1	-1,414	0	0	2	0	y_6
7	+1	0	+1,414	0	0	2	y_7
8	+1	0	-1,414	0	0	2	y_8

Однако матрица ротатабельного планирования второго порядка неортогональна, т.к.

$$\sum_{j=1}^n x_{0j} \cdot x_{uj}^2 \neq 0; \sum_{j=1}^n x_{ij}^2 \cdot x_{uj}^2 = 0; i \neq u. \quad (7.4.2.1)$$

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

$$S_1^2 = \sum_{j=1}^n (y_j - \hat{y}_j)^2 \quad (7.4.2.2)$$

с числом степеней свободы

$$m_1 = n - 1 = n - \frac{(k+2)(k+1)}{2}.$$

По опытам в центре плана определяют сумму квадратов воспроизводимости

$$S_2^2 = \sum_{j=1}^{n_0} (y_{0j} - \hat{y}_{0j})^2 \quad (7.4.2.3)$$

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 -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.

7.5. 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. 7.5.1а,б. Здесь точка А соответствует оптимальным значениям факторов x_1^* и x_2^* , обеспечивающим максимум функции отклика y_{\max} . Замкнутые линии на рис. 7.5.1б 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. 7.5.2, требуется определить при заданных ее конструктивных и режимных параметрах расположение фурмы по высоте печи h , ее диаметр d и высоту 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.

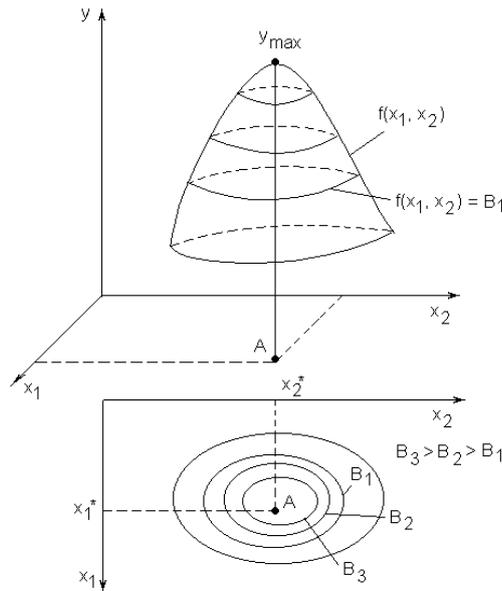


Рис. 7.5.1. Поверхность отклика (а) и линии равного уровня (б): $y = f(x_1, x_2) = B = const$ для $n = 2$

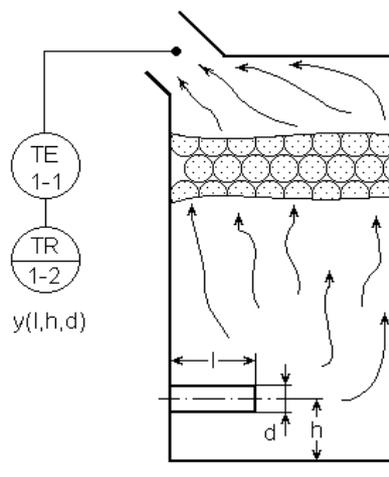


Рис. 7.5.2. Схема шахтной печи

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.

7.5.1. Метод покоординатной оптимизации (Гаусса — Зейделя)

Процесс поиска оптимума методом покоординатной оптимизации (Гаусса-Зейделя) в графическом виде для двумерного случая представлен на рис. 7.5.1.1. 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.

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. 7.5.1.1 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^* .

В дальнейшем рассмотрим более совершенные методы.

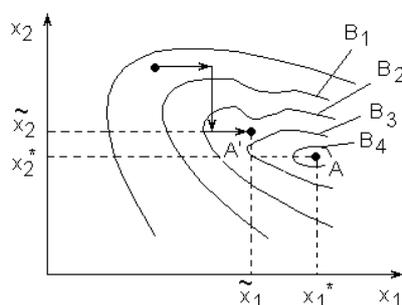


Рис. 7.5.1.1. Графическая иллюстрация метода Гаусса-Зейделя

7.5.2. 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).

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.

$$\hat{y} = b_0 + b_1x_1 + b_2x_2.$$

Градиент функции отклика в этой точке определяется как

$$\text{grad } y = \frac{\partial y}{\partial x_1} \cdot \vec{i} + \frac{\partial y}{\partial x_2} \cdot \vec{j},$$

где \vec{i}, \vec{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. В точке последнего находят новое направление градиента, осуществляя опять же ПФЭ, и далее процедура повторяется.

Практически алгоритм сводится к следующей последовательности операций:

1. Планирование и постановка ПФЭ (или ДФЭ) в окрестности точки начального состояния x_{i0} . Расчет коэффициентов b_i линейной математической модели с целью определения направления градиента.

2. Расчет произведений $b_i \Delta x_i$, где Δx_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

$$x_{ik} = x_{i0} + kh_i, k = 1, 2, \dots$$

находят координаты опытов 5, 6, 7, 8, 9. Часть этих опытов пролагают "мысленными". "Мысленный" опыт заключается в получении предсказанных (расчетных) значений функции отклика \hat{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.

7. В окрестности локального экстремума ставят новую серию опытов (ПФЭ или ДФЭ) для определения новых значений коэффициентов уравнения регрессии и нового направления градиента. 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.

7.5.3. 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. 7.5.3.1 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 вначале, а в дальнейшем каждый шаг сопровождается проведением только одного дополнительного опыта, условия которого выбираются на основе предшествующих результатов.

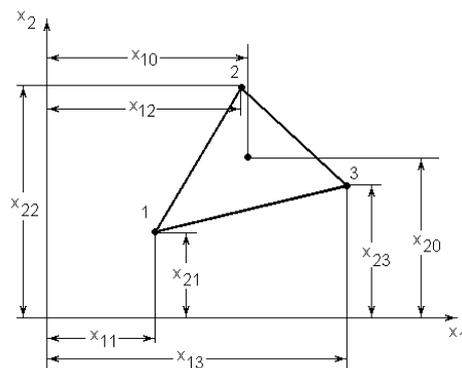


Рис. 7.5.3.1. Схема построения начального симплекса

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:

$$x_{ij} = x_{i0} + C_{ij}\Delta x_i, \quad (7.5.3.1)$$

где x_{i0} — координаты центра начального симплекса; Δx_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 таблице 7.5.3.1.

Таблица 7.5.3.1

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:

$$k_i = \frac{1}{i+1} \sqrt{\frac{i+1}{2i}} = \sqrt{\frac{1}{2i(i+1)}}; R_i = \sqrt{\frac{i}{2(i+1)}}; i = 1, 2, \dots, n; n — \text{число факторов}$$

Для определения условий проведения опыта в отраженной точке (координат новой вершины симплекса) используется формула:

$$x_{i'n} = \frac{2}{n} \sum_{j=1}^n x_{ij} - x_{i'n}, i \neq j, \quad (7.5.3.2)$$

где $x_{i'n}$ — координата новой точки (новой вершины) симплекса для i -й переменной; $x_{i'n}$ — 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 $\hat{y} = f(x_1, \dots, x_n) = const$ 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 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.

Контрольные вопросы.

1. С какой целью используют теорию планирования эксперимента?
2. Из каких соображений выбирают основные факторы, их уровни, а также интервалы варьирования факторов при проведении ПФЭ и ДФЭ?
3. В чем заключается основная идея ДФЭ?
4. В чем заключаются причины неадекватности математической модели? Как производится оценка адекватности?
5. Каковы принципы ротатабельного планирования эксперимента?
6. С какой целью композиционные планы приводят к ортогональному виду?
7. В чем заключается сущность планирования экспериментов при поиске оптимальных условий? Какие методы при этом используют?
8. На чем основан метод покоординатной оптимизации?
9. Из каких этапов состоит алгоритм оптимизации методом крутого восхождения?
10. В чем заключаются основная идея метода симплексного планирования?
11. Какие преимущества дает экспериментатору использование средств вычислительной техники?
12. Каковы возможности современных программ по обработке экспериментальных данных?
13. На каких принципах основана организация современных статистических пакетов?

ОСНОВНАЯ ЛИТЕРАТУРА

1. Планирование экстремальных экспериментов : методическое пособие по курсу "Планирование эксперимента" /Том. гос. ун-т ; сост. А. Н. Голованов
2. Планирование эксперимента : учебное пособие : [для студентов по специальности "Механика", для аспирантов] /А. Н. Голованов ; Том. гос. ун-т
<http://vital.lib.tsu.ru/vital/access/manager/Repository/vtls:000427154>
3. Планирование эксперимента /Г. И. Красовский, Г. Ф. Филаретов // Минск : Издательство БГУ им. В. И. Ленина , 1982
4. Статистика и планирование эксперимента в технике и науке : Методы планирования эксперимента /Н. Джонсон; Ред. Э. К. Лецкого // М. : Мир , 1981

ДОПОЛНИТЕЛЬНАЯ ЛИТЕРАТУРА

1. Порсев, Е.Г. Организация и планирование экспериментов : учебное пособие / Е.Г. Порсев. - Новосибирск : НГТУ, 2010. - 155 с. - Доступ из ЭБС «Университетская библиотека ONLINE».
2. Ли Р.И. Основы научных исследований : учебное пособие/ Ли Р.И. – Липецк : Липецкий государственный технический университет, ЭБС АСВ, 2013. — 190 с. — Доступ из ЭБС «IPRbooks». Режим доступа: <http://www.iprbookshop.ru/22903>.
3. ГОСТ Р 8.736–2011 «Государственная система обеспечения единства измерений. Измерения прямые многократные. Методы обработки результатов измерений. Основные положения». Введен впервые, введен 2013–01–01. – М.: Стандартинформ, 2013. — 25 с.
4. Метрология. Стандартизация. Сертификация : учебник для студентов вузов/А. В. Архипов [и др.]; под ред. В. М. Мишина. — М.: ЮНИТИ-ДАНА, 2009. – 495 с.
5. ГОСТ 5725–2002 Точность (правильность и прецизионность) методов и результатов измерений. — Введен 2002–04–23. — М.: Изд-во стандартов, 2002. — 108 с.

ЭЛЕКТРОННЫЕ РЕСУРСЫ

1. Берикашвили В. Ш., Оськин С. П.-СТАТИСТИЧЕСКАЯ ОБРАБОТКА ДАННЫХ, ПЛАНИРОВАНИЕ ЭКСПЕРИМЕНТА И СЛУЧАЙНЫЕ ПРОЦЕССЫ 2-е изд., испр. и доп. Учебное пособие для бакалавриата и магистратуры-М.:Издательство Юрайт,2019-164-Бакалавр и магистр. Академический курс-978-5-534-09216-5: -Текст электронный // ЭБС Юрайт - <https://biblio-online.ru/book/statisticheskaya-obrabotka-dannyh-planirovanie-eksperimenta-i-sluchaynye-processy-427449>
2. Гарькина, И.А. Планирование эксперимента. Обработка опытных данных : практическое пособие / Гарькина И.А., Данилов А.М., Прошин А.П., Соколова Ю.А. — Москва : Палеотип, 2005. — 273 с. — ISBN 5-94727-117-6. — URL: <https://book.ru/book/901182> (дата обращения: 10.10.2019). — Текст : электронный.
3. Григорьев, Ю.Д. Методы оптимального планирования эксперимента: линейные модели : учебное пособие. — СПб. : Лань, 2015. — 320 с. — Доступ из ЭБС «Лань». Режим доступа: http://e.lanbook.com/books/element.php?pl1_id=65949

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. Планирование эксперимента : учебное пособие : [для студентов по специальности "Механика", для аспирантов] /А. Н. Голованов ; Том. гос. ун-т <http://vital.lib.tsu.ru/vital/access/manager/Repository/vtls:000427154>
2. Сафин, Р.Г. Основы научных исследований. Организация и планирование эксперимента : учебное пособие / Р.Г. Сафин, А.И. Иванов, Н.Ф. Тимербаев. - Казань : Издательство КНИТУ, 2013. - 154 с. : ил., табл., схем.
3. Порсев, Е.Г. Организация и планирование экспериментов : учебное пособие / Е.Г. Порсев. - Новосибирск : НГТУ, 2010. - 155 с.

Учебное издание

Касымов Денис Петрович

Планирование и обработка результатов эксперимента

Учебно-методическое пособие

Издание вышло в свет в авторской редакции

Подписано к печати 22.12.2021 г. Формат 60×84¹/₈.

Бумага для офисной техники.

Печ. л. 11,75.

Тираж 50 экз. Заказ № 4874.

Отпечатано на оборудовании
Издательства Томского государственного университета
634050, г. Томск, пр. Ленина, 36
Тел. 8+(382-2)–52-98-49
Сайт: <http://publish.tsu.ru>
E-mail: rio.tsu@mail.ru