

Лекция 1

Планирование эксперимента и обработка результатов

Литература

1. Планирование эксперимента в исследовании технологических процессов / Под ред. Э.К.Лецкого. - М.: Мир, 1977.
2. Математическая статистика с элементами теории планирования эксперимента: учеб. пособие: Л.В. Горская, В.Н. Пиунова, В.С. Смирнова. Саратовский политех. институт, 1975, 103 с.
3. Математическая теория планирования эксперимента./ Под ред. С.М. Ермакова.- М.: Наука. Главная редакция физико-математической литературы, 1983. 392 с.
4. Надёжность технических систем: Справочник/ Ю.К. Беляев, В.А. Богатырёв и др. Под ред. И.А. Ушакова М.: Радио и связь, 1985. 608 с.
5. Дьяконов В.П. MATLAB 6/6.1/6.5 +Simulink 4/5 в математике и моделировании. Полное руководство пользователя.- М.: СОЛОН-Пресс, 2003. 576 с.
6. ГОСТ 24026-80 « Исследовательские испытания. Планирование эксперимента. Термины и определения. М.,1980.

Дополнительная

1. Исследования и изобретательство в машиностроении. Практикум/ Под общ. Ред. М.М. Кане. Мн.: УП «Технопринт», 2003. 237 с.
2. Адлер Ю.П. Введение в планирование эксперимента. М., 1969. Минько А.А. Статистический анализ в MS Excel. – М.: Изд. дом «Вильямс», 2004. – 448 с.
3. Тюрин Ю.Н. Анализ данных на компьютере / Ю.Н.Тюрин, А.А. Макаров. - М.: Финансы и статистика, 1995.

Основные понятия и определения

Эксперимент - система операций, воздействий и (или) наблюдений, направленных на получение информации об объекте при исследовательских испытаниях

Опыт- воспроизведение исследуемого явления в определенных условиях проведения эксперимента при возможности регистрации его результатов

План эксперимента - совокупность данных, определяющих число, условия и порядок реализации опытов

Планирование эксперимента - выбор плана эксперимента, удовлетворяющего заданным требованиям.

Фактор (ндп. *Параметр*) - переменная величина, по предположению влияющая на результаты эксперимента.

Уровень фактора - фиксированное значение фактора относительно начала отсчета

Основной уровень фактора - натуральное значение фактора, соответствующее нулю в безразмерной шкале.

Нормализация факторов - преобразование натуральных значений факторов в безразмерные значения.

Априорное ранжирование факторов - метод выбора наиболее важных факторов, основанный на экспертной оценке

Размах варьирования фактора - разность между максимальным и минимальным натуральными значениями фактора в данном плане

Интервал варьирования фактора - половина размаха варьирования фактора

Эффект взаимодействия факторов - показатель зависимости изменения эффекта одного фактора от уровней других факторов.

Факторное пространство - пространство, координатные оси которого соответствуют значениям факторов.

Область экспериментирования (область планирования) - область факторного пространства, где могут размещаться точки, отвечающие условиям проведения опытов.

Активный эксперимент - эксперимент, в котором уровни факторов в каждом опыте задаются исследователем.

Пассивный эксперимент - эксперимент, при котором уровни факторов в каждом опыте регистрируются исследователем, но не задаются.

Последовательный эксперимент (ндп. *шаговый эксперимент*) - эксперимент, реализуемый в виде серий, в котором условия проведения каждой последующей серии определяются результатами предыдущих.

Отклик (ндп. *Реакция на Параметр*) - наблюдаемая случайная переменная, по предположению, зависящая от факторов.

Функция отклика - зависимость математического ожидания отклика от факторов.

Оценка функции отклика - зависимость, получаемая при подстановке в функцию отклика оценок значений ее параметров.

Дисперсия оценки функции отклика - дисперсия оценки математического ожидания отклика в некоторой данной точке факторного пространства.

Поверхность отклика (ндп. *Поверхность регрессии*) - геометрическое представление функции отклика.

Поверхность уровня функции отклика - геометрическое место точек в факторном пространстве, которому соответствует некоторое фиксированное значение функции отклика.

Область оптимума - область факторного пространства в окрестности точки, в которой функция отклика достигает экстремального значения.

Рандомизация плана - один из приемов планирования эксперимента, имеющий целью свести эффект некоторого неслучайного фактора к случайной ошибке.

Параллельные опыты - рандомизированные во времени опыты, в которых уровни всех факторов сохраняются неизменными.

Временный дрейф - случайное или неслучайное изменение функции отклика во времени.

Исследователь на этапе планирования эксперимента должен:

- помнить, к какому классу относится моделируемая система (статическая или динамическая, детерминированная или стохастическая и т.д.);
- определить, какой режим работы его интересует, стационарный (установившийся) или нестационарный;
- знать, в течение какого промежутка времени следует наблюдать за поведением (функционированием) системы;

- знать, какой объём испытаний (то есть повторных экспериментов) сможет обеспечить требуемую точность оценок (в статистическом смысле) исследуемых характеристик системы.

Таким образом, планирование модельных экспериментов преследует 2 основные цели:

1. Сокращение общего объёма испытаний при соблюдении требований к достоверности и точности их результатов;
2. повышение информативности каждого из экспериментов в отдельности.

Поиск плана эксперимента проводится в **факторном пространстве**.

Факторное пространство (ФП) – это множество внешних и внутренних параметров модели, значения которых исследователь может контролировать в ходе подготовки и проведения эксперимента.

Значения факторов обычно называются **уровнями**.

Если при проведении эксперимента исследователь может изменять уровни факторов, эксперимент называется **активным**, в противном случае – **пассивным**.

Введём ещё несколько терминов, используемых в теории планирования эксперимента.

Каждый из факторов имеет верхний и нижний уровни, расположенные симметрично относительно некоторого нулевого уровня. **Точка в ФП, соответствующая нулевым уровням всех факторов, называется центром плана.**

Интервалом варьирования фактора называется некоторое число J , прибавление которого к нулевому уровню даёт верхний уровень, а вычитание – нижний.

Как правило, план эксперимента строится относительно одного (основного) выходного скалярного параметра Y , который называется **наблюдаемой переменной**.

Если моделирование используется как инструмент принятия решения, то в роли наблюдаемой переменной выступает **показатель эффективности**. При этом предполагается, что значение наблюдаемой переменной, полученное в ходе эксперимента, складывается из двух составляющих:

$$Y = f(x) + e(x),$$

где $f(x)$ - функция отклика (неслучайная функция факторов),

$e(x)$ - ошибка эксперимента (случайная величина).

Дисперсия D_Y наблюдаемой переменной, которая характеризует точность измерений, равна дисперсии ошибки опыта, т.е.

$$D_Y = D_e.$$

D_Y называют **дисперсией воспроизводимости эксперимента**. Она характеризует качество эксперимента. **Эксперимент называется идеальным при $D_Y = 0$.**

Существует два основных варианта постановки задачи планирования имитационного эксперимента:

- Из всех допустимых требуется выбрать такой план, который позволил бы получить наиболее достоверное значение функции отклика $f(x)$ при фиксированном числе опытов;

- *Из всех допустимых требуется выбрать такой план, при котором статистическая оценка функции отклика может быть получена с заданной точностью при минимальном объёме испытаний.*

Решение задачи планирования в первой постановке называется **стратегическим планированием**, во второй – **тактическим планированием**.

При проведении опытных исследований различают **пассивный и активный эксперимент**.

Методология **пассивного экспериментирования** предполагает проведение большой серии опытных исследований с поочередным варьированием значений входных переменных \bar{x} и анализом результатов измерений выходной переменной y (лабораторный эксперимент или эксперимент на пилотной установке).

К **пассивному эксперименту** принято относить также и сбор опытных данных в режиме эксплуатации промышленной установки – т.н. **промышленный эксперимент**.

Обработка результатов, и выбор вида эмпирической модели (уравнения регрессии), т.е. решение задачи структурной идентификации является достаточно сложной задачей.

Это связано с тем, что вид *уравнения регрессии необходимо определять по характеру изменения переменных на графике эмпирической линии регрессии, полученной по выборке экспериментальных данных.*

Для решения этой задачи для одной входной переменной x предложены эффективные методы, в которых предусматривается преобразование системы координат как для входной (x), так и для выходной переменной (y). При большем числе входных переменных (x_1, \dots, x_m) надёжных методов определения вида уравнения регрессии (вида эмпирической модели) в настоящее время не существует.

Активный эксперимент проводится по заранее составленному плану, в соответствии с которым ставится задача *не только определения оптимальных условий проведения эксперимента, но и оптимизации процесса (оптимальное планирование эксперимента).*

При определении оптимальных условий проведения процесса с использованием эмпирических моделей (например, методом Бокса-Вильсона) выходная переменная \hat{y} является критерием оптимальности или целевой функцией.

В теории активного экспериментирования выходную (зависимую) переменную принято называть **функцией отклика**, а входные (независимые) переменные – **факторами**. Соответственно - координатное пространство с координатами (x_1, x_2, \dots, x_m) - **факторным пространством**, а геометрическое изображение функции отклика в факторном пространстве – **поверхностью отклика**.

Активный эксперимент планируется таким образом, чтобы упростить обработку его результатов методами регрессионного и корреляционного анализа.

К достоинствам активного экспериментирования относятся:

- возможность предсказания количества опытов, которые следуют провести;
- определение точек факторного пространства, где следует проводить опыты;
- отсутствие проблем, связанных с выбором вида уравнения регрессии;

- возможность определения оптимальных параметров процесса экспериментально-статистическим методом;
- сокращение объёма опытных исследований.

Основы планирования многофакторного эксперимента

В общем случае объект исследования можно представить в виде структурной схемы, показанной на рис.1.

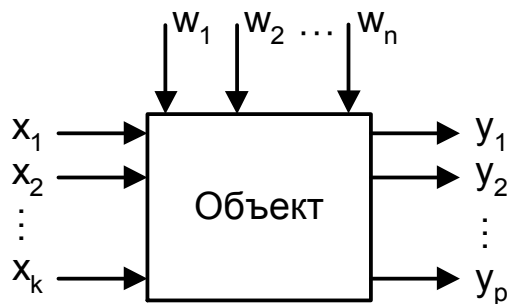


Рис.1

Представление объекта в виде такой схемы основано на принципе «черного ящика». Имеем следующие группы параметров:

- 1) управляющие (входные) x_i которые называются *факторами*;
- 2) выходные параметры y_i , которые называются *параметрами состояния*;
- 3) w_i - *возмущающие воздействия*.

Предполагается, что возмущающие воздействия не поддаются контролю и либо являются случайными, либо меняются во времени.

Каждый фактор x_i имеет область определения, которая должна быть установлена до проведения эксперимента.

Комбинацию факторов можно представить как точку в многомерном пространстве, характеризующую состояние системы.

На практике целью многофакторного эксперимента является установление зависимости

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

описывающей поведение объекта. Чаще всего функция (1) строится в виде полинома

$$y = a_0 + a_1 x_1 + a_2 x_2 \quad (2)$$

или

$$y = a_0 + a_1 x_1 + a_2 x_2 + a_{11} x_1^2 + a_{22} x_2^2 + a_{12} x_1 x_2. \quad (3)$$

Целью эксперимента может быть, например, построение зависимости (1) при минимальном количестве измерений значений управляющих параметров x_i .

На первом этапе планирования эксперимента необходимо выбрать область определения факторов x_i . Выбор этой области производится исходя из априорной информации. Значения x_i называются *уровнями управляющего параметра*.

Если выбрана линейная модель (2), то для построения аппроксимирующей функции достаточно выбрать **основной уровень и интервал варьирования** управляющего параметра x_i .

Для линейной модели **интервал варьирования** можно определить как

$$I = \frac{x_{\max} - x_{\min}}{2},$$

а **основной (нулевой) уровень** - как среднее значение

$$x_0 = \frac{x_{\max} + x_{\min}}{2}.$$

Для упрощения планирования эксперимента принято вместо реальных (натуральных) уровней x_i использовать **кодированные** значения факторов. Для факторов с непрерывной областью определения это можно сделать при помощи следующего преобразования (нормирования факторов)

$$x_j = \frac{\tilde{x}_j - x_{j0}}{I_j},$$

где \tilde{x}_j - натуральное значение фактора; I_j - интервал варьирования; x_{j0} - основной уровень; x_j - кодированное значение. В результате x_j принимает значения на границах $x_j = \pm 1$, на основном уровне $x_j = 0$. Основная проблема состоит в выборе области варьирования, поскольку эта задача является неформализованной.

Рассмотрим **полный факторный эксперимент** на примере линейной модели (2). Если число факторов k , то для проведения полного факторного эксперимента нужно $N = 2^k$ опытов, где 2 - число уровней, которого достаточно для построения линейной модели.

Условие проведения этого эксперимента можно зафиксировать в матрице планирования (табл.1).

Таблица 1

Номер опыта	x_1	x_2	y
1	-1	-1	y_1
2	+1	-1	y_2
3	-1	+1	y_3
4	+1	+1	y_4

Таким образом, для двух факторов построение матрицы планирования элементарно. Для большего числа факторов необходимо разработать *правила построения таких матриц*. Например, при появлении фактора x_3 в табл.1 произойдут следующие

щие изменения (табл.2): при появлении нового столбца каждая комбинация уровней исходной таблицы проявится дважды.

Таблица 2

Номер опыта	x_1	x_2	x_3	y
1	-1	-1	+1	y_1
2	+1	-1	+1	y_2
3	-1	+1	+1	y_3
4	+1	+1	+1	y_4
5	-1	-1	-1	y_5
6	+1	-1	-1	y_6
7	-1	+1	-1	y_7
8	+1	+1	-1	y_8

Это не единственный способ расширения матрицы планирования. Используют также перемножение столбцов, правило чередования знаков.

Очень важны **общие свойства матрицы планирования**:

- **симметричность матрицы** относительно центра эксперимента: $x_i = 0$. Тогда

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

- $\sum_{i=1}^N x_{ij}^2 = N$ - условие нормировки, то есть сумма квадратов элементов каждого столбца равна числу опытов.

Первые два свойства относятся к построению отдельных столбцов матрицы.

- $\sum_{i=1}^N x_{ij} \cdot x_{in} = 0$ - совокупность столбцов имеет следующее свойство, где $j \neq n$.
- **Ротатабельность**. Это означает, что точки (значения факторов) в матрице планирования подбираются так, что точность предсказания выходного параметра должна быть одинакова на равных расстояниях от центра эксперимента (нулевого уровня) и не зависеть от направления.

Планирование эксперимента первого порядка для двух переменных.

План эксперимента первого порядка для двух переменных показан на рис. 2. То есть искомая функция $y = f(x_1, x_2)$ описывается модельно в виде плоскости

$$y = a_0 + a_1 x_1 + a_2 x_2 \quad (4)$$

или гиперboloида

$$y = a_0 + a_1 x_1 + a_2 x_2 + a_3 x_1 x_2. \quad (5)$$

Расположение этой модели в пространстве показано на рис.2 поверхностью, проходящей через точки 1 – 2 – 3 – 4.

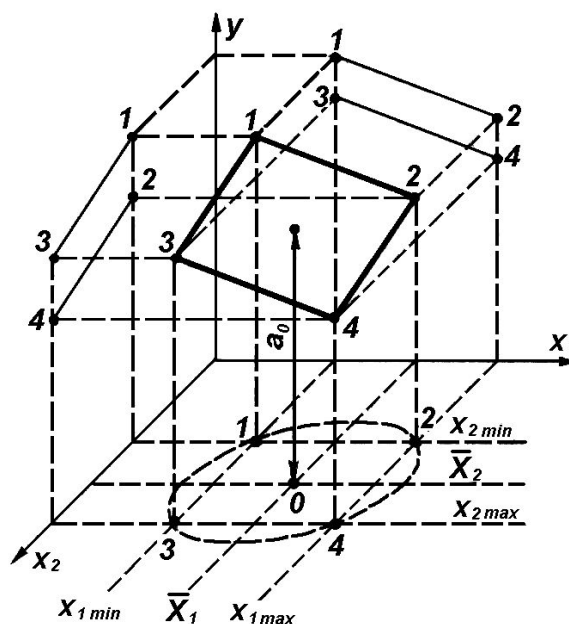


Рис.2

Необходимые уровни для полного факторного эксперимента расположены в плоскости (x_1, x_2) . Для модели в виде гиперboloида этот план является предельно экономным. Для построения гиперboloида необходимо определить четыре коэффициента в модели (5). Это можно сделать, решая систему из четырех уравнений. Следовательно, необходимы все четыре опыта. В теории планирования эксперимента используется термин **насыщенности**.

Если рассматривать модель (4) в виде плоскости, то план эксперимента является **ненасыщенным (избыточным)**, так как необходимо определить только три коэффициента a_0 , a_1 и a_2 . В случае модели (5) (насыщенный эксперимент) решение системы единственно, и поверхность гиперboloида пройдет через все четыре экспериментальных значения y_i . Следствием этого является то, что насыщенный эксперимент не позволяет усреднить случайные погрешности и не дает сведения об их размере.

Для ненасыщенного плана (4) избыточное число опытов позволяет произвести усреднение и оценить размеры погрешности. Проведя плоскость через точки 1, 2 и 3, можно оценить погрешность, определив, на каком расстоянии от плоскости находится точка 4. Погрешность в других точках может быть оценена проведением плоскостей 1 – 3 – 4, 1 – 2 – 4 и 2 – 3 – 4. С другой стороны коэффициент a_1 наклона поверхности к оси x_1 может быть найден как из наклона прямой 1 – 2, так и из наклона прямой 3 – 4. Аналогично коэффициент a_2 при x_2 можно определить из наклона прямых 1 – 3 и 2 – 4.

Поскольку полученные таким образом значения a_1 и a_2 могут отличаться, насыщенный эксперимент позволяет провести их усреднение и оценить погрешность.

Если уравнение плоскости представить в виде

$$y = a_0 + a_1(x_1 - \bar{x}_1) + a_2(x_2 - \bar{x}_2), \quad (6)$$

где $\bar{x}_1 = \frac{x_{1\min} + x_{1\max}}{2}$; $\bar{x}_2 = \frac{x_{2\min} + x_{2\max}}{2}$, то мы переносим начало координат в точку с координатами (\bar{x}_1, \bar{x}_2) . Тогда коэффициент a_0 находится усреднением всех четырех значений y_i как высота центра плоскости 1 – 2 – 3 – 4.

Процесс переноса начало координат в центр пространства факторов с координатами $(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_k)$ очень важен при обработке данных любых экспериментов, описываемых моделью в виде гиперплоскости, так как позволяет получить более устойчивое усредненное значение для a_0 .

Важнейшим фактором является то, что в результате такого усреднения построенная плоскость удовлетворяет всем четырем значениям y_i лишь в среднем. В любой точке может быть найдена погрешность отклонения экспериментальных данных относительно модели, и по этим четырем отклонениям можно вычислить СКО.

Таким образом, один из четырех опытов является избыточным и может быть исключен. Но тогда план эксперимента становится неротатабельным, то есть неравноточным по всем направлениям. Если исключена точка 4 на рис.2, то в направлении 3 – 2 в плоскости факторов будет обеспечена большая точность, чем в направлении 1 – 0. В этом случае для восстановления ротатабельности точки 1, 2 и 3 в плоскости факторов должны быть равноудалены как друг от друга, так и от центра, то есть располагаться в вершинах равностороннего треугольника с центром в точке 0. В общем случае для линейной модели (4), эксперимент содержащий конечное число опытов позволяет получить только оценки для коэффициентов a_0 , a_1 и a_2 .

Подставив в уравнение модели (4) известные значения факторов x_{ij} и результаты опытов y_i получим систему линейных алгебраических уравнений для определения a_i . Если количество этих уравнений больше трех, то значения оценок a_0 , a_1 и a_2 могут быть получены при помощи МНК:

$$a_j = \left(\sum_{i=1}^N x_{ij} y_i \right) / N, \quad (7)$$

где N - количество опытов. Здесь учтено, что x_{ij} принимают значения -1,+1. Для вычисления коэффициентов линейной модели по формуле (7) получим:

$$a_1 = \frac{[(-1)y_1 + (+1)y_2 + (-1)y_3 + (+1)y_4]}{4}, \quad (8)$$

$$a_2 = \frac{[(-1)y_1 + (-1)y_2 + (+1)y_3 + (+1)y_4]}{4}$$

Таким образом, для вычисления a_1 и a_2 можно использовать (8). Для определения a_0 в формуле (6.2.4) найдем среднее значение $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$, равное $\bar{y} = a_0 + a_1 \bar{x}_1 + a_2 \bar{x}_2$, где $\bar{x}_1 = \frac{1}{n} \sum_{i=1}^n x_{1i}$, $\bar{x}_2 = \frac{1}{n} \sum_{i=1}^n x_{2i}$.

В случае симметричности матрицы планирования $\bar{x}_1 = \bar{x}_2 = 0$, откуда $\bar{y} = a_0$. Чтобы коэффициент модели вычислялся по единой формуле (7) в матрице планирования вводят фиктивную переменную x_0 , которая принимает значение 1 во всех опытах и соответствует коэффициенту a_0 . Коэффициент при независимых переменных x_i указывает на силу влияния факторов: чем больше значение имеет коэффициент a_i , тем большее влияние оказывает соответствующий фактор. В этом смысле результаты планирования эксперимента аналогичны факторному анализу. Для пассивных экспериментов факторный анализ может использоваться в качестве априорных данных при планировании.

Планируя эксперимент, стремятся получить линейную модель, однако в выбранных интервалах варьирования априори не известно, что линейная модель адекватно описывает поведение системы.

Нелинейность связана со смешанным взаимодействием. Формула (5) всегда может быть оценена по полному факторному эксперименту. Для полного факторного эксперимента $N = 2^2$ матрица планирования с учетом эффекта взаимодействия приведена в табл.3.

Таблица 3

Номер опыта	x_0	x_1	x_2	$x_1 x_2$	y
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

В этом случае коэффициент a_{12} также может быть вычислен по формуле (9):

$$a_{12} = \frac{[(+1)y_1 + (-1)y_2 + (-1)y_3 + (+1)y_4]}{4} \quad (9)$$

Столбцы x_1, x_2 задают планирование эксперимента – по ним определяют результаты опыта; столбцы x_0, x_1x_2 служат только для расчета.

С ростом числа факторов число возможных взаимодействий возрастает. Например, для факторного эксперимента $N = 2^3$ кроме x_0, x_1, x_2, x_3 в матрице планирования появляются столбцы $x_1x_2, x_1x_3, x_2x_3, x_1x_2x_3$. Всего в матрице планирования оказывается восемь столбцов, следовательно, необходимо определять восемь коэффициентов. Все восемь коэффициентов необходимо определять в том случае, если учитывать смешанное взаимодействие. Если же модель задается в виде гиперплоскости (линейная модель), то достаточно определить четыре коэффициента: x_0, x_1, x_2, x_3 . Полный факторный эксперимент оказывается избыточным и у экспериментатора возникает *выбор*:

1. Построить гиперплоскость по четырем экспериментам, а остальные четыре опыта использовать для оценки погрешности.

2. Провести эксперимент, состоящий из 4-х опытов, то есть реализовать экономный план эксперимента.

Таким образом, в отличие от модели гиперboloида, которая требует определение 2^k неизвестных коэффициентов, модель гиперплоскости, содержит $k+1$ коэффициент и требует соответствующего числа опытов, то есть **полный факторный план (ПФП)** для модели гиперплоскости сильно избыточен.

Для построения гиперплоскости, следовательно, достаточно использовать лишь некоторую часть из ПФП. Эту часть в теории планирования эксперимента называют *дробной репликой* или **дробным факторным планом (ДФП)**. Если дробление ПФП производится последовательным делением числа опытов на 2, то реплику называют *регулярной*. Число P последовательного деления называют *дробностью* реплики.

Число опытов регулярного ДФП равняется $n = 2^{k-p}$. При $p = 1$ ДФП называют полурепликой (или 1/2 реплика), при $p = 2$ – 1/4 реплика и т.д.

Соответствующее число опытов и параметров планирования приведены в таблице 4.

Таблица 4

Число факторов, k	Число коэфф. модели, $k+1$	Число опытов ПФП	Вид плана	Число опытов плана	Избыточность
2	3	4	ПФП	4	1
3	4	8	Полуреплика	4	0
4	5	16	Полуреплика	8	3
5	6	32	Четвертьреплика	8	2
6	7	64	1/8 реплика	8	1

7	8	128	1/16 реплика	8	0
8	9	256	1/16 реплика	16	7

Для составления планов-таблиц регулярных дробных реплик часто используют так называемое **правило двоичного кода**. Оно гласит, что для модели в виде гиперболоида знаки “+” и “–” в столбцах плана должны чередоваться по правилу чередования двоичных чисел в разряде двоичного кода, то есть в столбце x_1 - через 1, в столбце x_2 - через 2, в столбце x_3 - через 4, в столбце x_k - через 2^{k-1} .

Лекция 2

Построение эмпирических моделей по данным активного эксперимента

При проведении опытных исследований различают пассивный и активный эксперимент.

Методология пассивного экспериментирования предполагает проведение большой серии опытных исследований с поочередным варьированием значений входных переменных x и анализом результатов измерений выходной переменной y (лабораторный эксперимент).

К пассивному эксперименту принято относить также и сбор опытных данных в режиме эксплуатации промышленной установки – т.н. промышленный эксперимент.

Обработка результатов пассивного эксперимента проводится методами *регрессионного и корреляционного анализа*, и выбор вида эмпирической модели (уравнения регрессии), т.е. решение задачи структурной идентификации является достаточно сложной задачей.

Это связано с тем, что вид уравнения регрессии необходимо определять по характеру изменения переменных на графике эмпирической линии регрессии, полученной по выборке экспериментальных данных.

Для решения этой задачи для одной входной переменной x предложены эффективные методы, в которых предусматривается преобразование системы координат как для входной (x), так и для выходной переменной (y). При большем числе входных переменных

(x_1, \dots, x_m) надёжных методов определения вида уравнения регрессии (вида эмпирической модели) в настоящее время не существует.

Активный эксперимент проводится по заранее составленному плану, в соответствии с которым ставится задача не только определения оптимальных условий проведения эксперимента, но и оптимизации процесса (оптимальное планирование эксперимента).

При этом **уравнения регрессии** (эмпирические модели) описывают данные активного эксперимента, в основном, в двух ограниченных областях и имеют следующий вид:

- **вдали от экстремального значения выходной переменной y :**

$$\hat{y}' = a_0 + \sum_{j=1}^m a_j x_j + \sum_{j=1}^{m-1} \sum_{\substack{u=2 \\ u > j}}^m a_{ju} x_j x_u$$

- **вблизи** экстремального значения выходной переменной y («в почти стационарной области»):

$$\hat{y}'' = a_0 + \sum_{j=1}^m a_j x_j + \sum_{j=1}^{m-1} \sum_{\substack{u=2 \\ u > j}}^m a_{ju} x_j x_u + \sum_{j=1}^m a_j x_j^2.$$

Приведённые уравнения являются линейными относительно коэффициентов регрессии \bar{a} и имеют достаточно простой вид. Они включают слагаемые с двойным взаимодействием входных переменных

$$\left(\sum_{j=1}^{m-1} \sum_{\substack{u=2 \\ u > j}}^m a_{ju} x_j x_u \right)$$

и не учитывают взаимодействия более высоких порядков (тройные, четверные и т.д.), вероятность которых существенно меньше.

Последнее уравнение включает слагаемые с квадратами входных переменных

$$\left(\sum_{j=1}^m a_j x_j^2 \right)$$

и его коэффициенты получаются при обработке результатов активных экспериментов **II-го** порядка (верхний индекс II при $\hat{y} : \hat{y}''$) – например, **ОЦКП – ортогонального центрального композиционного плана эксперимента.**

Предпоследнее уравнение не включает слагаемые с квадратами входных переменных и его коэффициенты получаются при обработке результатов активных экспериментов **I-го** порядка – верхний индекс I при $\hat{y} : \hat{y}'$ - например, **ПФЭ – полный факторный эксперимент.**

При определении оптимальных условий проведения процесса с использованием эмпирических моделей (например, методом Бокса-Вильсона) выходная переменная \hat{y} является **критерием оптимальности или целевой функцией.**

В теории активного экспериментирования выходную (зависимую) переменную принято называть **функцией отклика**, а входные (независимые) переменные – **факторами**. Соответственно - координатное пространство с координатами (x_1, x_2, \dots, x_m) - **факторным пространством**, а геометрическое изображение функции отклика в факторном пространстве – **поверхностью отклика.**

Активный эксперимент планируется таким образом, чтобы упростить обработку его результатов методами регрессионного и корреляционного анализа.

Ортогональные планы экспериментов, используемые при активном экспериментировании, обеспечивают диагональный вид корреляционной матрицы $\bar{\bar{C}}$ при регрессионном анализе и, соответственно, статистическую независимость коэффициентов регрессии.

К другим достоинствам активного экспериментирования относятся:

- возможность предсказания количества опытов, которые следуют провести;
- определение точек факторного пространства, где следует проводить опыты;
- отсутствие проблем, связанных с выбором вида уравнения регрессии;
- возможность определения оптимальных параметров процесса экспериментально-статистическим методом;
- сокращение объёма опытных исследований.

1. ПФЭ и обработка его результатов.

Полный факторный эксперимент (ПФЭ) относится к экспериментам *I*-го порядка, т.к. описывающее его уравнение \hat{y}^I не включает факторы в квадрате.

Для двух факторов (x_1 и x_2) и без учёта взаимодействия факторов соответствующая эмпирическая модель может быть записана:

$$\hat{y}^I = a_0 + a_1x_1 + a_2x_2$$

В соответствии с теорией ПФЭ при проведении опытных исследований каждый из факторов варьируется только на двух уровнях – *минимальном* (кодированное значение -1) и *максимальном* (кодированное значение +1).

При этом реализуются возможные комбинации минимальных и максимальных значений факторов, в результате чего общее число опытов (n) в ПФЭ равно 2^m и полный факторный эксперимент обычно называется ПФЭ типа 2^m .

Для определения числа опытов применяется формула:

$$n = 2^m$$

В последнее уравнение включаются кодированные значения факторов z_j вместо x_j , значения которых получаются по следующей схеме кодирования:

$$z_j = \frac{x_j - x_j^{(0)}}{\Delta x_j}, \quad j = 1, \dots, m$$

где

$$x_j^{(0)} = 0.5(x_j^{\min} + x_j^{\max})$$
$$\Delta x_j = \frac{x_j^{\max} - x_j^{\min}}{2}, \quad j = 1, \dots, m$$

В результате план проведения эксперимента с учётом вышесказанного и кодирования факторов имеет вид:

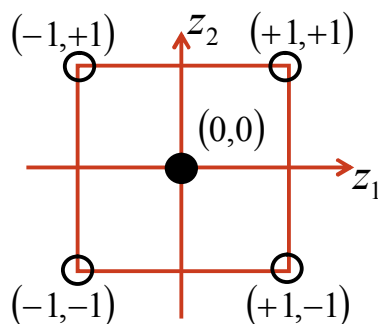
(число факторов равно $2 - m = 2$, число опытов $n = 2^m = 2^2 = 4$)

$n \backslash p$	z_0	z_1	z_2	y^3
1	+ 1	- 1	- 1	y_1^3
2	+ 1	+ 1	- 1	y_2^3
3	+ 1	- 1	+ 1	y_3^3
4	+ 1	+ 1	+ 1	y_4^3

При этом уравнение регрессии, описывающее эти опытные данные, записывается с использованием кодированных факторов z_j ($j = 0, 1, 2$) и соответственно кодированных коэффициентов регрессии $\tilde{a}_0, \tilde{a}_1, \tilde{a}_2$:

$$\hat{y} = \tilde{a}_0 + \tilde{a}_1 z_1 + \tilde{a}_2 z_2.$$

В кодированном факторном пространстве в соответствии с указанным планом проведения эксперимента проведённые опыты представляются точками вершин квадрата:



Для параметрической идентификации кодированного уравнения регрессии используется метод регрессионного анализа, включающий три этапа:

- **определение кодированных коэффициентов уравнения регрессии \tilde{a} методом наименьших квадратов;**
- **оценка значимости кодированных коэффициентов регрессии с использованием t – критерия Стьюдента;**
- **проверка адекватности кодированного уравнения регрессии с использованием F – критерия Фишера.**

Реализация двух последних этапов возможна при выполнении свойства *однородности дисперсий* (одно из требований регрессионного анализа) и проведении *параллельных опытов*, например, в точке с координатами $z_1 = 0$ и $z_2 = 0$ (центр плана, на рисунке – тёмная точка).

При проведении k параллельных опытов в центре плана $(y_{0s}^3, s = 1, \dots, k)$ среднее значение y_c^3 определяется как **среднее арифметическое результатов измерений** во всех параллельных опытах:

$$y_c^3 = \frac{\sum_{s=1}^k y_{0s}^3}{k}.$$

2. Определение кодированных коэффициентов регрессии (ПФЭ)

В этом случае используется применяемая при линейном регрессионном анализе матричная формула метода наименьших квадратов (МНК), которая с учётом кодирования факторов имеет вид:

$$\underset{(m+1) \times 1}{\bar{\bar{a}}} = \left(\underset{(m+1) \times n}{\bar{\bar{\Phi}}}^T \underset{n \times (m+1)}{\bar{\bar{\Phi}}} \right)^{-1} \underset{(m+1) \times n}{\bar{\bar{\Phi}}}^T \underset{n \times 1}{\bar{\bar{y}}},$$

где кодированная матрица, зависящая от независимых переменных для двух факторов включает только +1 и -1 и имеет вид:

$$\underset{(4 \times 3)}{\bar{\bar{\Phi}}} = \underset{(4 \times 3)}{z} = \begin{bmatrix} z_{10} & z_{11} & z_{12} \\ z_{20} & z_{21} & z_{22} \\ z_{30} & z_{31} & z_{32} \\ z_{40} & z_{41} & z_{42} \end{bmatrix} = \begin{bmatrix} +1 & -1 & -1 \\ +1 & +1 & -1 \\ +1 & -1 & +1 \\ +1 & +1 & +1 \end{bmatrix}$$

Матрица $\bar{\bar{z}}$ при активном экспериментировании называется **матрицей планирования** и обладает тремя оптимальными свойствами:

- **симметричности:** сумма элементов всех столбцов матрицы, кроме первого (точнее, нулевого) равна нулю

$$\sum_{i=1}^n z_{ij} = 0, \quad j = 1, \dots, m;$$

- **ортогональности:** скалярное произведение двух любых столбцов матрицы равно нулю

$$\bar{\bar{z}}_j^T \bar{\bar{z}}_u = \sum_{i=1}^n z_{ij} z_{iu} = 0$$

$$j, u = 0, 1, \dots, m \quad u \neq j;$$

- **нормировки:** скалярное произведение двух одинаковых столбцов матрицы равно n ($n = 2^m$ в ПФЭ)

$$\bar{\bar{z}}_j^T \bar{\bar{z}}_j = \sum_{i=1}^n z_{ij}^2 = n$$

$$j = 0, 1, \dots, m.$$

Благодаря перечисленным оптимальным свойствам матрицы планирования $\bar{\bar{z}}$ **информационная матрица** в ПФЭ при $m=2$ равна

$$\begin{matrix} \bar{\bar{I}} \\ (3 \times 3) \end{matrix} = \begin{matrix} \bar{\bar{\Phi}}^T \bar{\bar{\Phi}} \\ (3 \times 4) \quad (4 \times 3) \end{matrix} = \begin{matrix} \bar{\bar{z}}^T \bar{\bar{z}} \\ (3 \times 4) \quad (4 \times 3) \end{matrix} = \begin{bmatrix} n & 0 & 0 \\ 0 & n & 0 \\ 0 & 0 & n \end{bmatrix},$$

т.е. она является диагональной с одинаковыми элементами на главной диагонали, равными $n=2^2=4$.

Соответственно, **корреляционная матрица** $\bar{\bar{C}}$ также будет диагональной и с одинаковыми элементами главной диагонали:

$$\begin{matrix} \bar{\bar{C}} \\ (3 \times 3) \end{matrix} = \left(\begin{matrix} \bar{\bar{\Phi}}^T \bar{\bar{\Phi}} \\ (3 \times 4) \quad (4 \times 3) \end{matrix} \right)^{-1} = \left(\begin{matrix} \bar{\bar{z}}^T \bar{\bar{z}} \\ (3 \times 4) \quad (4 \times 3) \end{matrix} \right)^{-1} = \begin{bmatrix} n^{-1} & 0 & 0 \\ 0 & n^{-1} & 0 \\ 0 & 0 & n^{-1} \end{bmatrix}$$

Результатом подстановки последних соотношений в матричную формулу для определения кодированных коэффициентов регрессии будет простая формула:

$$\tilde{a}_j = \frac{\sum_{i=1}^n z_{ij} y_i^3}{n}, \quad j = 0, 1, \dots, m$$

При учёте взаимодействия двух факторов z_1 и z_2 кодированное уравнение регрессии принимает вид:

$$\hat{y} = \tilde{a}_0 + \tilde{a}_1 z_1 + \tilde{a}_2 z_2 + \tilde{a}_{12} z_1 z_2$$

и в матрицу планирования $\bar{\bar{z}}$ включается ещё один дополнительный последний столбец, каждый элемент которого равен произведению элементов столбцов, соответствующих взаимодействующим факторам:

$$\begin{matrix} \bar{\bar{\Phi}} \\ (4 \times 4) \end{matrix} = \begin{matrix} \bar{\bar{z}} \\ (4 \times 4) \end{matrix} = \begin{bmatrix} z_{10} & z_{11} & z_{12} & (z_{11} z_{12}) \\ z_{20} & z_{21} & z_{22} & (z_{21} z_{22}) \\ z_{30} & z_{31} & z_{32} & (z_{31} z_{32}) \\ z_{40} & z_{41} & z_{42} & (z_{41} z_{42}) \end{bmatrix} = \begin{bmatrix} +1 & -1 & -1 & +1 \\ +1 & +1 & -1 & -1 \\ +1 & -1 & +1 & -1 \\ +1 & +1 & +1 & +1 \end{bmatrix}$$

При этом матрица планирования сохраняет все три оптимальных свойства – **симметричности, ортогональности и нормировки**, а кодированный коэффициент уравнения регрессии при члене, характеризующем взаимодействие факторов, определяется по формуле:

$$\tilde{a}_{ju} = \frac{\sum_{i=1}^n (z_{ij} z_{iu}) y_i^3}{n}, \quad j, u = 1, \dots, m \quad u > j$$

В теории ПФЭ доказывается, что при увеличении числа факторов ($m > 2$) матрица планирования $\bar{\bar{z}}_{(n \times p)}$ строится с использованием рассмотренной методики, в том числе и с учётом взаимодействия факторов (не только двойного, но и тройного, четверного и т.д.).

В этом случае число столбцов матрицы p зависит от числа учёта взаимодействий факторов $n = 2^m$ и матрица планирования сохраняет перечисленные оптимальные свойства.

Поэтому для определения кодированных коэффициентов регрессии используются приведённые выше формулы.

Для расчёта натуральных значений коэффициентов в кодированное уравнение регрессии вместо кодированных факторов z_j ($j = 1, \dots, m$) следует подставить выражения для последних через натуральные значения факторов x_j ($j = 1, \dots, m$) в соответствии с приведённой выше схемой кодирования.

3. Определение значимости кодированных коэффициентов регрессии (ПФЭ)

Незначимость кодированных коэффициентов регрессии определяется с использованием квантиля t – распределения Стьюдента $t_{\beta(f_e)}^{\text{табл}}$ при помощи неравенства:

$$\frac{|\tilde{a}_j|}{S_{\tilde{a}_j}} \leq t_{\beta(f_e)}^{\text{табл}},$$

где

β – доверительная вероятность (в инженерных расчётах равная 0,95);

f_e – число степеней свободы дисперсии воспроизводимости (при одной серии параллельных опытов равная $k - 1$).

Выборочное значение квадратного корня дисперсии кодированного коэффициента регрессии определяется по формуле:

$$S_{\tilde{a}_j} = \sqrt{\tilde{C}_{jj}} S_e,$$

где

S_e – квадратный корень из дисперсии воспроизводимости, определяемой по k параллельным опытам в центре плана эксперимента:

$$S_e^2 = \frac{\sum_{j=1}^k (y_{0S}^j - y_c^j)^2}{k - 1} \equiv \frac{SS_e}{f_e},$$

где

SS_e – сумма квадратов дисперсии воспроизводимости;

f_e – число степеней свободы дисперсии воспроизводимости.

Как было показано выше, диагональные элементы корреляционной матрицы в ПФЭ при кодировании факторов одинаковы и равны $1/n$, вследствие чего

$$S_{\tilde{a}_j} = \frac{S_e}{\sqrt{n}}.$$

В результате условие не значимости кодированных коэффициентов регрессии принимает вид:

$$\frac{|\tilde{a}_j|}{S_e} \sqrt{n} \leq t_{\beta(f_e)}^{\text{табл}}.$$

Так как корреляционная матрица $\bar{\bar{C}}$ в этом случае является диагональной, то кодированные коэффициенты регрессии статистически независимы и при одновременной не значимости нескольких кодированных коэффициентов регрессии они (в отличие от процедуры обработки пассивного эксперимента) могут быть сразу, все вместе, исключены из кодированного уравнения регрессии.

4. Проверка адекватности уравнения регрессии (ПФЭ)

Проводится так же, как и при проведении пассивного эксперимента, с использованием табличного значения критерия Фишера, выбранного при доверительной вероятности β (чаще всего равной 0,95) и числе степеней свободы остаточной дисперсии f_R и дисперсии воспроизводимости f_e .

Условие адекватности проверяется с использованием неравенства:

$$\frac{S_R^2}{S_e^2} \leq F_{\beta(f_R, f_e)}^{\text{табл}},$$

где остаточная дисперсия, характеризующая точность уравнения, определяется по формуле:

$$S_R^2 = \frac{\sum_{i=1}^n (\hat{y}_i^I - y_i^e)^2}{n - p} = \frac{SS_R}{f_R}.$$

При этом $f_R = n - p$, где n – число экспериментов при различных значениях факторов; p – число значимых коэффициентов регрессии.

К недостаткам ПФЭ относится резкое увеличение числа опытов при возрастании количества факторов больше, чем 5 (при $m = 5$ $n = 2^5 = 32$).

Для проведения регрессионного анализа при пренебрежении целым рядом несущественных взаимодействий факторов достаточно проводить меньшее число опытов. В этом случае можно реализовать часть ПФЭ, т.н. дробный факторный эксперимент (ДФЭ), который здесь не рассматривается.

5. ОЦКП и обработка его результатов

Ортогональный центральный композиционный эксперимент (ОЦКП) относится к экспериментам II – го порядка, так как описывающее его уравнение \hat{y}^{II} включает факторы в квадрате и поэтому может описывать поверхности функций отклика в окрестности их экстремальных значений.

Для двух факторов (x_1 и x_2) с учётом только двойного взаимодействия факторов соответствующая эмпирическая модель может быть записана:

$$\hat{y}^{\text{II}} = a_0 + a_1 x_1 + a_2 x_2 + a_{12} x_1 x_2 + a_{11} x_1^2 + a_{22} x_2^2.$$

В соответствии с методикой ортогонального центрального композиционного плана эксперимента (ОЦКП) здесь, также как и для ПФЭ, осуществляется кодирование факторов по приведённой выше схеме, и для обеспечения ортогонального

свойства матрицы планирования эксперимента в уравнение регрессии \hat{y} включается некоторая постоянная S .

В результате уравнение регрессии при $m = 2$ принимает вид:

$$\hat{y} = \tilde{a}_0 + \tilde{a}_1 z_1 + \tilde{a}_2 z_2 + \tilde{a}_{12} z_1 z_2 + \tilde{a}_{11} (z_1^2 - S) + \tilde{a}_{22} (z_2^2 - S)$$

Для определения большего числа кодированных коэффициентов, чем при обработке ПФЭ, и описания поверхности функции отклика вблизи её экстремума («почти стационарной области»), количество опытов в этом случае увеличивается.

При этом опыты, проводимые при ПФЭ $n = 2^m$, дополняются опытами в «звёздных» точках факторного пространства $n_\alpha = 2m$ и опытами в центре плана с координатами $z_1 = 0$ и $z_2 = 0$ (n_c).

«Звёздные» точки в факторном пространстве располагаются на осях координат на расстоянии $+\alpha$ и $-\alpha$ от центра плана эксперимента; причём величина α называется «звёздным» плечом и её значения, так же как величина S , определяются из условия ортогональности матрицы планирования \bar{z} для ОЦКП.

Общее число опытов N в ортогональном центральном композиционном эксперименте определяется по формуле:

$$N = n + n_\alpha + n_c,$$

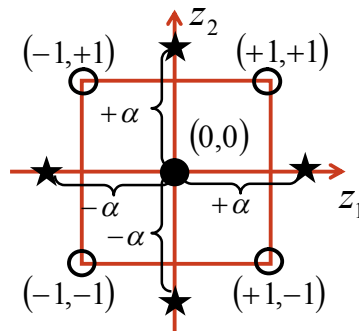
или с учётом приведённых выше равенств:

$$N = 2^m + 2m + n_c.$$

Для случая двух факторов ($m = 2$):

$$N = 8 + n_c.$$

Расположение опытных точек в факторном пространстве для случая двух факторов в приведённой ранее кодированной системе координат может быть представлено:



План проведения экспериментов в этом случае может быть представлен:

$n \backslash p$	z_0	z_1	z_2	$z_1 z_2$	$z_1^2 - S$	$z_2^2 - S$	y^3
2^m	1	+1	-1	-1	+1	$1-S$	y_1^3
	2	+1	+1	-1	-1	$1-S$	y_2^3
	3	+1	-1	+1	-1	$1-S$	y_3^3
	4	+1	+1	+1	+1	$1-S$	y_4^3
$2m$	5	+1	$-\alpha$	0	0	$\alpha^2 - S$	y_5^3
	6	+1	$+\alpha$	0	0	$\alpha^2 - S$	y_6^3
	7	+1	0	$-\alpha$	0	$-S$	y_7^3
	8	+1	0	$+\alpha$	0	$-S$	y_8^3
n_c	9	+1	0	0	0	$-S$	y_9^3
	\vdots	+1	0	0	0	$-S$	\vdots
	N	+1	0	0	0	$-S$	y_N^3

Матрица планирования $\bar{\bar{z}}$ представляет собой часть плана проведения эксперимента без горизонтальных и вертикальных заголовков таблицы и вектора наблюдения \bar{y}^3 (правого столбца).

6. Определение величины «звёздного плеча» α и S из условия ортогональности матрицы планирования $\bar{\bar{z}}$.

Матрица планирования $\bar{\bar{z}}_{(N \times 6)}$ была бы ортогональной, если бы выполнялись следующие равенства:

$$\begin{cases} \bar{\bar{z}}_0^T (\bar{\bar{z}}_j - \bar{S}) = 0 \\ j = 1, 2 \end{cases}$$

и

$$(\bar{\bar{z}}_1^2 - \bar{S})^T (\bar{\bar{z}}_2^2 - \bar{S}) = 0.$$

Раскрывая первое равенство, можно получить:

$$\bar{\bar{z}}_0^T (\bar{\bar{z}}_j - \bar{S}) = \sum_{i=1}^N z_{i0} z_{ij}^2 - \sum_{i=1}^n z_{i0} S = n + 2\alpha^2 - NS = 0$$

$$j = 1, 2.$$

Откуда:

$$S = \frac{n + 2\alpha^2}{N} \quad (A).$$

Раскрывая второе равенство, получаем:

$$\begin{aligned} (\bar{\bar{z}}_1^2 - \bar{S})^T (\bar{\bar{z}}_2^2 - \bar{S}) &= (\bar{\bar{z}}_1^2)^T \bar{\bar{z}}_2^2 - (\bar{\bar{z}}_1^2)^T \bar{S} - \bar{S}^T \bar{\bar{z}}_2^2 + \bar{S}^T \bar{S} = \\ &= n - (n + 2\alpha^2)S - S(n + 2\alpha^2) + NS^2 = n - 2NS^2 + NS^2 = \\ &= n - NS^2 = 0 \end{aligned}$$

Откуда:

$$S = \sqrt{\frac{n}{N}} \quad (B).$$

Последнее выражение используется для определения S .

Приравняв правые части двух выражений для S , можно найти формулу для определения α :

$$\frac{n + 2\alpha^2}{N} = \sqrt{\frac{n}{N}}$$

$$\alpha^2 = \frac{1}{2}(\sqrt{Nn} - n) = \frac{n}{2}\left(\sqrt{\frac{N}{n}} - 1\right).$$

В результате звёздное плечо α можно определить по формуле:

$$\alpha = \sqrt{\frac{n}{2}\left(\sqrt{\frac{N}{n}} - 1\right)}.$$

4. Определение кодированных коэффициентов регрессии (ОЦКП)

В соответствии с методом наименьших квадратов эти коэффициенты определяются по матричной формуле:

$$\bar{\tilde{a}} = \bar{\tilde{C}}^T \bar{y},$$

где

$$\bar{\tilde{C}} = \begin{pmatrix} \bar{z}^T & \bar{z} \end{pmatrix}^{-1}$$

Из-за свойства ортогональности матрицы планирования \bar{z} необходимо определить только диагональные элементы информационной матрицы:

$$\bar{\tilde{I}} = \bar{z} \bar{z}^T,$$

а затем диагональные элементы корреляционной матрицы:

$$\bar{\tilde{C}} = \bar{\tilde{I}}^{-1}.$$

8. Определение диагональных элементов информационной и корреляционной матриц.

Обобщая уравнение регрессии на случай m факторов и учитывая только все двойные взаимодействия факторов, число которых определяется по формуле:

$$C_m^2 = \frac{m(m-1)}{2!},$$

общее число коэффициентов уравнения регрессии для t факторов равно:

$$p = 1 + m + \frac{m(m-1)}{2!} + m,$$

диагональные элементы информационной матрицы \hat{I} определяются:

$i_{00} = N$ - число таких элементов равно 1;

$$i_{jj} = n + 2\alpha^2 \quad (j = 1, \dots, m);$$

$$i_{ju} = n \quad (u > j) - \text{число таких элементов равно: } \frac{m(m-1)}{2!}.$$

Для определения при квадратах факторов i_{jj} можно записать:

$$\begin{aligned} i_{jj} &= n(1-S)^2 + 2(\alpha^2 - S)^2 + (N-n-2)S^2 = \\ &= n - 2nS + nS^2 + 2\alpha^4 - 4\alpha^2 S + 2S^2 + NS^2 - nS^2 - 2S^2 = \\ &= 2\alpha^4 + n - 2S \underbrace{(n + 2\alpha^2 S)}_{\text{из равенства (A) } NS} + NS^2 = 2\alpha^4 + \underbrace{n - NS^2}_{\text{из равенства (B) } =0} = 2\alpha^4 \end{aligned}$$

Количество таких диагональных элементов — m .

Диагональная матрица \bar{I} имеет размер:

$$p = 1 + m + \frac{m(m-1)}{2} + m = \frac{(m+1)(m+2)}{2},$$

что соответствует числу определяемых параметров p .

В результате диагональная корреляционная матрица $\tilde{C} = \tilde{\Gamma}^{-1}$ размером $p \times p$ для m факторов и с учётом их двойных взаимодействий имеет вид:

[illegible]

Элементы корреляционной матрицы определяются по матричной формуле МНК:

$$\bar{\bar{a}} = \bar{\bar{C}} \bar{\bar{z}}^T \bar{\bar{y}}.$$

Кодированные коэффициенты регрессии определяются:

$$\begin{aligned} \tilde{a}_0 &= \frac{\sum_{i=1}^N y_i^3}{N}; & \tilde{a}_j &= \frac{\sum_{i=1}^N z_{ij} y_i^3}{n + 2\alpha^2} \quad (j = 1, \dots, m); \\ \tilde{a}_{ju} &= \frac{\sum_{i=1}^N (z_{ij} z_{iu}) y_i^3}{n} \quad u > j & & \left(\text{число коэффициентов} \frac{m(m-1)}{2} \right) \end{aligned}$$

$$\tilde{a}_{jj} = \frac{\sum_{i=1}^n (z_{ij}^2 - S) y_i^3}{2\alpha^4} \quad (j = 1, \dots, m).$$

Для пересчёта этих коэффициентов регрессии в натуральные значения необходимо вместо кодированных факторов z подставить их натуральные величины x_j в соответствии с приведённой схемой кодирования.

9. Определение значимости кодированных коэффициентов регрессии (ОЦКП)

В отличие от ПФЭ значимость коэффициентов регрессии определяется по разным формулам для различных коэффициентов, так как диагональные элементы корреляционной матрицы $\bar{\bar{C}}$ отличаются друг от друга.

С учётом общей формулы для определения не значимости коэффициентов регрессии

$$\frac{|\tilde{a}_j|}{\sqrt{\tilde{C}_{jj}} S_e} \leq t_{\beta(S_e)}^{\text{табл}},$$

не значимость каждого вида коэффициентов регрессии определяется:

$$\frac{|\tilde{a}_0|}{S_e} \sqrt{N} \leq t_{\beta(S_e)}^{\text{табл}};$$

$$\frac{|\tilde{a}_j|}{S_e} \sqrt{n + 2\alpha^2} \leq t_{\beta(S_e)}^{\text{табл}} \quad (j = 1, \dots, m)$$

$$\frac{|\tilde{a}_{ju}|}{S_e} \sqrt{n} \leq t_{\beta(S_e)}^{\text{табл}} \quad \left(\text{число коэффициентов} \frac{m(m-1)}{2} \right)$$

$$\frac{|\tilde{a}_{jj}|}{S_e} \sqrt{2\alpha^4} \leq t_{\beta(S_e)}^{\text{табл}} \quad (j = 1, \dots, m).$$

10. Проверка адекватности уравнения регрессии (ОЦКП)

Осуществляется с использованием критерия Фишера – так же, как и в случае с ПФЭ.

11. Определение экстремума функции отклика

Уравнение регрессии с m факторами вида:

$$\hat{y}^{\Pi} = \tilde{a}_0 + \sum_{j=1}^m \tilde{a}_j z_j + \sum_{j=1}^{m-1} \sum_{u=2}^m \tilde{a}_{ju} z_j z_u + \sum_{j=1}^m \tilde{a}_{jj} (z_j^2 - S)$$
$$u > j$$

может применяться для определения экстремума функции отклика с использованием необходимого условия экстремума функции многих переменных:

$$\frac{\partial \hat{y}^{\Pi}}{\partial z_1} = 0$$
$$\dots \dots \dots$$
$$\frac{\partial \hat{y}^{\Pi}}{\partial z_m} = 0 .$$

Полученная система линейных уравнений (СЛАУ) позволяет расчётным путём определить $z_j^{opt} (j=1, \dots, m)$ и после подстановки их величин в исходное уравнение \hat{y}^{Π} получить максимальное или минимальное значение функции отклика.

Лекция 3

Планы дробного факторного эксперимента (планы ДФЭ)

При многофакторном эксперименте, особенно когда число факторов больше шести ($n > 6$), число опытов планов ПФЭ 2^n ($N = 2^n$) становится чрезмерным. Если нам не требуется определение всех коэффициентов неполного квадратичного полинома, то переходят к дробному факторному эксперименту (ДФЭ) – части полного факторного эксперимента. Так, например, если требуется определить лишь коэффициенты при самих факторах

$$Y = b_0 + b_1 x_1 + b_2 x_2 + \dots + b_n x_n,$$

то план ПФЭ 2^n дает избыточную информацию. Так при $n=6$, в этом случае требуется определить $n+1=7$ коэффициентов, тогда как по плану ПФЭ необходимо провести $N = 2^6 = 64$ опыта.

Хотя эта избыточная информация не является бесполезной, она позволяет более точно определить коэффициенты, но все же часто используют планы ДФЭ 2^{n-k} , где k – показатель дробности плана ПФЭ. При $k=1$ число опытов в плане ДФЭ в два раза меньше, чем в плане ПФЭ, поэтому такие планы называют полуреплика плана ПФЭ. Так при $k=1$ для плана ДФЭ 2^{6-1} $N = 2^{6-1} = 32$, при $k=2$ для плана ДФЭ 2^{6-2} $N = 2^{6-2} = 16$ и такой план называют четвертьрепликой, при $k=3$ для плана ДФЭ 2^{6-3} $N = 2^{6-3} = 8$.

При выборе дробности плана k необходимо учитывать, что число опытов должно быть больше числа членов уравнения. В рассматриваемом случае величина k должна быть такой, что бы удовлетворялось условие

$$n+1 \leq 2^{n-k}.$$

План ДФЭ строится, как и для плана ПФЭ, но с меньшим числом факторов. Оставшиеся факторы варьируются не произвольно, а так чтобы сохранялась ортогональность плана. Это обеспечивается, если оставшиеся факторы варьируются по выбранному генерирующему соотношению, например как произведение каких-либо факторов из первой группы. Но это приводит к тому, что в матрице X будут существовать одинаковые столбцы. Следовательно, мы не сможем найти в чистом виде все коэффициенты неполного квадратичного полинома, а лишь определим совместную величину коэффициентов для одинаковых столбцов.

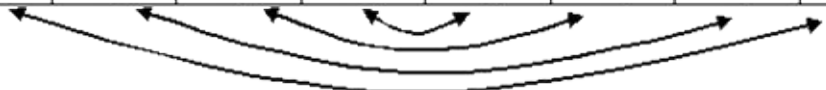
Рассмотрим построение плана ДФЭ 2^{3-1} . Здесь $n = 3$, $\kappa = 1$, $N = 2^{3-1} = 4$. Первые два фактора варьируем как и ранее для плана ПФЭ 2^2 , а для третьего фактора выбираем генерирующее соотношение в виде $x_3 = x_1 \cdot x_2$.

Для неполного квадратичного полинома

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

количество столбцов плана составляет восемь.

i	0	1	2	3	4	5	6	7
U	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
2	+1	+1	-1	-1	-1	-1	+1	+1
3	+1	-1	+1	-1	-1	+1	-1	+1
4	+1	+1	+1	+1	+1	+1	+1	+1



План является ортогональным, но в нем оказались четыре пары одинаковых столбцов. Поэтому можно определить только четыре коэффициента, отражающие совместные влияния двух одинаковых столбцов

$$b_0 + b_{123} = \frac{\sum_{U=1}^N x_{0U} Y_U}{N}.$$

Суммарные значения коэффициентов $b_1 + b_{23}$; $b_2 + b_{13}$; $b_3 + b_{12}$ определяются аналогично. Это следствие того, что мы пытаемся определить полное количество коэффициентов – 8 по недостаточному числу опытов - 4. Однако, если заранее известно, что некоторые из членов уравнения равны нулю (пренебрежимо малы) или имеется априорная информация о величинах некоторых коэффициентов, то полученные коэффициенты могут быть вычленены. Так если $b_{123} = 0$, то

$$b_0 = \frac{\sum_{U=1}^N x_{0U} Y_U}{N}.$$

Если можно допустить, что коэффициенты из их смешанной оценки сопоставимы, то для рассмотренного плана

$$b_0 = b_{123} = \frac{1}{2} \frac{\sum_{v=1}^N x_{0v} Y_v}{N}.$$

Графическое изображение планов ПФЭ 2^3 и ДФЭ 2^{3-1} в факторном пространстве (для трех факторов - трехмерное пространство) представлено на рис. 1. План ПФЭ 2^3 представлен кубом с восемью узлами (точками плана), а возможные планы ДФЭ 2^{3-1} – проекциями этого куба на три плоскости. То есть из восьми узлов выбираются четыре (рис. 1, а). Из куба можно также выбрать четыре точки из восьми, не лежащие в одной плоскости, и сформировать план ДФЭ 2^{3-1} (рис. 1, б).

Насыщенные планы первого порядка.

Насыщенным планом первого порядка – называется план, содержащий $n+1$ точку (опыт). Например, при $n = 4$, $N=n + 1 = 5$.

То есть полином формируется в виде

$$Y = b_0 + b_1 x_1 + \dots + b_n x_n.$$

Таким образом, насыщенный план – это предельно минимальный случай плана ДФЭ. Такие планы называются **симплекс-планами**. Для симплекс-плана при $n = 1$ $N = 2$ его геометрическое изображение представлено на рис. 2, а; при $n=2$, $N=3$ – на рис. 2, б; при $n=3$, $N=4$ – на рис. 2, в. Симплекс-планы обычно используются на стадии предварительного исследования.

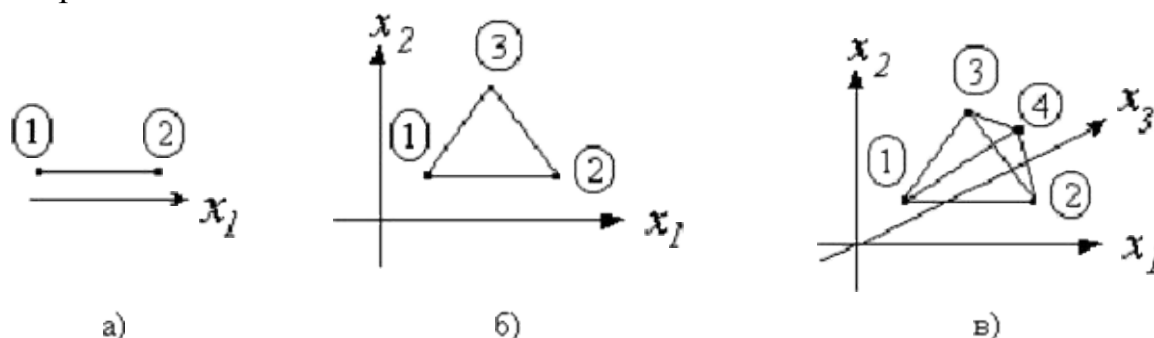


Рис. 2. Симплекс-план для $n=1$, $N=2$ (а); $n=2$, $N=3$ (б); $n=3$, $N=4$ (в)

Симплекс-план не всегда является ортогональным. Симплекс-план называется правильным, если расстояние между двумя любыми точками плана одинаковое. Симплекс-план называется центрированным, если

$$\sum_{v=1}^{n+1} x_{iv} = 0,$$

для $i=1, 2, \dots, n$.

Применимость планов ПФЭ и пути повышения точности полиномов

По каким же признакам можно судить о допустимости использования неполного квадратичного полинома, построенного на основе планов ПФЭ 2^n ?

Такие полиномы дают поверхность отклика, которая проходит точно через все экспериментальные точки, по которым определяются коэффициенты. Так как точки планов ПФЭ располагаются на границах диапазонов варьирования факторов, то это означает, что поверхность отклика проходит через граничные точки. В любом сечении поверхности отклика, полученной по такому полиному, плоскостью при фиксированных всех факторах кроме одного и параллельной оси Y получается след в виде прямой линии.

Возможны случаи, когда реальная поверхность отклика определяется полиномами второго и выше порядков. В этом случае поверхность плана ПФЭ, совпадая с реальной поверхностью в граничных точках, может отличаться в других точках факторного пространства, например в центральной точке плана, т.е. $Y_0 \neq \hat{Y}_0$. Поэтому одним из признаков неудовлетворительной аппроксимации полиномами по плану ПФЭ является расхождение результатов функции отклика с реальной функцией в центральной точке плана.

Однако при многофакторном эксперименте возможны случаи, когда в реальности функция отклика зависит, в том числе, от квадратов факторов, у которых коэффициенты имеют разные знаки, например, для “седловидной” поверхности. При этом, несмотря на то, что эта поверхность явно нелинейная, результат опыта в центральной точке может оказаться достаточно близким к полученному результату по неполному квадратичному полиному плана ПФЭ. Однако расхождения будут возникать во всех других точках плана эксперимента. Поэтому нецелесообразность использования плана ПФЭ определяется нелинейностью каких-либо сечений поверхности отклика. Косвенным признаком может служить расхождение Y_0 и \hat{Y}_0 в центральной точке плана.

Если не удастся получить полином по плану ПФЭ, хорошо аппроксимирующей реальную поверхность, то какие пути можно предложить для повышения точности полиномов?

- Уменьшение диапазона варьирования факторов или его разбиение на поддиапазоны, для каждого из которых строится свой план ПФЭ и определяется свой полином. Путь достаточно трудоемок, но погрешность семейства планов ПФЭ снижается.

- Выделение фактора, порождающий нелинейность, и построение для оставшихся $n-1$ факторов k планов ПФЭ, в каждом из которых выделенный фактор зафиксирован при некотором значении. На основе полученных k полиномов можно попытаться сформировать общий полином, коэффициенты которого являются функциями выделенного фактора. Этот путь также достаточно трудоемок.

- Переход к плану ПФЭ с большим числом уровней варьирования факторов, например к планам с варьированием факторов на трех уровнях - планам ПФЭ 3^n (рис. 3). В этом случае происходит резкое увеличение количества точек по сравнению с планом ПФЭ 2^n . Так при $n = 2$ для ПФЭ 2^n $N=4$, для ПФЭ 3^n $N=9$; при $n = 3$ для ПФЭ 2^n $N=8$, для ПФЭ 3^n $N=27$; при $n = 4$ для ПФЭ 2^n $N=16$, для ПФЭ 3^n $N=81$ и т.д.

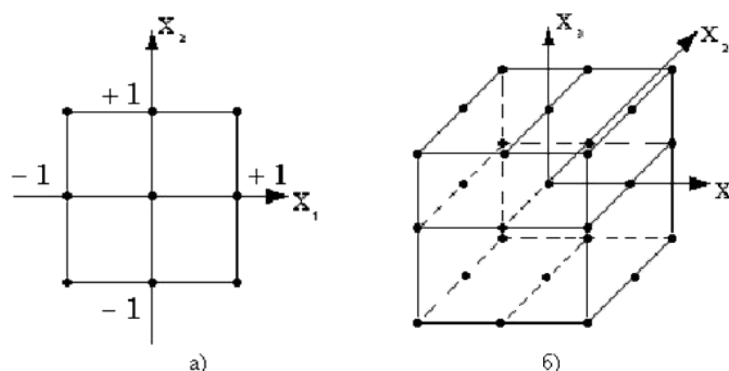


Рис. 3. Планы ПФЭ 3^2 (а) и ПФЭ 3^3 (б)

- Достраивание планов ПФЭ 2^n до планов более высокого порядка (чаще всего второго) и построение полных квадратичных полиномов (с наличием квадратов факторов).
- Преобразование метрики матричного пространства, то есть переход к новым факторам функционально связанным с прежними факторами, но не порождающими нелинейности.

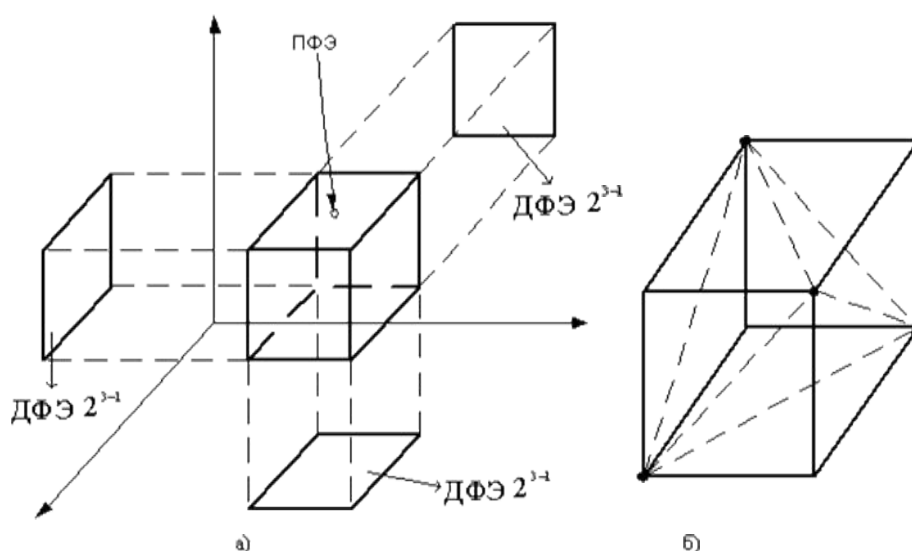


Рис. 4. Графическое изображение планов ПФЭ 2^3 и ДФЭ 2^{3-1} в факторном пространстве

Планы ДФЭ, как и планы ПФЭ, являются ротатабельными. Планы ДФЭ могут быть как **насыщенными** так и **ненасыщенными**.

Достоинство планов ДФЭ заключается и в том, что если построенный на его основе неполный полином не удовлетворяет требованиям по точности, то план ДФЭ легко достраивается до плана ПФЭ, без потери информации прежних опытах, с формированием более точного полинома.

Пример построения плана ДФЭ.

Построить план ДФЭ 2^{4-1} и определить полином

$$Y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4 + b_{12}x_1x_2 + b_{13}x_1x_3 + b_{14}x_1x_4 + b_{23}x_2x_3 + b_{24}x_2x_4 + b_{34}x_3x_4.$$

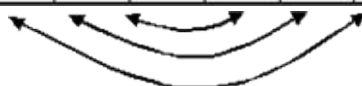
Число факторов – 4. Нужно найти 8 коэффициентов полинома. Выбираем 8 из 16 опытов плана ПФЭ 2^4 таким образом, чтобы были определены независимые коэффициенты при самих факторах, смешанные коэффициенты при парных сочетаниях

факторов и в пренебрежении тройными и четверным сочетаниями факторов и при этом сохранялась ортогональность плана.

U		x_1	x_2	x_3	x_4
ПФЭ 2^4	ДФЭ 2^{4-1}				
1	1	-1	-1	-1	-1
2	-	+1	-1	-1	-1
3	-	-1	+1	-1	-1
4	4	+1	+1	-1	-1
5	-	-1	-1	+1	-1
6	6	+1	-1	+1	-1
7	7	-1	+1	+1	-1
8	-	+1	+1	+1	-1
9	-	-1	-1	-1	+1
10	2	+1	-1	-1	+1
11	3	-1	+1	-1	+1
12	-	+1	+1	-1	+1
13	5	-1	-1	+1	+1
14	-	+1	-1	+1	+1
15	-	-1	+1	+1	+1
16	8	+1	+1	+1	+1

Такой выбор позволяет сформировать план ДФЭ 2^{4-1} как и план ПФЭ 2^3 , но с $x_4 = x_1 x_2 x_3$. План ДФЭ 2^{4-1} представляется в виде

U	x_0	x_1	x_2	x_3	x_4	$x_1 x_2$	$x_1 x_3$	$x_1 x_4$	$x_2 x_3$	$x_2 x_4$	$x_3 x_4$	Y	\bar{Y}
1	+1	-1	-1	-1	-1	+1	+1	+1	+1	+1	+1	10	10
2	+1	+1	-1	-1	+1	-1	-1	+1	+1	-1	-1	8	8
3	+1	-1	+1	-1	+1	-1	+1	-1	-1	+1	-1	8	8
4	+1	+1	+1	-1	-1	+1	-1	-1	-1	-1	+1	7	6,8
5	+1	-1	-1	+1	+1	+1	-1	-1	-1	-1	+1	9	9,2
6	+1	+1	-1	+1	-1	-1	+1	-1	-1	+1	-1	8	8
7	+1	-1	+1	+1	-1	-1	-1	+1	+1	-1	-1	8	8
8	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1	6,5	6,68



Значения коэффициентов полинома составляют:

$$b_{123} = b_{124} = b_{134} = b_{234} = b_{1234} = 0$$

$$b_0 = \frac{10+8+8+7+9+8+8+6,5}{8} = 8,06;$$

$$b_1 = \frac{-10+8-8+7-9+8-8+6,5}{8} = -0,69;$$

$$b_2 = -0,69; \quad b_3 = -0,19; \quad b_4 = -0,19;$$

$$(b_{12} + b_{34}) = \frac{10 - 8 - 8 + 7 + 9 - 8 - 8 + 6,5}{8} = 0,06;$$

$$b_{13} + b_{24} = 0,06, \quad b_{14} + b_{23} = 0,06.$$

Если принять, что

$$b_{12} \approx b_{34} = \frac{1}{2}(b_{12} + b_{34}) = 0,03,$$

$$b_{13} \approx b_{24} = \frac{1}{2}(b_{13} + b_{24}) = 0,03,$$

$$b_{14} \approx b_{23} = \frac{1}{2}(b_{14} + b_{23}) = 0,03,$$

то полином имеет вид

$$\hat{y} = 8,06 - 0,69x_1 - 0,69x_2 - 0,19x_3 - 0,19x_4 + 0,03x_1x_2 + \\ + 0,03x_1x_3 + 0,03x_1x_4 + 0,03x_2x_3 + 0,03x_2x_4 + 0,03x_3x_4.$$

Значения полинома в точках плана приведены в последнем столбце плана ДФЭ 2^{4-1} . В нашем случае точность его достаточно высокая.

Лекция 4

Дробный факторный эксперимент

При большом числе факторов ($k > 3$) проведение полного факторного эксперимента связано с большим числом, экспериментов, значительно превосходящим число коэффициентов линейной модели. Если при получении модели можно ограничиться, линейным приближением, т. е. получить адекватную модель в виде полинома $y = b_0 + b_1x_1 + b_2x_2 + \dots + b_kx_k$, то число экспериментов можно резко сократить в результате использования дробного факторного эксперимента. Так, например, в полном факторном эксперименте типа 2^2 при линейном приближении коэффициент регрессии b_{12} можно принять равным нулю, а столбец x_1x_2 матрицы (табл. 4) использовать для третьего фактора x_3 .

Таблица 4

Матрица планирования

Номер эксп- пер.	x_0	x_1	x_2	$x_3 (x_1x_2)$	y
1	+	+	+	+	y_1
2	+	-	+	-	y_2
3	+	+	-	-	y_3
4	+	-	-	+	y_4

В этом случае линейная модель будет определяться уравнением $y = b_0 + b_1x_1 + b_2x_2 + b_3x_3$. Для определения коэффициентов этого уравнения достаточно провести четыре эксперимента вместо восьми в полном факторном эксперименте типа

2^3 . План эксперимента, предусматривающий реализацию половины экспериментов полного факторного эксперимента, называют **полурепликой**. При увеличении числа факторов ($k > 3$) возможно применение реплик большей дробности.

Дробной репликой называют план эксперимента, являющийся частью плана полного факторного эксперимента. Дробные реплики обозначают зависимостью 2^{k-p} , где p - число линейных эффектов, приравненных к эффектам взаимодействия. При $p = 1$ получают **полуреплику**; при $p = 2$ получают **1/4 - реплику**; при $p = 3$ получают **1/8 - реплику** и т. д. по степеням двойки. Так, например, если в полном факторном эксперименте 2^3 (табл. 5) один из эффектов взаимодействия (x_1x_2 , x_1x_3 , x_2x_3 , $x_1x_2x_3$) заменим четвертым фактором x_4 , то получим полуреплику 2^{4-1} от полного факторного эксперимента 2^4 . Если два эффекта взаимодействия заменить факторами x_4 и x_5 , то получим $1/4$ -реплику 2^{5-2} от полного факторного эксперимента 2^5 .

Таблица 5

Матрица полного факторного эксперимента типа 2^3

Номер экспер.	x_0	x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3	$x_1x_2x_3$	y
1	+	-	-	+	+	-	-	+	y_1
2	+	+	-	+	-	+	-	-	y_2
3	+	-	+	+	-	-	+	-	y_3
4	+	+	+	+	+	+	+	+	y_4
5	+	-	-	-	+	+	+	-	y_5
6	+	+	-	-	-	-	+	+	y_6
7	+	-	+	-	-	+	-	+	y_7
8	+	+	+	-	+	-	-	-	y_8

Можно получать $1/8$ -реплику от полного факторного эксперимента 2^6 , заменив три эффекта взаимодействия факторами x_4 , x_5 и x_6 . Если заменить четыре эффекта взаимодействия факторами x_4 , x_5 , x_6 и x_7 , то получим $1/16$ -реплику 2^{7-4} от полного факторного эксперимента 2^7 .

Реплики, которые используют для сокращения числа экспериментов в 2^m раз, где $m=1, 2, 3 \dots$, называют регулярными.

В связи с тем, что в дробных репликах часть взаимодействий заменена новыми факторами, найденные коэффициенты уравнения регрессии будут являться совместными оценками линейных эффектов и эффектов взаимодействия. Так, например, если в матрице (табл. 4) вычислим элементы столбцов для произведений x_1x_3 и x_2x_3 , то увидим, что элементы столбца x_1x_2 совпадают с элементами столбца x_2 , а элементы столбца x_2x_3 - с элементами столбца x_1 . Следовательно, коэффициенты b_1 , b_2 , b_3 будут оценками совместных эффектов, а именно: $b_1 \rightarrow \beta_1 + \beta_{23}$; $b_2 \rightarrow \beta_2 + \beta_{13}$; $b_3 \rightarrow \beta_3 + \beta_{12}$.

Коэффициент b_1 является оценкой влияния фактора x_1 и парного взаимодействия x_2x_3 на функцию отклика. Влияние фактора x_1 в этом случае характеризуется величиной β_1 , а влияние взаимодействия - величиной β_{23} . *Оценки, в которых невозможно разделить линейный эффект и эффект взаимодействия, называют сме-*

шанными. Линейные эффекты рекомендуется смешивать, прежде всего, с теми взаимодействиями, которые согласно априорной информации незначимы.

Число несмешанных линейных эффектов в дробной реплике называют ее разрешающей способностью.

Часто приходится решать задачи, в которых заранее можно полагать, что эффекты взаимодействия, хотя и малы по сравнению с линейными, но все же не равны нулю. В таких случаях необходимо заранее определить, какие коэффициенты являются смешанными оценками. Тогда в зависимости от условий поставленной задачи, подбирается такая дробная реплика, с помощью которой можно извлечь максимальную информацию из эксперимента.

Прямая оценка разрешающей способности дробной реплики затруднена. Поэтому дробные реплики задают с помощью генерирующих соотношений. **Генерирующим называют соотношение, которое показывает, какое из взаимодействий принято незначимым и заменено новым фактором.**

План типа 2^{3-1} может быть представлен двумя полурепликами (табл. 6), которые задаются одним из следующих генерирующих соотношений:

$$x_3 = x_1x_2; \quad x_3 = -x_1x_2.$$

Генерирующие соотношения умножим на новую независимую переменную x_3 :

$$x_3^2 = x_1x_2x_3; \quad x_3^2 = -x_1x_2x_3.$$

Таблица 6

Две полуреплики плана типа 2^{3-1}

Номер экспер.	$x_3 = x_1x_2$			Номер экспер.	$x_3 = -x_1x_2$		
	x_1	x_2	x_3		x_1	x_2	x_3
1	-	+	-	1	-	+	+
2	+	+	+	2	+	+	-
3	-	-	+	3	-	-	-
4	+	-	-	4	+	-	+

Поскольку всегда $x_i^2 = 1$, получим следующие соотношения:

$$1 = x_1x_2x_3; \quad 1 = -x_1x_2x_3. \quad (3)$$

В результате умножения генерирующего соотношения на новую переменную получают так называемый определяющий контраст. Для указанных выше полуреplik определяющими контрастами будут зависимости (3). Зная определяющий контраст, можно найти соотношения, задающие совместные оценки. Для этого необходимо умножить независимые переменные x_1 , x_2 и x_3 на определяющий контраст. Умножая определяющие контрасты (3) на x_1 , получим соотношения

$$x_1 \cdot 1 = x_1^2x_2x_3; \quad x_1 \cdot 1 = -x_1^2x_2x_3;$$

так как $x_1^2 = 1$, то

$$x_1 = x_2x_3; \quad x_1 = -x_2x_3.$$

Умножая определяющие контрасты на x_2 и x_3 , получаем следующие соотношения:

$$x_2 = x_1x_3; \quad x_2 = -x_1x_3;$$

$$x_3 = x_1x_2; \quad x_3 = -x_1x_2.$$

Это означает, что коэффициенты регрессии будут оценками

$$b_1 \rightarrow \beta_1 + \beta_{23}; \quad b_1 \rightarrow \beta_1 - \beta_{23};$$

$$b_2 \rightarrow \beta_2 + \beta_{13}; \quad b_2 \rightarrow \beta_2 - \beta_{13};$$

$$b_3 \rightarrow \beta_3 + \beta_{12}; \quad b_3 \rightarrow \beta_3 - \beta_{12}.$$

Полуреплика 2^{4-1} может быть задана генерирующим соотношением $x_4 = x_1x_2x_3$. Матрица планирования этой полуреплики представлена табл. 7.

Определяющим контрастом полуреплики является соотношение

$$1 = x_1x_2x_3x_4.$$

Совместные оценки будут определяться следующим образом:

$$x_1 = x_2x_3x_4 \quad b_1 \rightarrow \beta_1 + \beta_{234};$$

$$x_2 = x_1x_3x_4 \quad b_2 \rightarrow \beta_2 + \beta_{134};$$

$$x_3 = x_1x_2x_4 \quad b_3 \rightarrow \beta_3 + \beta_{124};$$

$$x_4 = x_1x_2x_3 \quad b_4 \rightarrow \beta_4 + \beta_{123};$$

$$x_1x_2 = x_2x_4 \quad b_{12} \rightarrow \beta_{12} + \beta_{34};$$

$$x_1x_3 = x_3x_4 \quad b_{13} \rightarrow \beta_{13} + \beta_{24};$$

$$x_1x_4 = x_2x_3 \quad b_{14} \rightarrow \beta_{14} + \beta_{23}.$$

Полуреплика 2^{4-1} может быть также задана генерирующим соотношением $x_4 = x_1x_2$. Матрица планирования этой полуреплики представлена табл. 8.

Таблица 7

Полуреплика 2^{4-1} с определяющим контрастом $1 = x_1x_2x_3x_4$.

Номер экспер.	x_0	x_1	x_2	x_3	x_4	y
1	+	-	-	+	+	y_1
2	+	+	-	+	-	y_2
3	+	-	+	+	-	y_3
4	+	+	+	+	+	y_4
5	+	-	-	-	-	y_5
6	+	+	-	-	+	y_6
7	+	-	+	-	+	y_7
8	+	+	+	-	-	y_8

Таблица 8

Полуреплика 2^{4-1} с определяющим контрастом $1 = x_1x_2x_4$

кон-

Номер экспер.	x_0	x_1	x_2	x_3	x_4	y
1	+	-	-	+	+	y_1
2	+	+	-	+	-	y_2
3	+	-	+	+	-	y_3
4	+	+	+	+	+	y_4
5	+	-	-	-	+	y_5
6	+	+	-	-	-	y_6
7	+	-	+	-	-	y_7
8	+	+	+	-	+	y_8

Определяющим контрастом полуреплики является соотношение

$$1 = x_1x_2x_4.$$

Совместные оценки в этом случае будут определяться следующим образом:

$$x_1 = x_2x_4 \quad b_1 \rightarrow \beta_1 + \beta_{24};$$

$$x_2 = x_1x_4 \quad b_2 \rightarrow \beta_2 + \beta_{14};$$

$$x_3 = x_1x_2x_3x_4 \quad b_3 \rightarrow \beta_3 + \beta_{1234};$$

$$x_4 = x_1x_2 \quad b_4 \rightarrow \beta_4 + \beta_{12};$$

$$x_1x_3 = x_2x_3x_4 \quad b_{13} \rightarrow \beta_{13} + \beta_{234};$$

$$x_2x_3 = x_1x_3x_4 \quad b_{23} \rightarrow \beta_{23} + \beta_{134};$$

$$x_3x_4 = x_1x_2x_3 \quad b_{34} \rightarrow \beta_{34} + \beta_{123}.$$

В практических задачах тройные и более высокого порядка взаимодействия значительно чаще, чем двойные, бывают равны нулю и ими обычно можно пренебречь. Полуреплика 2^{4-1} , заданная генерирующим соотношением $x_4 = x_1x_2x_3$, позволяет получить отдельные оценки четырех линейных эффектов и три совместные оценки парных взаимодействий. В этом случае отдельными оценками будут b_1, b_2, b_3 и b_4 , так как тройными взаимодействиями $\beta_{234}, \beta_{134}, \beta_{124}$ и β_{123} вследствие их незначимости можно пренебречь. В полуреплике, заданной генерирующим соотношением $x_4 = x_1x_2$, три линейных эффекта, а именно b_1, b_2, b_4 - оказались смешанными с парными взаимодействиями. Разрешающая способность полуреплики, заданной генерирующим соотношением $x_4 = x_1x_2x_3$, получилась значительно выше, чем у полуреплики, заданной генерирующим соотношением $x_4 = x_1x_2$. Следовательно, разрешающая способность полуреплики зависит от генерирующего соотношения, которым она задана.

Для оценки разрешающей способности реплик (большой дробности ($1/4, 1/8$ и т. д.) используют обобщающие определяющие контрасты. $1/4$ -реплика 2^{5-2} может быть задана следующими генерирующими соотношениями: $x_4 = x_1x_2x_3, x_5 = x_2x_3$. Матрица планирования этой реплики представлена табл. 9.

Таблица 9
Матрица планирования 2^{5-2}

Номер экспер.	x_0	x_1	x_2	x_3	x_4	x_5	y
1	+	+	+	-	-	-	y_1
2	+	-	+	-	+	-	y_2
3	+	-	-	-	+	+	y_3
4	+	+	-	-	-	+	y_4
5	+	-	+	+	+	+	y_5
6	+	+	+	+	-	+	y_6
7	+	+	-	+	-	-	y_7
8	+	-	-	+	+	-	y_8

Таблица 10
Матрица планирования 2^{7-4}

Номер экспер.	x_0	x_1	x_2	x_3	x_4	x_5	x_6	x_7	y
1	+	+	-	+	-	-	+	-	y_1
2	+	-	-	+	+	+	-	-	y_2
3	+	-	+	+	+	+	+	+	y_3
4	+	+	+	+	-	-	-	+	y_4
5	+	-	-	-	+	-	-	+	y_5
6	+	+	-	-	-	+	+	+	y_6
7	+	+	+	-	-	+	-	-	y_7
8	+	-	+	-	+	-	+	-	y_8

Определяющими контрастами реплики являются соотношения

$$1 = x_1x_2x_3x_4; \quad 1 = x_2x_3x_5.$$

Перемножив определяющие контрасты, получим третье соотношение

$$1 = x_1x_4x_5.$$

Полная характеристика разрешающей способности рассматриваемой реплики будет определяться обобщающим определяющим контрастом, имеющим вид

$$1 = x_1x_2x_3x_4 = x_2x_3x_5 = x_1x_4x_5.$$

Схему смешивания оценок находим последовательным умножением обобщающего определяющего контраста на x_1, x_2, x_3 и т. д.

$$\begin{aligned} x_1 &= x_2x_3x_4 = x_1x_2x_3x_5 = x_4x_5 & b_1 &\rightarrow \beta_1 + \beta_{234} + \beta_{1235} + \beta_{45}; \\ x_2 &= x_1x_3x_4 = x_3x_5 = x_1x_2x_4x_5 & b_2 &\rightarrow \beta_2 + \beta_{134} + \beta_{35} + \beta_{1245}; \\ x_3 &= x_1x_2x_4 = x_2x_5 = x_1x_3x_4x_5 & b_3 &\rightarrow \beta_3 + \beta_{124} + \beta_{25} + \beta_{1345}; \end{aligned}$$

$$\begin{array}{ll}
x_4 = x_1x_2x_3 = x_2x_3x_4x_5 = x_1x_5 & b_4 \rightarrow \beta_4 + \beta_{123} + \beta_{2345} + \beta_{15}; \\
x_5 = x_1x_2x_3x_4x_5 = x_2x_3 = x_1x_4 & b_5 \rightarrow \beta_5 + \beta_{12345} + \beta_{23} + \beta_{14}; \\
x_1x_2 = x_3x_4 = x_1x_3x_5 = x_2x_4x_5 & b_{12} \rightarrow \beta_{12} + \beta_{34} + \beta_{135} + \beta_{245}; \\
x_1x_3 = x_2x_4 = x_1x_2x_5 = x_3x_4x_5 & b_{13} \rightarrow \beta_{13} + \beta_{24} + \beta_{125} + \beta_{345}.
\end{array}$$

Для 1/16 реплики генерирующими соотношениями

$$x_4 = x_1x_2x_3; \quad x_5 = x_1x_2; \quad x_6 = x_1x_3; \quad x_7 = x_2x_3$$

матрица планирования представлена табл. 10. Определяющими контрастами этой реплики будут соотношения

$$1) 1 = x_1x_2x_3x_4; \quad 2) 1 = x_1x_2x_5; \quad 3) 1 = x_1x_3x_6; \quad 4) 1 = x_2x_3x_7.$$

Если попарно перемножить определяющие контрасты $1x_2; 1x_3; 1x_4; 2x_3; 2x_4; 3x_4$, то получим

$$\begin{array}{llll}
1 = x_3x_4x_5; & 1 = x_2x_4x_6; & 1 = x_1x_4x_7; & 1 = x_2x_3x_5x_6; \\
& 1 = x_1x_3x_5x_7; & 1 = x_1x_2x_6x_7. &
\end{array}$$

Произведения определяющих контрастов по три: $1x_2x_3; 1x_2x_4; 2x_3x_4; 1x_3x_4$ - будут равны

$$\begin{array}{ll}
1 = x_1x_4x_5x_6; & 1 = x_2x_4x_5x_7; \\
1 = x_5x_6x_7; & 1 = x_3x_4x_6x_7.
\end{array}$$

Умножая определяющие контрасты по четыре, получим

$$1 = x_1x_2x_3x_4x_5x_6x_7.$$

Чтобы полностью характеризовать разрешающую способность данной реплики, запишем обобщающий определяющий контраст

$$\begin{aligned}
1 = x_1x_2x_3x_4 = x_1x_2x_5 = x_1x_3x_6 = x_2x_3x_7 = x_3x_4x_5 = x_2x_4x_6 = x_1x_4x_7 = x_2x_3x_5x_6 = x_1x_3x_5x_7 = \\
x_1x_2x_6x_7 = x_1x_4x_5x_6 = x_2x_4x_5x_7 = x_5x_6x_7 = x_3x_4x_6x_7 = x_1x_2x_3x_4x_5x_6x_7.
\end{aligned}$$

Если эффектами взаимодействия, начиная с тройных, можно пренебречь, то коэффициенты будут оценками:

$$\begin{aligned}
b_1 \rightarrow \beta_1 + \beta_{25} + \beta_{36} + \beta_{47}; \quad b_2 \rightarrow \beta_2 + \beta_{15} + \beta_{37} + \beta_{46}; \quad b_3 \rightarrow \beta_3 + \beta_{16} + \beta_{27} + \beta_{45}; \\
b_4 \rightarrow \beta_4 + \beta_{35} + \beta_{26} + \beta_{17}; \quad b_5 \rightarrow \beta_5 + \beta_{12} + \beta_{34} + \beta_{67}; \quad b_6 \rightarrow \beta_6 + \beta_{13} + \beta_{24} + \beta_{57}; \\
b_7 \rightarrow \beta_7 + \beta_{23} + \beta_{14} + \beta_{56}.
\end{aligned}$$

Таким образом, получаем весьма сложную систему смешивания. Все линейные эффекты оказались смешанными с несколькими парными взаимодействиями, поэтому разрешающая способность этой дробной реплики очень низкая. Пользоваться такой репликой можно лишь в том случае, если все парные взаимодействия близки к нулю.

Выбор дробной реплики зависит от конкретной задачи. Для получения линейной модели рекомендуют выбирать дробные реплики с возможно большей разрешающей способностью, т. е. реплики, у которых линейные эффекты смешаны с эффектами взаимодействия близкими к нулю. При выборе дробной реплики важно учитывать насыщенность плана, т. е. соотношение между числом опытов и числом коэффициентов, определяемых по результатам этих экспериментов. **Дробная реплика, полученная заменой всех эффектов взаимодействия новыми факторами, называется насыщенной.** Применение насыщенных планов требует минимального числа экспериментов. Число экспериментов в матрице насыщенной дробной репли-

ки равно числу коэффициентов линейной модели. Гипотезу адекватности модели в этом случае проверить невозможно, так как число степеней свободы равно нулю.

Например, 1/16-реплика от полного факторного эксперимента 2^7 (табл. 10) является насыщенной, так как линейная модель не содержит коэффициентов, которые необходимо определить по результатам восьми экспериментов. При этом не остается степеней свободы для проверки адекватности модели.

Дробные реплики широко применяют при получении линейных моделей. Эффективность применения дробных реплик зависит от удачного выбора системы смешивания линейных эффектов с эффектами взаимодействия. **При построении дробных реплик используют следующее правило: новый фактор, введенный в планирование, нужно поместить в столбец матрицы, принадлежащий взаимодействию, которым можно пренебречь.**

Лекция 5

Проведение эксперимента и обработка его результатов

После выбора плана эксперимента, основных уровней и интервалов варьирования факторов переходят к эксперименту. Каждая строка матрицы – это условия эксперимента. Для исключения систематических ошибок рекомендуется эксперименты, предусмотренные матрицей, проводить в случайной последовательности. Порядок проведения следует выбирать по таблице случайных чисел (табл. 11). Например, если требуется провести восемь экспериментов, то из случайного места таблицы последовательно выписывают числа, лежащие в интервале от 1 до 8, при этом не учитываются уже выписанные и числа больше восьми. Так, например, начиная с числа 87 (1-я строка табл. 11), получаем следующую последовательность реализации экспериментов:

Номер опыта в матрице	1	2	3	4	5	6	7	8
Порядок реализации экспериментов	7	2	8	3	1	4	5	6

Таблица 11

Фрагмент таблицы случайных чисел																	
87	63	88	23	62	51	07	69	59	02	89	49	14	98	53	41	92	36
07	76	85	37	84	37	47	32	25	21	15	08	82	34	57	57	35	22
03	33	48	84	37	37	29	38	37	89	76	25	09	69	44	61	88	23
13	01	59	47	64	04	99	59	96	20	30	87	31	33	69	45	58	48
00	83	48	94	44	08	67	79	41	61	41	15	60	11	88	83	24	82
24	07	78	61	89	42	58	88	22	16	13	24	40	09	00	65	46	38
61	12	90	62	41	11	59	85	18	42	61	29	88	76	04	21	80	78
27	84	05	99	85	75	67	80	05	57	05	71	70	21	31	99	99	06
96	53	99	25	13	63												

Для компенсации влияния случайных погрешностей каждый эксперимент рекомендуется повторить n раз. **Эксперименты, повторенные несколько раз при од-**

них и тех же значениях факторов, называют параллельными. Под дублированием понимают постановку параллельных экспериментов. Обычно число n параллельных экспериментов принимают равным 2-3, иногда – 4-5. При проведении исследований приходится иметь дело с тремя вариантами дублирования экспериментов:

1. с равномерным дублированием экспериментов;
2. с неравномерным дублированием экспериментов;
3. без дублирования экспериментов.

При *равномерном дублировании* все строки матрицы планирования имеют одинаковые числа параллельных экспериментов. В случае *неравномерного дублирования* числа параллельных экспериментов неодинаковы. При отсутствии дублирования параллельные эксперименты не проводятся. Наиболее предпочтительным из трех вариантов дублирования является первый. При этом варианте эксперимент отличается повышенной точностью, а математическая обработка экспериментальных данных - простотой. Характер дублирования влияет на содержание математической обработки результатов наблюдений. Рассмотрим методику обработки результатов эксперимента для каждого из трех вариантов дублирования.

Обработка результатов эксперимента при равномерном дублировании.

Для каждой строки матрицы планирования по результатам n параллельных экспериментов находят \bar{y}_j среднее арифметическое значение параметра оптимизации:

$$\bar{y}_j = \frac{1}{n} \sum_{u=1}^n y_{ju},$$

где u - номер параллельного эксперимента; y_{ju} - значение параметра оптимизации в u -м параллельном эксперименте j -й строки матрицы.

С целью оценки отклонений параметра оптимизации от его среднего значения для каждой строки матрицы планирования вычисляют дисперсию s_j^2 эксперимента по данным n параллельных экспериментов. Статистической дисперсией называют среднее значение квадрата отклонений случайной величины от ее среднего значения:

$$s_j^2 = \frac{1}{n-1} \sum_{u=1}^n (y_{ju} - \bar{y}_j)^2. \quad (4)$$

Ошибка s_j эксперимента определяется как корень квадратный из дисперсии

$$s_j = +\sqrt{\frac{1}{n-1} \sum_{u=1}^n (y_{ju} - \bar{y}_j)^2}.$$

В этом случае ошибка при большом рассеянии будет значительной. *Рассеяние результатов эксперимента определяется влиянием неуправляемых факторов, погрешностями измерений и другими причинами.* Большое рассеяние изучаемой величины может произойти из-за наличия в эксперименте сомнительных результатов. Для проверки сомнительных, т. е. резко выделяющихся результатов, используют специальные критерии; одним из таких критериев является отношение U (ГОСТ 11.002-73). Чтобы оценить принадлежность резко выделяющихся результатов $y_{j \max}$

или $y_{j \min}$ к данной нормальной совокупности и принять решение об исключении или оставлении их в составе выборки, находят отношение

$$U_{\max} = \frac{y_{j \max} - \bar{y}_j}{s_j} \quad \text{или} \quad U_{\min} = \frac{\bar{y}_j - y_{j \min}}{s_j},$$

где $y_{j \max}$ - наибольшее значение параметра оптимизации среди его значений, полученных в n параллельных экспериментах j -й строки матрицы планирования; $y_{j \min}$ - наименьшее значение параметра оптимизации среди его значений, полученных в n параллельных экспериментах j -й строки матрицы планирования.

Результат сравнивают с величиной β , взятой из ГОСТ 11.002-73 (табл. 1) для числа n параллельных экспериментов и принятого уровня значимости α . Число n параллельных экспериментов и объем выборки n в рассматриваемом случае понятия равноценные. Если $U_{\max} \geq \beta$, то сомнительный результат может быть исключен, в противном случае его считают нормальным и не исключают.

Аналогично производится оценка результата $y_{j \min}$: если $U_{\min} \geq \beta$, то сомнительный результат признают аномальным; при $U_{\min} < \beta$ подозреваемый в аномальности результат считают нормальным. Чтобы числа параллельных экспериментов были одинаковы во всех строках матрицы, необходимо повторить те, результаты которых были признаны аномальными. В математической статистике для проверки гипотез пользуются критериями согласия. Для того чтобы принять или забраковать гипотезу при помощи этих критериев, устанавливают уровни значимости их. *Уровень значимости представляет собой достаточно малое значение вероятности, отвечающее событиям, которые в данной обстановке исследования можно считать практически невозможными.*

Обычно принимают 5%-, 2%- или 1%-ный уровень значимости. В технике чаще всего принимают 5%-ный уровень. *Уровень значимости α называют также уровнем риска или доверительным уровнем вероятности*, который соответственно может быть принят равным 0,05, 0,02 или 0,01. Так, например, при уровне значимости (риска) $\alpha = 0,05$ вероятность P верного ответа при проверке нашей гипотезы $P = 1 - \alpha = 1 - 0,05 = 0,95$, или 95%. Это значит, что в среднем только в 5% случаев возможна ошибка при проверке гипотезы. После вычисления по формуле (4) дисперсий проверяют гипотезу их однородности. Проверка однородности двух дисперсий производится с помощью *F-критерия Фишера*, который представляет собой отношение большей дисперсии к меньшей:

$$F = \frac{s_1^2}{s_2^2},$$

где $s_1^2 > s_2^2$.

Если наблюдаемое значение F_p -критерия меньше табличного F_T (табл. 12) для соответствующих чисел степеней свободы и принятого уровня значимости, то дисперсии однородны. ***Однородность ряда дисперсий проверяют по критерию Кохрена или по критерию Бартлета.*** При равномерном дублировании экспериментов однородность ряда дисперсий проверяют с помощью *G-критерия Кохрена*,

представляющего собой отношение максимальной дисперсии к сумме всех дисперсий:

$$G_p = \frac{s_{\max}^2}{s_1^2 + s_2^2 + \dots + s_N^2} = s_{\max}^2 / \sum_{j=1}^N s_j^2.$$

Дисперсии однородны, если расчетное значение G_p -критерия не превышает табличного значения G_m -критерия. В табл. 13 - N показывает число сравниваемых дисперсий, а n - число параллельных опытов. Если $G_p > G_t$, то дисперсии неоднородны, а это указывает на то, что исследуемая величина y не подчиняется нормальному закону. В этом случае нужно попытаться заменить y случайной величиной $q = f(y)$, достаточно близко следующей нормальному закону. Если дисперсии s_j^2 экспериментов однородны, то дисперсию s_y^2 воспроизводимости вычисляют по зависимости

$$s_y^2 = \frac{1}{N} \sum_{j=1}^N s_j^2, \quad (5)$$

где N - число экспериментов или число строк матрицы планирования.

Таблица 12

Значения F -критерия Фишера при 5 %- ном уровне значимости

Число степеней свободы для меньшей дисперсии	Значения критерия при числе степеней свободы для большей дисперсии								
	1	2	3	4	5	6	12	24	≈
1	164,4	199,5	215,7	224,6	230,2	234,0	224,9	249,0	254,3
2	18,5	19,2	19,2	19,3	19,3	19,3	19,4	19,4	19,5
3	10,1	9,6	9,3	9,1	9,0	8,9	8,7	8,6	8,5
4	7,7	6,9	6,6	6,4	6,3	6,2	5,9	5,8	5,6
5	6,6	5,8	5,4	5,2	5,1	5,0	4,7	4,5	4,4
6	6,0	5,1	4,8	4,5	4,4	4,3	4,0	3,8	3,7
7	5,5	4,7	4,4	4,1	4,0	3,9	3,6	3,4	3,2
8	5,3	4,5	4,1	3,8	3,7	3,6	3,3	3,1	2,9
9	5,1	4,3	3,9	3,6	3,5	3,4	3,1	2,9	2,7
10	5,0	4,1	3,7	3,5	3,3	3,2	2,9	2,7	2,5
11	4,8	4,0	3,6	3,4	3,2	3,1	2,8	2,6	2,4
12	4,8	3,9	3,5	3,3	3,1	3,0	2,7	2,5	2,3
13	4,7	3,8	3,4	3,2	3,0	2,9	2,6	2,4	2,2
14	4,6	3,7	3,3	3,1	3,0	2,9	2,5	2,3	2,1
15	4,5	3,7	3,3	3,1	2,9	2,8	2,5	2,3	2,1
16	4,5	3,6	3,2	3,0	2,9	2,7	2,4	2,2	2,0
17	4,5	3,6	3,2	3,0	2,8	2,7	2,4	2,2	2,0
18	4,4	3,6	3,2	2,9	2,8	2,7	2,3	2,1	1,9
19	4,4	3,5	3,1	2,9	2,7	2,6	2,3	2,1	1,9
20	4,4	3,5	3,1	2,9	2,7	2,6	2,3	2,1	1,8

22	4,3	3,4	3,1	2,8	2,7	2,6	2,2	2,0	1,8
24	4,3	3,4	3,0	2,8	2,6	2,5	2,2	2,0	1,7
26	4,2	3,4	3,0	2,7	2,6	2,5	2,2	2,0	1,7
28	4,2	3,3	3,0	2,7	2,6	2,4	2,1	1,9	1,7
30	4,2	3,3	2,9	2,7	2,5	2,4	2,1	1,9	1,6
40	4,1	3,2	2,9	2,6	2,5	2,3	2,0	1,8	1,5
60	4,0	3,2	2,8	2,5	2,4	2,3	1,9	1,7	1,4
120	3,9	3,1	2,7	2,5	2,3	2,2	1,8	1,6	1,3
∞	3,8	3,0	2,6	2,4	2,2	2,1	1,8	1,5	1,0

Таблица 13

Значения G-критерия при 5 % - ном уровне значимости

N	n-1								
	1	2	3	4	5	6	7	8	9
4	0,9065	0,7679	0,6841	0,6287	0,5895	0,5598	0,5365	0,5175	0,5017
6	0,7808	0,6161	0,5321	0,4803	0,4447	0,4184	0,3980	0,3817	0,3682
8	0,6798	0,5157	0,4377	0,3910	0,3595	0,3362	0,3185	0,3043	0,2926
10	0,6020	0,4450	0,3733	0,3311	0,3029	0,2823	0,2666	0,2541	0,2439
12	0,5410	0,3924	0,3624	0,2880	0,2624	0,2439	0,2299	0,2187	0,2098
15	0,4709	0,3346	0,2758	0,2419	0,2195	0,2034	0,1911	0,1815	0,1736
20	0,3894	0,2705	0,2205	0,1921	0,1735	0,1602	0,1501	0,1422	0,1357

По результатам эксперимента вычисляют коэффициенты модели. Свободный член b_0 определяют по формуле

$$b_0 = \frac{1}{N} \sum_{j=1}^N \bar{y}_j. \quad (6)$$

Коэффициенты регрессии, характеризующие линейные эффекты, вычисляют по зависимости

$$b_i = \frac{1}{N} \sum_{j=1}^N x_{ij} \bar{y}_j. \quad (7)$$

Коэффициенты регрессии, характеризующие эффекты взаимодействия, определяют по формуле

$$b_{il} = \frac{1}{N} \sum_{j=1}^N x_{ij} x_{lj} \bar{y}_j, \quad (8)$$

где i, l - номера факторов; x_{ij}, x_{lj} - кодированные значения факторов i и l в j -м эксперименте. Формулы (6), (7), (8) получены в результате использования метода наименьших квадратов.

Коэффициенты b_0, b_i, b_{ij} - это оценки теоретических коэффициентов $\beta_0, \beta_i, \beta_{ij}$ регрессии. Оценки, найденные с помощью метода наименьших квадратов, являются наилучшими в том смысле, что они *распределены нормально со средними значениями*, равными теоретическим коэффициентам, и с наименьшими возможными дисперсиями. Вычислив коэффициенты модели, проверяют их значимость. Проверку значимости коэффициентов можно производить двумя способами:

1. сравнением абсолютной величины коэффициента с доверительным интервалом;
2. с помощью *t-критерия Стьюдента*.

При проверке значимости коэффициентов первым способом для определения доверительного интервала вычисляют дисперсии коэффициентов регрессии. Дисперсию $s^2\{b_i\}$ i -го коэффициента определяют по зависимости

$$s^2\{b_i\} = \frac{1}{nN} s_y^2. \quad (9)$$

Доверительный интервал Δb_i - находят по формуле

$$\Delta b_i = \pm t_\tau s\{b_i\}, \quad (10)$$

где t_τ - табличное значение критерия при принятом уровне значимости и числе степеней свободы f , с которым определялась дисперсия s_y^2 ; при равномерном дублировании экспериментов число степеней свободы находится по зависимости

$$f = (n - 1) N,$$

где N - число экспериментов в матрице планирования, а n - число параллельных экспериментов; $s\{b_i\}$ - ошибка в определении i -го коэффициента регрессии, вычисляемая по формуле $s\{b_i\} = +\sqrt{s^2\{b_i\}}$. Значения t приведены в табл. 14.

Таблица 14

Значения t - критерия при 5% - ном уровне значимости

Число степеней свободы	1	2	3	4	5	6	7	8
Значения t	12,71	4,30	3,18	2,78	2,57	2,45	2,37	2,30
Число степеней свободы	9	10	11	12	13	14	15	16
Значения t	2,26	2,23	2,20	2,18	2,16	2,14	2,13	2,12
Число степеней свободы	17	18	19	20	21	22	23	24
Значения t	2,11	2,10	2,09	2,09	2,08	2,07	2,07	2,06
Число степеней свободы	25	26	27	28	29	30	40	60
Значения t	2,06	2,06	2,05	2,05	2,05	2,04	2,02	2,00

Коэффициент значим, если его абсолютная величина больше доверительного интервала.

При проверке значимости коэффициентов вторым способом вычисляют t_p - критерий по зависимости

$$t_p = \frac{|b_i|}{s\{b_i\}}$$

и сравнивают его с табличным t_m . *Коэффициент значим, если $t_p > t_m$ для принятого уровня значимости и числа степеней свободы, с которым определялась дисперсия s_y^2 .*

Критерий Стьюдента t вычисляют для каждого коэффициента регрессии. Статистически незначимые коэффициенты могут быть исключены из уравнения.

После расчета коэффициентов модели и проверки их значимости определяют дисперсию $s_{ад}^2$ адекватности. Остаточная дисперсия, или дисперсия адекватности, характеризует рассеяние эмпирических значений y относительно расчетных \hat{y} , определенных по найденному уравнению регрессии. Дисперсию адекватности определяют по формуле

$$s_{ад}^2 = \frac{n \sum_{j=1}^N (\bar{y}_j - \hat{y}_j)^2}{f} = \frac{n \sum_{j=1}^N (\bar{y}_j - \hat{y}_j)^2}{N - (k + 1)}, \quad (16.11)$$

где \bar{y}_j - среднее арифметическое значение параметра оптимизации в j -м эксперименте; \hat{y}_j - значение параметра оптимизации, вычисленное по модели для условий j -го опыта; f - число степеней свободы, равное $N - (k + 1)$; k - число факторов.

Проверку **гипотезы адекватности** найденной модели производят по F -критерию Фишера:

$$F_p = \frac{s_{ад}^2}{s_y^2}. \quad (12)$$

Если значение $F_p < F_m$ для принятого уровня значимости и соответствующих чисел степеней свободы, то модель считают адекватной. При $F_p > F_m$ гипотеза адекватности отвергается.

Таким образом, обработка результатов эксперимента при равномерном дублировании экспериментов может быть представлена следующей схемой:

1) для каждой строки матрицы планирования по формуле

$$\bar{y} = \frac{1}{n} \sum_{u=1}^n y_{ju}$$

вычисляют среднее арифметическое значение \bar{y}_j параметра оптимизации;

2) по формуле (4) определяют дисперсию s_j^2 каждого опыта матрицы планирования;

3) используя критерий Кохрена, проверяют гипотезу однородности дисперсий s_j^2 опытов;

- 4) если дисперсии опытов однородны, то по формуле (5) вычисляют дисперсию s_y^2 воспроизводимости эксперимента;
- 5) по формулам (6), (7), (8) определяют коэффициенты уравнения регрессии;
- 6) по зависимости (9) находят дисперсии $s^2\{b_i\}$ коэффициентов регрессии;
- 7) по формуле (10) устанавливают величину доверительного интервала Δb_i ;
- 8) проверяют статистическую значимость коэффициентов регрессии;
- 9) по зависимости (11) определяют дисперсию $s_{ад}^2$ адекватности;
- 10) с помощью F -критерия проверяют гипотезу адекватности модели.

В заключение необходимо отметить, что использование критериев Кохрена, Стьюдента и Фишера предполагает нормальное распределение результатов эксперимента.

Обработка результатов эксперимента при неравномерном дублировании.

Результаты отдельных экспериментов иногда получаются ошибочными, и их приходится исключать. *Вследствие этого числа параллельных экспериментов оказываются неодинаковыми.* Бывают и другие случаи, когда по тем или иным причинам не удастся провести одинаковое число параллельных опытов в каждом из основных экспериментов. При неодинаковых числах параллельных экспериментов нарушается ортогональность матрицы планирования и, как следствие, изменяются формулы для определения коэффициентов регрессии и их ошибок. Расчет коэффициентов регрессии и их ошибок при неодинаковых числах параллельных опытов усложняется.

Обработка результатов эксперимента при неравномерном дублировании производится по следующей схеме:

1. Для каждой строки матрицы планирования находят \bar{y}_j - среднее арифметическое значение параметра оптимизации

$$\bar{y}_j = \frac{1}{n_j} \sum_{u=1}^{n_j} y_{ju},$$

где n_j - число параллельных экспериментов в j -й строке матрицы.

2. Для каждой строки матрицы вычисляют дисперсию s_j^2 эксперимента;

$$s_j^2 = \frac{1}{n_j - 1} \sum_{u=1}^{n_j} (y_{ju} - \bar{y}_j)^2.$$

3. Проверяют с помощью критерия Бартлета гипотезу однородности дисперсий.

Для этого подсчитывают дисперсию s_y^2 воспроизводимости эксперимента по формуле

$$s_y^2 = \frac{1}{\sum_{j=1}^N f_j} \left(\sum_{j=1}^N s_j^2 f_j \right)$$

где f_j - число степеней свободы, с которым определялась дисперсия s_j^2 i -го эксперимента.

После этого определяют величину

$$Q = c^{-1} \left(f \lg s_y^2 - \sum_{j=1}^N f_j \lg s_j^2 \right),$$

где

$$c = 0,4343 \left[1 + \frac{1}{3(N-1)} \left(\sum_{j=1}^n \frac{1}{f_j} - \frac{1}{f} \right) \right]; \quad f = \sum_{j=1}^N f_j.$$

Бартлет показал, что величина Q приближенно подчиняется χ^2 - распределению с $(N - 1)$ степенями свободы, где N - число сравниваемых дисперсий.

Если Q меньше χ^2_{α} (табл. 15) для данного числа $(N - 1)$ степеней свободы и принятого уравнения значимости, то дисперсии однородны, и наоборот. *Критерий Бартлета* основан на нормальном распределении. Если распределение случайной величины не подчиняется нормальному закону, то проверка однородности дисперсий может привести к ошибочным результатам.

Таблица 15

Значения χ^2 при 5 % - ном уровне значимости

Число степеней свободы	Значения χ^2	Число степеней свободы	Значения χ^2
1	3,84	16	26,3
2	5,99	17	27,6
3	7,82	18	28,9
4	9,49	19	30,1
5	11,07	20	31,4
6	12,59	21	32,7
7	14,07	22	33,9
8	15,51	23	35,2
9	16,92	24	36,4
10	18,31	25	37,7
11	19,68	26	38,9
12	21,0	27	40,1
13	22,4	28	41,3
14	23,7	29	42,6
15	25,0	30	43,8

Рассмотрим применение *критерия Бартлета* для проверки однородности дисперсий. Матрица планирования предусматривала выполнение четырех экспериментов. Первый был повторен пять раз, второй - шесть, третий и четвертый - по четыре раза. При этом дисперсия первого эксперимента равна 3,5; второго – 4,22; третьего

— 5,88; четвертого — 11,36. Необходимо проверить, верна ли гипотеза об однородности дисперсий.

Дисперсия s_y^2 параметра оптимизации

$$s_y^2 = \frac{1}{\sum_{j=1}^N f_j} \left(\sum_{j=1}^N s_j^2 f_j \right) = \frac{3,5 \cdot 4 + 4,22 \cdot 5 + 5,88 \cdot 3 + 11,36 \cdot 3}{15} = 5,79.$$

Вычисляем величину c :

$$c = 0,4343 \left[1 + \frac{1}{3(4-1)} \left(\frac{1}{4} + \frac{1}{5} + \frac{1}{3} + \frac{1}{3} - \frac{1}{15} \right) \right] = 0,485.$$

Определяем Q :

$$Q = \frac{1}{c} \left(f \lg s_y^2 - \sum_{j=1}^N f_j \lg s_j^2 \right) = \\ = \frac{1}{0,485} [15 \lg 5,79 - 4 \lg 3,5 - 5 \lg 4,22 - 3 \lg 5,88 - 3 \lg 11,36] = 1,37$$

Табличное значение χ^2 для трех степеней свободы ($N-1=3$) и 5 % уровня значимости равно 7,82. Так как $Q < \chi^2$, то гипотеза однородности дисперсий принимается.

4. Вычисляют коэффициенты b_i уравнения регрессии, дисперсии $s^2\{b_i\}$ коэффициентов регрессии и ошибки $s\{b_i\}$ в определении коэффициентов.

5. Для каждого коэффициента регрессии находят расчетное значение t -критерия

$$t_p = \frac{\{b_i\}}{s\{b_i\}}.$$

Сравнивают расчетное значение t_p с табличным значением t_m критерия. Табличное значение критерия находят для принятого уровня значимости и числа степеней свободы f , которое в рассматриваемом случае определяют по зависимости

$$f = \sum_{j=1}^N f_j = \sum_{j=1}^N (n_j - 1).$$

Коэффициент значим при $t_p > t_m$ и незначим при $t_p < t_m$. Статистически незначимые коэффициенты могут быть исключены из уравнения регрессии. При исключении статистически незначимых коэффициентов из уравнения оставшиеся коэффициенты пересчитывают с использованием метода наименьших квадратов.

6. Определяют дисперсию адекватности

$$s_{ад}^2 = \frac{\sum_{j=1}^N n_j (\bar{y}_j - \hat{y}_j)^2}{f} = \frac{\sum_{j=1}^N n_j (\bar{y}_j - \hat{y}_j)^2}{N - (k+1)},$$

где n_j - число параллельных экспериментов в j -й строке матрицы.

7. Проверяют гипотезу адекватности полученной модели с помощью F -критерия, используя для этого формулу (12). Если $F_p < F_m$ для принятого уровня значимости и соответствующих чисел степеней свободы, то модель считают адекватной. При $F_p > F_m$ гипотеза адекватности отвергается.

Обработка результатов эксперимента при отсутствии дублирования

Обработку результатов эксперимента в этом случае производят по следующей схеме.

1. Для вычисления дисперсии s_y^2 воспроизводимости эксперимента выполняют несколько *параллельных опытов в нулевой точке* (в центре плана). При постановке опытов в нулевой точке все факторы находятся на нулевых уровнях. По результатам исследований в центре плана вычисляют дисперсию s_y^2 воспроизводимости эксперимента

$$s_y^2 = \frac{1}{n_0 - 1} \left[\sum_{u=1}^{n_0} (y_u - \bar{y})^2 \right],$$

где n_0 - число параллельных экспериментов в нулевой точке; y_u - значение параметра оптимизации в u -м опыте; \bar{y} - среднее арифметическое значение параметра оптимизации в n_0 параллельных экспериментах.

2. Закончив эксперимент, вычисляют коэффициенты модели. Свободный член b_0 определяют по формуле

$$b_0 = \frac{1}{N} \sum_{j=1}^N y_j.$$

Коэффициенты регрессии, характеризующие линейные эффекты, вычисляют по зависимости

$$b_i = \frac{1}{N} \sum_{j=1}^N x_{ij} y_j.$$

Коэффициенты регрессии, характеризующие эффекты взаимодействия, определяют по формуле

$$b_{il} = \frac{1}{N} \sum_{j=1}^N x_{ij} x_{lj} y_j,$$

где i, l - номера факторов; j - номер строки или опыта в матрице планирования; y_j - значение параметра оптимизации в j -м опыте; x_{ij}, x_{lj} - кодированные значения (± 1) факторов i и l в j -м опыте.

3. Проверяют статистическую значимость коэффициентов уравнения регрессии. Проверку значимости коэффициентов можно производить двумя способами: 1) сравнением абсолютной величины коэффициента с доверительным интервалом; 2) с помощью t -критерия.

При проверке значимости коэффициентов первым способом для определения доверительного интервала вычисляют дисперсии коэффициентов регрессии по зависимости

$$s^2\{b_i\} = \frac{1}{N} s_y^2, \quad (13)$$

где $s^2\{b_i\}$ - дисперсия i -го коэффициента регрессии; N - число строк или опытов в матрице планирования.

Из формулы (13) следует, что дисперсии всех коэффициентов равны. Доверительный интервал Δb_i определяют по формуле (10). Значение t -критерия, входящего в эту формулу, находят по таблице для принятого уровня значимости и числа степеней свободы f , которое определяют по зависимости $f = n_0 - 1$. Коэффициент регрессии значим, если его абсолютная величина больше доверительного интервала. При проверке значимости коэффициентов вторым способом вычисляют критерий t_p

$$t_p = \frac{|b_i|}{s\{b_i\}}$$

и сравнивают его с табличным t_m . Коэффициент значим, если $t_p > t_m$ для принятого уровня значимости и числа степеней свободы, определенного по формуле $f = n_0 - 1$. Критерий Стьюдента t вычисляют для каждого коэффициента регрессии. Статистически незначимые коэффициенты регрессии могут быть исключены из уравнения.

4. Определяют дисперсию s_a^2 , адекватности по формуле

$$s_{ад}^2 = \frac{\sum_{j=1}^N (y_j - \hat{y}_j)^2}{f} = \frac{\sum_{j=1}^N (y_j - \hat{y}_j)^2}{N - (k + 1)}.$$

где y_j - значение параметра оптимизации в j -м эксперименте; \hat{y}_j - значение параметра оптимизации, вычисленное по модели для условий j -го эксперимента; f - число степеней свободы, которое для линейной модели определяется по зависимости $f = N - (k + 1)$, где k - число факторов.

5. Проверяют гипотезу адекватности модели по F -критерию, используя для определения F_p - критерия формулу (16.12).

Если $F_p < F_m$ для принятого уровня значимости и соответствующих чисел степеней свободы, то модель считают адекватной. При $F_p > F_m$ гипотеза адекватности отвергается. В этом случае для получения адекватной модели принимают одно из следующих решений:

1. переходят к планированию второго или более высокого порядка;
2. уменьшают интервалы варьирования и ставят новый эксперимент, повторяя эти действия до получения адекватной линейной модели.

Если линейная модель адекватна, то переходят к методу крутого восхождения. Необходимо заметить, что крутое восхождение эффективно тогда, когда *все коэффициенты при факторах значимы*. Незначимость некоторых коэффициентов может получиться вследствие:

1. неудачно выбранных интервалов варьирования;

2. включения факторов, не влияющих на параметр оптимизации;
3. большой ошибки эксперимента.

Принятие решения в данной ситуации зависит от того, какая из трех указанных причин имеет место. Если имеет место первая причина, то изменяют интервалы варьирования по незначимым факторам и ставят новую серию экспериментов. Если принята вторая, то не влияющие факторы стабилизируют и исключают из экспериментов. Если принята третья причина, то увеличивают число параллельных экспериментов. Увеличение их числа приводит к уменьшению дисперсии коэффициентов и величины доверительного интервала, в результате все или часть коэффициентов могут оказаться значимыми. Возможен случай, когда все коэффициенты, кроме b_0 , незначимы, а модель адекватна. Такая ситуация чаще всего возникает из-за слишком узких интервалов варьирования или вследствие большой ошибки эксперимента. В этой ситуации возможны два решения: 1) расширение интервалов варьирования или 2) повышение точности эксперимента путем улучшения методики проведения и увеличения числа параллельных экспериментов.

Лекция 6

БАЗОВЫЕ ПОНЯТИЯ И ОПЕРАЦИИ ОБРАБОТКИ ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ (ЭД)

1. Эмпирическая функция распределения

Методы обработки ЭД опираются на базовые понятия теории вероятностей и математической статистики. К их числу относятся понятия *генеральной совокупности, выборки, эмпирической функции распределения*.

Под *генеральной совокупностью* понимают все возможные значения параметра, которые могут быть зарегистрированы в ходе неограниченного по времени наблюдения за объектом.

Такая совокупность состоит из *бесконечного множества* элементов. В результате наблюдения за объектом формируется ограниченная по объему совокупность значений параметра x_1, x_2, \dots, x_n . С формальной точки зрения такие данные представляют собой *выборку* из генеральной совокупности. Будем считать, что выборка содержит полные наработки до системных событий.

Наблюдаемые значения x_i называют *вариантами*, а их количество — *объемом выборки n* . Для того чтобы по результатам наблюдения можно было делать какие-либо выводы, выборка должна быть *репрезентативной* (представительной), т. е. правильно представлять пропорции генеральной совокупности. Это требование выполняется, если объем выборки достаточно велик, а каждый элемент генеральной совокупности имеет одинаковую вероятность попасть в выборку.

Пусть в полученной выборке значение x_1 параметра наблюдалось n_1 раз, значение x_2 — n_2 раз, значение x_k — n_k раз, $n_1 + n_2 + \dots + n_k = n$. Совокупность значений, записанных в порядке их возрастания, называют *вариационным рядом*, величины n_i — *частотами*, а их отношения к объему выборки $\omega_i = n_i/n$ — *относительными частотами* (частостями). Очевидно, что сумма относительных частот равна единице.

Под распределением понимают соответствие между наблюдаемыми вариантами и их частотами или частостями. Пусть n_x – количество наблюдений, при которых случайные значения параметра X меньше x . Частость события $X < x$ равна n_x/n . Это отношение является функцией от x и от объема выборки: $F_n(x) = n_x/n$. Величина $F_n(x)$ обладает всеми свойствами функции распределения:

$F_n(x)$ неубывающая функция, ее значения принадлежат отрезку $[0 - 1]$;

если x_1 – наименьшее значение параметра, а x_k – наибольшее, то $F_n(x) = 0$, когда $x < x_1$, и $F_n(x_k) = 1$, когда $x \geq x_k$.

Функция $F_n(x)$ определяется по ЭД, поэтому ее называют **эмпирической функцией распределения**. В отличие от эмпирической функции $F_n(x)$ функцию распределения $F(x)$ генеральной совокупности называют теоретической функцией распределения, она характеризует не частость, а вероятность события $X < x$. **Из теоремы Бернулли вытекает, что частость $F_n(x)$ стремится по вероятности к вероятности $F(x)$ при неограниченном увеличении n .** Следовательно, при большом объеме наблюдений теоретическую функцию распределения $F(x)$ можно заменить эмпирической функцией $F_n(x)$.

График эмпирической функции $F_n(x)$ представляет собой ломаную линию. В промежутках между соседними членами вариационного ряда $F_n(x)$ сохраняет постоянное значение. При переходе через точки оси x , равные членам выборки, $F_n(x)$ претерпевает разрыв, скачком возрастаая на величину $1/n$, а при совпадении l наблюдений – на l/n .

Пример 1. Построить вариационный ряд и график эмпирической функции распределения по результатам наблюдений, табл. 1.

Таблица 1

i	1	2	3	4	5	6
x_i	51	43	56	60	64	56

Решение. Построим вариационный ряд, упорядочив по возрастанию значения варианты, табл. 2.

Таблица 2

i	1	2	3	4	5	6
x_i	43	51	56	56	60	64

Искомая эмпирическая функция представлена рис. 1:

$$F_n(x) = \begin{cases} 0, & \text{при } x < 43, \\ 0,16, & \text{при } 43 \leq x < 51, \\ 0,33, & \text{при } 51 \leq x < 56, \\ 0,67, & \text{при } 56 \leq x < 60, \\ 0,84, & \text{при } 60 \leq x < 64, \\ 1, & \text{при } x \geq 64. \end{cases}$$

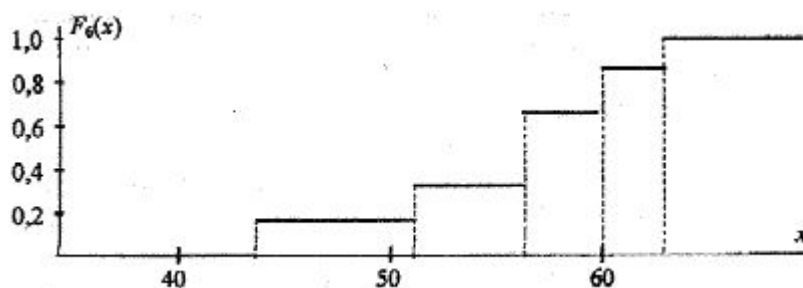


Рис. 1. Эмпирическая функция распределения

При большом объеме выборки (понятие «большой объем» зависит от целей и методов обработки, в данном случае будем считать n большим, если $n > 40$) в целях удобства обработки и хранения сведений прибегают к группированию ЭД в интервалы. Количество интервалов следует выбрать так, чтобы в необходимой мере отразилось разнообразие значений параметра в совокупности и в то же время закономерность распределения не искажалась случайными колебаниями частот по отдельным разрядам. Существуют нестрогие рекомендации по выбору количества y и размера h таких интервалов, в частности:

в каждом интервале должно находиться не менее 5 – 7 элементов. В крайних разрядах допустимо всего два элемента;

количество интервалов не должно быть очень большим или очень маленьким. Минимальное значение интервалов должно быть не менее 6 – 7. При объеме выборки, не превышающем несколько сотен элементов, количество интервалов задают в пределах от 10 до 20. Для очень большого объема выборки ($n > 1000$) количество интервалов может превышать указанные значения. Некоторые исследователи рекомендуют пользоваться соотношением $y = 1,441 \ln(n) + 1$;

при относительно небольшой неравномерности длины интервалов удобно выбирать одинаковыми и равными величине

$$h = (x_{\max} - x_{\min}) / y,$$

где x_{\max} – максимальное и x_{\min} – минимальное значение параметра. При существенной неравномерности закона распределения длины интервалов можно задавать меньшего размера в области быстрого изменения плотности распределения; при значительной неравномерности лучше в каждый разряд назначать примерно одинаковое количество элементов выборки. Тогда длина конкретного интервала будет определяться крайними значениями элементов выборки, сгруппированными в этот интервал, т.е. будет различна для разных интервалов (в этом случае при построении гистограммы нормировка по длине интервала обязательна - в противном случае высота каждого элемента гистограммы будет одинакова).

Группирование результатов наблюдений по интервалам предусматривает: определение размаха изменений параметра x ; выбор количества интервалов и их величины; подсчет для каждого i -го интервала $[x_i - x_{i+1}]$ частоты n_i или относительной частоты (частоты n_i) попадания варианты в интервал. В результате формируется представление ЭД в виде *интервального или статистического ряда*.

Графически статистический ряд отображают в виде гистограммы, полигона и ступенчатой линии. Часто *гистограмму* представляют как фигуру, состоящую из прямоугольников, основаниями которых служат интервалы длиной h , а высоты равны соответствующей частоте. Однако такой подход неточен. Высоту i -го прямоугольника z_i следует выбрать равной $n_i/(nh)$. Такую гистограмму можно интерпретировать как графическое представление эмпирической функции плотности распределения $f_n(x)$, в ней суммарная площадь всех прямоугольников составит единицу. Гистограмма помогает подобрать вид теоретической функции распределения для аппроксимации ЭД.

Полигоном называют ломаную линию, отрезки которой соединяют точки с координатами по оси абсцисс, равными серединам интервалов, а по оси ординат – соответствующим частотам. Эмпирическая функция распределения отображается ступенчатой ломаной линией: над каждым интервалом проводится отрезок горизонтальной линии на высоте, пропорциональной накопленной частоте в текущем интервале. Накопленная частота равна сумме всех частот, начиная с первого и до данного интервала включительно.

Пример 2. Имеются результаты регистрации значений затухания сигнала x_i на частоте 1000 Гц коммутируемого канала телефонной сети. Эти значения, измеренные в дБ, в виде вариационного ряда представлены в табл. 3. Необходимо построить статистический ряд.

Таблица 3

i	1	2	3	4	5	6	7	8	9	10	11
x_i	25,79	25,98	25,98	26,12	26,13	26,49	26,52	26,60	26,66	26,69	26,74
i	12	13	14	15	16	17	18	19	20	21	22
x_i	26,85	26,90	26,91	26,96	27,02	27,11	27,19	27,21	27,28	27,30	27,38
i	23	24	25	26	27	28	29	30	31	32	33
x_i	27,40	27,49	27,64	27,66	27,71	27,78	27,89	27,89	28,01	28,10	28,11
i	34	35	36	37	38	39	40	41	42	43	44
x_i	28,37	28,38	28,50	28,63	28,67	28,90	28,99	28,99	29,03	29,12	29,28

Решение. Количество разрядов статистического ряда следует выбрать минимальным, чтобы обеспечить достаточное количество попаданий в каждый из них, возьмем $y = 6$. Определим размер разряда

$$h = (x_{\max} - x_{\min})/y = (29,28 - 25,79)/6 = 0,58.$$

Сгруппируем наблюдения по разрядам, табл. 4.

Таблица 4

i	1	2	3	4	5	6
x_i	25,79	26,37	26,95	27,53	28,12	28,70
n_i	5	9	10	9	5	6
$F_i = n_i/n$	0,114	0,205	0,227	0,205	0,114	0,136
$f_i = n_i/h$	0,196	0,353	0,392	0,353	0,196	0,235

На основе статистического ряда построим гистограмму, рис. 2, и график эмпирической функции распределения, рис. 3.

График эмпирической функции распределения, рис. 3, отличается от графика, представленного на рис. 1 равенством шага изменения варианты и величиной шага

приращения функции (при построении по вариационному ряду шаг приращения кратен $1/n$, а по статистическому ряду – зависит от частоты в конкретном разряде).

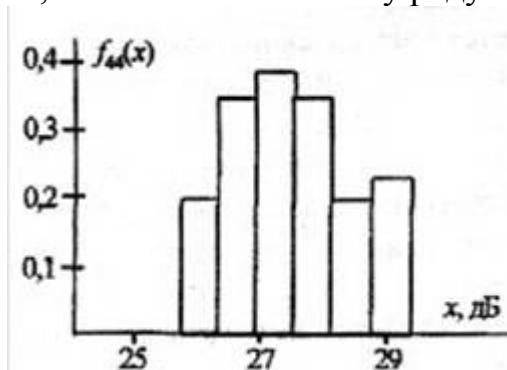


Рис. 2. Гистограмма распределения

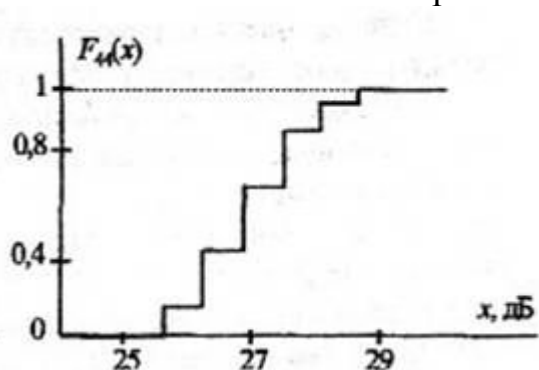


Рис. 3. Эмпирическая функция распределения

Рассмотренные представления ЭД являются исходными для последующей обработки и вычисления различных параметров.

2. Оценки параметров распределения и их свойства

Значение параметра, вычисленное по ограниченному объему ЭД, является случайной величиной, т. е. значение такой величины от выборки к выборке может меняться заранее не предвиденным образом. Следовательно, в результате обработки ЭД определяется не значение параметра θ , а только лишь его приближенное значение $\tilde{\theta}$ – статистическая оценка параметра. Получить статистическую оценку параметра теоретического распределения означает найти функцию от имеющихся результатов наблюдения, которая и даст приближенное значение искомого параметра. Различают два вида оценок – *точечные* и *интервальные*.

Точечными называют такие оценки, которые характеризуются одним числом. При малых объемах выборки точечные оценки могут значительно отличаться от истинных значений параметров, поэтому их применяют при большом объеме выборки.

Интервальные оценки задаются двумя числами, определяющими вероятный диапазон возможного значения параметра. Эти оценки применяются для малых и для больших выборок. Рассмотрим вначале точечные оценки.

Применительно к каждому оцениваемому параметру закона распределения генеральной совокупности существует множество функций, позволяющих вычислить искомые значения. Например, оценку математического ожидания можно вычислить, взяв среднее арифметическое выборочных значений, половину суммы крайних членов вариационного ряда, средний член выборки и т.д. Указанные функции отличаются качеством оценок и трудоемкостью реализации.

Качество оценок характеризуется такими свойствами, как *состоятельность*, *несмещенность*, *эффективность* и *достаточность*.

Состоятельность характеризует сходимость по вероятности оценки $\tilde{\theta}$ к истинному значению параметра θ при неограниченном увеличении объема выборки n .

- Смещение оценки должно стремиться к нулю при объеме выборки, стремящейся к бесконечности, т.е. $M(\tilde{\theta}_i) - \theta_i = \varphi_i \rightarrow 0$;

$$D(\tilde{\theta}_i) \rightarrow 0 \quad \text{при} \quad n \rightarrow \infty.$$

Оценка является состоятельной, если она удовлетворяет закону больших чисел, т.е. с увеличением объема выборки значение оценки стремится к значению параметра генеральной совокупности по вероятности:

$$P(|\tilde{\theta}_i - \theta_i| < \epsilon) \rightarrow 1, \quad \text{при} \quad n \rightarrow \infty.$$

- Для определения состоятельности оценки достаточно выполнения двух условий

Свойство состоятельности проявляется при неограниченном увеличении n , а при небольших объемах ЭД наличие этого свойства еще недостаточно для применения оценки.

Несмещенность характеризует отсутствие систематических (в среднем) отклонений оценки от параметра при любом конечном, в том числе и малом, объеме выборки, т. е. $M(\tilde{\theta}) = \theta$.

Использование статистической оценки, математическое ожидание которой не равно оцениваемому параметру, приводит к систематическим ошибкам. Не всегда наличие смещения плохо. Оно может быть существенно меньше погрешности регистрации значений параметра или давать дополнительную гарантию выполнения требований к значению параметра (если даже при положительном смещении оценка $\tilde{\theta}$ меньше предельно допустимого значения, то несмещенное значение тем более будет отвечать этому условию). В таких ситуациях допустимо применение смещенных оценок, если они вычисляются проще, чем несмещенные. Но даже несмещенная оценка может быть удалена от истинного значения.

Эффективность характеризует разброс случайных значений оценки около истинного значения параметра. Среди всех оценок следует выбрать ту, значения которой теснее сконцентрированы около оцениваемого параметра.

Для многих применяемых способов оценивания выборочные распределения параметров асимптотически нормальны, поэтому часто мерой эффективности служит **дисперсия оценки**.

Оценка $\tilde{\theta}$ является эффективной для параметра θ , если она имеет наименьшую из возможных дисперсию при заданном объеме выборки n .

Достаточность характеризует полноту использования информации, содержащейся в выборке. Другими словами, оценка $\tilde{\theta}$ будет достаточной, если все другие независимые оценки на основе данной выборки не дают дополнительной информации об оцениваемом параметре. Эффективная оценка обязательно является и достаточной.

Рассмотренные свойства применимы также и к ЭД, которые характеризуются многомерными распределениями вероятностей.

Подходы к формированию оценок разработаны в теории несмещенных оценок, предложенной А. Н. Колмогоровым и С. Рао. В данной теории предполагается известным с точностью до параметра T вид функции плотности распределения наблюдаемой величины $f(x, T)$. Вид распределения устанавливается исходя из априор-

ных соображений, например, на основе общепринятых суждений о характере безотказной работы технических средств. Тогда задача сводится к нахождению такой функции от результатов наблюдений, которая дает несмещенную и эффективную оценку.

23. Оценки моментов и квантилей распределения

Для характеристики эмпирического распределения можно использовать оценки центральных и начальных моментов. Применение находят моменты до четвертого порядка включительно, так как точность выборочных моментов резко падает с увеличением их порядка, в частности, дисперсия начальных моментов порядка p зависит от моментов порядка $2p$. Она становится значительной для моментов высокого порядка даже при больших объемах выборки. Выборочные значения моментов определяют непосредственно по выборке или по сгруппированным данным.

Пусть X – изучаемый количественный признак, x_1, x_2, \dots, x_k имеющаяся выборка объема $n = n_1 + n_2 + \dots + n_k$.

Для произвольного действительного числа x обозначим n_x – количество вариантов выборки, меньших числа x .

Относительная частота появления события ($X < x$) равна $\frac{n_x}{n}$.

Функция $F^*(x) = \frac{n_x}{n}$, определяющая для каждого $x \in n_x$ относительную частоту события ($X < x$), называется **эмпирической функцией распределения или выборочной функцией распределения**.

Функция распределения $F(x)$ генеральной совокупности называется **теоретической функцией распределения**.

Теоретическая функция распределения характеризует вероятность события ($X < x$), а эмпирическая функция характеризует относительную частоту данного события. Эмпирическая функция распределения обладает всеми свойствами теоретической функции распределения.

Рассмотрим выборочные числовые характеристики.

Выборочным средним называется число

$$\bar{x}_s = \sum_{i=1}^k x_i \frac{n_i}{n}.$$

Выборочной дисперсией называется число

$$D_s = \sum_{i=1}^k (x_i - \bar{x}_s)^2 \frac{n_i}{n}.$$

Начальный выборочный момент p -го порядка – это число

$$m_k^B = \sum_{i=1}^k (x_i)^p \cdot \frac{n_i}{n}.$$

Центральный выборочный момент p -го порядка – это число

$$m_{kC}^B = \sum_{i=1}^k (x_i - \bar{x}_s)^p \cdot \frac{n_i}{n}.$$

Важно заметить, что для интервального вариационного ряда в приведённых выше формулах вместо значений вариант x_i берётся середина соответствующего интервала \tilde{x}_i , а число n_i в этом случае есть количество наблюдаемых значений, попавших в i -ый интервал.

Главное назначение средних величин (оценок начальных моментов и в первую очередь первого момента распределения) состоит в их обобщающей функции. Это обобщение позволяет заменить множество различных индивидуальных значений показателя средней величиной, характеризующей всю однородную совокупность. Иначе говоря, средняя величина является типической характеристикой варианты в конкретной выборке. Иногда средняя величина обобщает и неоднородные совокупности данных. Например, может применяться такой показатель как среднее количество обработанных запросов на сервере в течение суток, хотя очевидно, что дневная загрузка сервера сильно отличается от загрузки в ночное время. Указанный показатель имеет смысл для оценки ресурса накопителей на жестких дисках.

Каждый элемент ЭД формируется под влиянием как общих закономерностей, так и особых условий и случайных событий. Следовательно, в обработке ЭД большой интерес представляют вопросы оценки величин, характеризующих вариацию значений параметра у разных объектов или у одного и того же объекта в разные моменты времени.

Вариацией какого-либо параметра (показателя) в совокупности наблюдений называется различие его значений у разных элементов этой совокупности.

Вариация является характерным свойством большинства информационных параметров АСОИУ. Именно это свойство является объектом исследования большинства методов обработки ЭД. Для характеристики вариации нет единого показателя, в этих целях применяются моменты распределения выше первого, производные от них величины, размах выборки, квантили и др.

Лекция 7

ПРОВЕРКА СТАТИСТИЧЕСКИХ ГИПОТЕЗ

1. Сущность задачи проверки статистических гипотез

Статистическая гипотеза представляет собой некоторое предположение о законе распределения случайной величины или о параметрах этого закона, формулируемое на основе выборки.

Примерами статистических гипотез являются предположения: **генеральная совокупность распределена по экспоненциальному закону; математические ожидания двух экспоненциально распределенных выборок равны друг другу.** В первой из них высказано предположение о виде закона распределения, а во второй – о параметрах двух распределений. Гипотезы, в основе которых нет никаких допущений о конкретном виде закона распределения, называют **непараметрическими**, в противном случае – **параметрическими**.

Гипотезу, утверждающую, что различие между сравниваемыми характеристиками отсутствует, а наблюдаемые отклонения объясняются лишь случайными ко-

лебаниями в выборках, на основании которых производится сравнение, называют **нулевой** (основной) гипотезой и обозначают H_0 .

Наряду с основной гипотезой рассматривают и **альтернативную** (конкурирующую, противоречащую) ей гипотезу H_1 . И если нулевая гипотеза будет отвергнута, то будет иметь место альтернативная гипотеза.

Различают *простые и сложные гипотезы*.

Гипотезу называют *простой*, если она однозначно характеризует параметр распределения случайной величины.

Например, если l является параметром экспоненциального распределения, то гипотеза H_0 о равенстве $l=10$ – простая гипотеза.

Сложной называют гипотезу, которая состоит из конечного или бесконечного множества простых гипотез. Сложная гипотеза H_0 о неравенстве $l>10$ состоит из бесконечного множества простых гипотез H_0 о равенстве $l=b_i$, где b_i – любое число, большее 10. Гипотеза H_0 о том, что математическое ожидание нормального распределения равно двум при неизвестной дисперсии, тоже является сложной. Сложной гипотезой будет предположение о распределении случайной величины X по нормальному закону, если не фиксируются конкретные значения математического ожидания и дисперсии.

Проверка гипотезы основывается на вычислении некоторой случайной величины – критерия, точное или приближенное распределение которого известно. Обозначим эту величину через z , ее значение является функцией от элементов выборки $z=z(x_1, x_2, \dots, x_n)$.

Процедура проверки гипотезы предписывает каждому значению критерия одно из двух решений – принять или отвергнуть гипотезу. Тем самым все выборочное пространство и соответственно множество значений критерия делятся на два непересекающихся подмножества S_0 и S_1 . Если значение критерия z попадает в область S_0 , то гипотеза принимается, а если в область S_1 , то гипотеза отклоняется.

Множество S_0 называется *областью принятия гипотезы или областью допустимых значений*, а множество S_1 – *областью отклонения гипотезы или критической областью*. Выбор одной области однозначно определяет и другую область.

Принятие или отклонение гипотезы H_0 по случайной выборке соответствует истине с некоторой вероятностью и, соответственно, возможны два рода ошибок.

Ошибка первого рода возникает с вероятностью α тогда, когда отвергается верная гипотеза H_0 и принимается конкурирующая гипотеза H_1 . *Ошибка второго рода* возникает с вероятностью β в том случае, когда принимается неверная гипотеза H_0 , в то время как справедлива конкурирующая гипотеза H_1 .

Доверительная вероятность – это вероятность не совершить ошибку первого рода и принять верную гипотезу H_0 .

Вероятность отвергнуть ложную гипотезу H_0 называется **мощностью критерия**. Следовательно, при проверке гипотезы возможны четыре варианта исходов, табл. 1.

Таблица 1.

Гипотеза H_0	Решение	Вероятность	Примечание
Верна	Принимается	$1 - \alpha$	Доверительная вероятность
	Отвергается	α	Вероятность ошибки первого рода

Неверна	Принимается	β	Вероятность ошибки второго рода
	Отвергается	$1 - \beta$	Мощность критерия

Например, рассмотрим случай, когда некоторая несмещенная оценка параметра θ вычислена по выборке объема n , и эта оценка имеет плотность распределения $f(\theta)$, рис. 1.

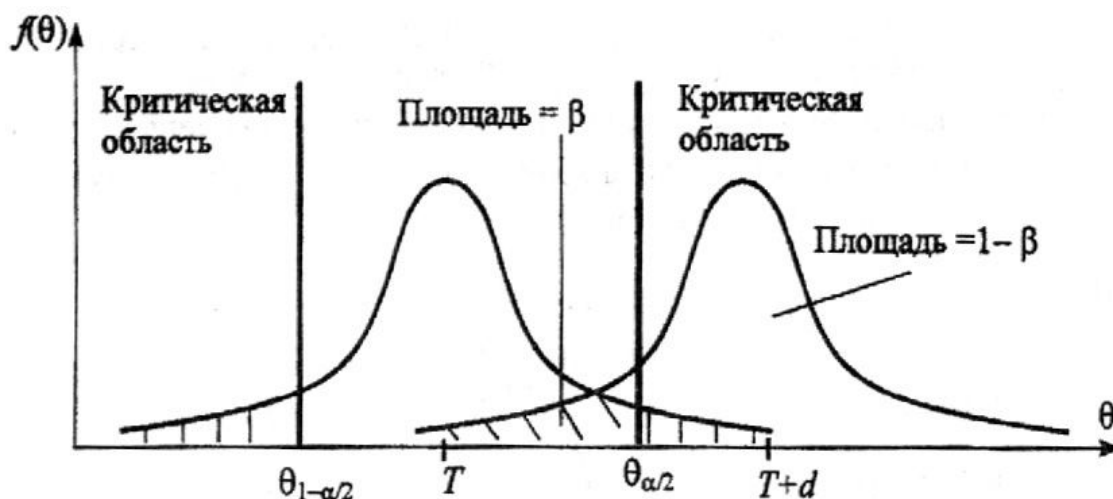


Рис. 1. Области принятия и отклонения гипотезы

Предположим, что истинное значение оцениваемого параметра равно T . Если рассматривать гипотезу H_0 о равенстве $\theta=T$, то насколько велико должно быть различие между θ и T , чтобы эту гипотезу отвергнуть. Ответить на данный вопрос можно в статистическом смысле, рассматривая вероятность достижения некоторой заданной разности между θ и T на основе выборочного распределения параметра θ .

Целесообразно полагать одинаковыми значения вероятности выхода параметра θ за нижний и верхний пределы интервала. Такое допущение во многих случаях позволяет минимизировать доверительный интервал, т.е. повысить мощность критерия проверки. Суммарная вероятность выхода параметра θ за пределы интервала с границами $\theta_{1-\alpha/2}$ и $\theta_{\alpha/2}$, составляет величину α . Эту величину следует выбирать такой, чтобы выход за пределы интервала был маловероятен. Если оценка параметра попала в заданный интервал, то нет основания подвергать сомнению проверяемую гипотезу. Гипотезу равенства $\theta=T$ можно принять. Но если после получения выборки окажется, что оценка выходит за установленные пределы, то в этом случае есть серьезные основания отвергнуть гипотезу H_0 . Отсюда следует, что вероятность допустить ошибку *первого* рода равна α (равна уровню значимости критерия).

Если предположить, например, что истинное значение параметра в действительности равно $T+d$, то согласно гипотезе H_0 о равенстве $\theta=T$ – вероятность того, что оценка параметра θ попадет в область принятия гипотезы, составит β , рис. 1.

При заданном объеме выборки вероятность совершения ошибки первого рода можно уменьшить, снижая уровень значимости α . Однако при этом увеличивается вероятность ошибки второго рода β (снижается мощность критерия). Аналогичные

рассуждения можно провести для случая, когда истинное значение параметра равно $T-d$.

Единственный способ уменьшить обе вероятности состоит в увеличении объема выборки (плотность распределения оценки параметра при этом становится более "узкой").

При выборе критической области руководствуются правилом Неймана – Пирсона: *следует так выбирать критическую область, чтобы вероятность α была мала, если гипотеза верна, и велика в противном случае*. Однако выбор конкретного значения α относительно произволен. Употребительные значения лежат в пределах от 0,001 до 0,2. В целях упрощения ручных расчетов составлены таблицы интервалов с границами $\theta_{1-\alpha/2}$ и $\theta_{\alpha/2}$ для типовых значений α и различных способов построения критерия.

При выборе уровня значимости необходимо учитывать мощность критерия при альтернативной гипотезе. Иногда большая мощность критерия оказывается существеннее малого уровня значимости, и его значение выбирают относительно большим, например 0,2. Такой выбор оправдан, если последствия ошибок второго рода более существенны, чем ошибок первого рода. Например, если отвергнуто правильное решение "продолжить работу пользователей с текущими паролями", то ошибка первого рода приведет к некоторой задержке в нормальном функционировании системы, связанной со сменой паролей. Если же принято решения не менять пароли, несмотря на опасность несанкционированного доступа посторонних лиц к информации, то эта ошибка повлечет более серьезные последствия.

В зависимости от сущности проверяемой гипотезы и используемых мер расхождения оценки характеристики от ее теоретического значения применяют различные критерии. К числу наиболее часто применяемых критериев для проверки гипотез о законах распределения относят критерии ***хи-квадрат Пирсона, Колмогорова, Мизеса, Вилкоксона, о значениях параметров – критерии Фишера, Стьюдента***.

Лекция 8

1. Типовые распределения

При проверке гипотез широкое применение находит ряд теоретических законов распределения. Наиболее важным из них является ***нормальное распределение***. С ним связаны распределения ***хи-квадрат, Стьюдента, Фишера***, а также интеграл вероятностей. Для указанных законов функции распределения аналитически не представимы. Значения функций определяются по таблицам или с использованием стандартных процедур пакетов прикладных программ. Указанные таблицы обычно построены в целях удобства проверки статистических гипотез в ущерб теории распределений – они содержат не значения функций распределения, а критические значения аргумента $z(\alpha)$.

Для односторонней критической области $z(\alpha)=z_{1-\alpha}$, т.е. критическое значение аргумента $z(\alpha)$ соответствует квантили $z_{1-\alpha}$ уровня $1-\alpha$, рис 3, так как

$$\int_z^{\infty} f(z)dz = \alpha = 1 - \int_{-\infty}^{z(\alpha)} f(z)dz$$

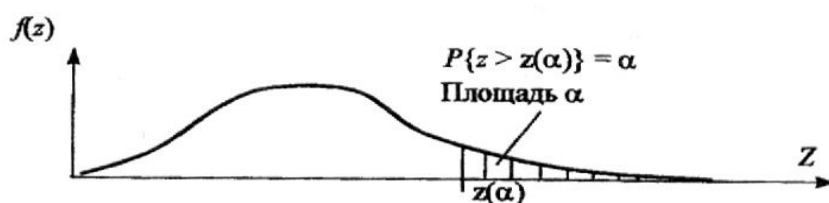


Рис. 3. Односторонняя критическая область

Для двусторонней критической области, с уровнем значимости α , размер левой области α_2 , правой α_1 ($\alpha_1 + \alpha_2 = \alpha$), рис.4. Значения $z(\alpha_2)$ и $z(\alpha_1)$ связаны с квантилями распределения соотношениями

$$z(\alpha_1) = z_{1-\alpha_1}, \quad z(\alpha_2) = z_{\alpha_2},$$

так как

$$\int_{-\infty}^{z(\alpha_1)} f(z) dz = 1 - \alpha_1,$$

$$\int_{-\infty}^{z(\alpha_2)} f(z) dz = 1 - \alpha_2$$

Для симметричной функции плотности распределения $f(z)$ критическую область выбирают из условия $\alpha_1 = \alpha_2 = \alpha/2$ (обеспечивается наибольшая мощность критерия). В таком случае левая и правая границы будут равны $|z(\alpha/2)|$.

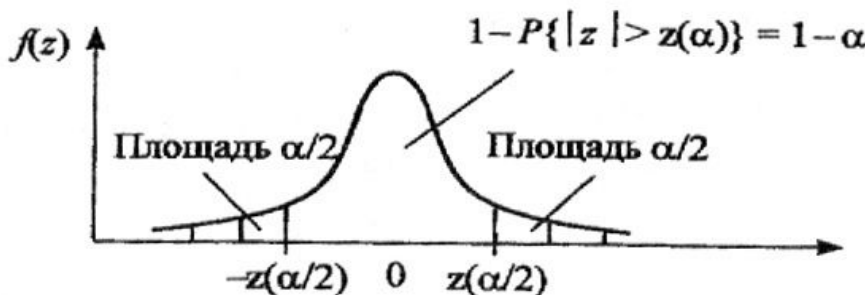


Рис. 4. Двусторонняя критическая область

Нормальное распределение

Этот вид распределения является наиболее важным в связи с центральной предельной теоремой теории вероятностей: **распределение суммы независимых случайных величин стремится к нормальному с увеличением их количества при произвольном законе распределения отдельных слагаемых, если слагаемые обладают конечной дисперсией.**

Кроме того, А.М. Ляпунов доказал, что распределение параметра стремится к нормальному, если на параметр оказывает влияние большое количество факторов и ни один из них не является преобладающим. Функция плотности нормального распределения

$$f(x) = \frac{1}{\sqrt{2\pi m_2}} \exp \left[-\frac{(x - m_1)^2}{2m_2} \right]$$

– унимодальная, симметричная, аргумент x может принимать любые действительные значения, рис. 3.5.

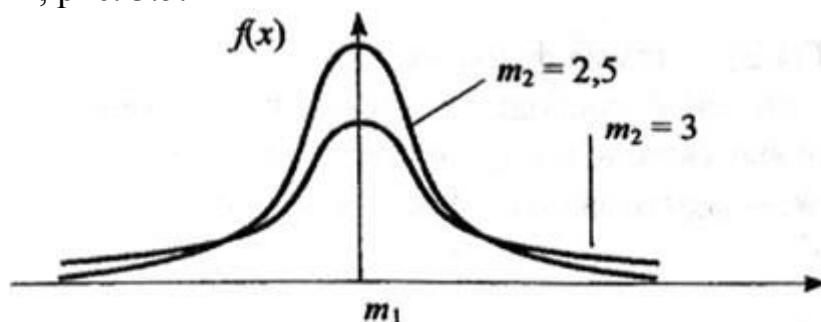


Рис. 5. Плотность нормального распределения

Функция плотности нормального распределения стандартизованной величины u имеет вид

$$f(u) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{u^2}{2}\right]$$

Вычисление значений функции распределения $\Phi(u)$ для стандартизованного неотрицательного аргумента u ($u \geq 0$) можно произвести с помощью полинома наилучшего приближения

$$\Phi(u) = 1 - 0,5(1 + 0,196854u + 0,115194u^2 + 0,000344u^3 + 0,019527u^4) - 4$$

Такая аппроксимация обеспечивает абсолютную ошибку не более 0,00025. Для вычисления $\Phi(u)$ в области отрицательных значений стандартизованного аргумента u ($u < 0$) следует воспользоваться свойством симметрии нормального распределения

$$\Phi(u) = 1 - \Phi(-u).$$

Интеграл вероятностей связан с функцией нормального распределения стандартизованной величины u соотношением

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

Распределение хи-квадрат

Распределению хи-квадрат (χ^2 -распределению) с k степенями свободы соответствует распределение суммы

$$\chi^2 = \sum_{i=1}^n u_i^2$$

квадратов n стандартизованных случайных величин u_i , каждая из которых распределена по нормальному закону, причем k из них независимы, $n \geq k$. Функция плотности распределения хи-квадрат с k степенями свободы

$$f(x) = \left[2^{k/2} \Gamma\left(\frac{k}{2}\right) \right]^{-1} (x)^{\frac{k}{2}-1} e^{-x/2}, \quad x \geq 0, \quad (4)$$

где $\Gamma(k/2)$ – гамма-функция.

Функция плотности при k , равном одному или двум, – монотонная, а при $k > 2$ – унимодальная, несимметричная, рис. 6.

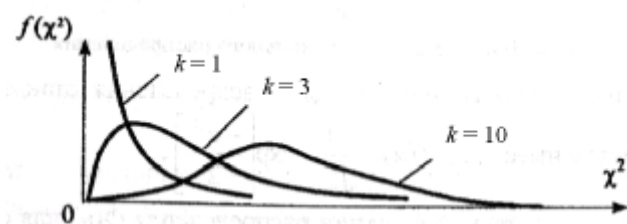


Рис. 6. Плотность распределения хи-квадрат

Математическое ожидание и дисперсия величины χ^2 равны соответственно k и $2k$. Распределение хи-квадрат является частным случаем более общего гамма-распределения, а величина, равная корню квадратному из хи-квадрат с двумя степенями свободы, подчиняется распределению Рэля.

С увеличением числа степеней свободы ($k > 30$) распределение хи-квадрат приближается к нормальному распределению с математическим ожиданием k и дисперсией $2k$.

Распределение Стьюдента

Распределение Стьюдента (t -распределение, предложено в 1908 г. английским статистиком В. Госсетом, публиковавшим научные труды под псевдонимом Student) характеризует распределение случайной величины

$$t = \frac{u_0}{\sqrt{(u_1^2 + u_2^2 + \dots + u_k^2)/k}},$$

где u_0, u_1, \dots, u_k взаимно независимые нормально распределенные случайные величины с нулевым средним и конечной дисперсией. Аргумент t не зависит от дисперсии слагаемых. Функция плотности распределения Стьюдента

$$f(t) = \frac{\Gamma[(k+1)/2]}{\sqrt{\pi k} \Gamma(k/2)} \left[1 + \frac{t^2}{k} \right]^{-\frac{(k+1)}{2}} \quad (5)$$

Величина k характеризует количество степеней свободы. Плотность распределения – унимодальная и симметричная функция, похожая на нормальное распределение, рис. 7.

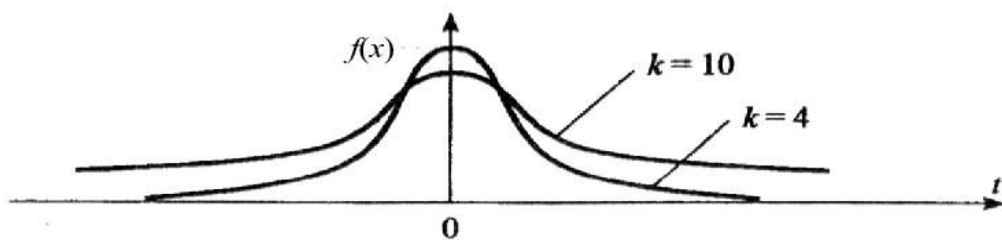


Рис. 7. Плотность распределения Стьюдента

Область изменения аргумента t от минус до плюс бесконечности. Математическое ожидание и дисперсия равны 0 и $k/(k-2)$ соответственно, при $k > 2$. По сравнению с нормальным распределением Стьюдента более пологое, оно имеет меньшую дисперсию. Это отличие заметно при небольших значениях k , что следует учитывать при проверке статистических гипотез (критические значения аргумента распределения Стьюдента превышают аналогичные показатели нормального распределения). Таблицы распределения содержат значения для односторонней (пределы интегрирования от $r(k; \alpha)$ до бесконечности)

$$\int_{r(k, \alpha)}^{\infty} f(t) dt = \alpha$$

или двусторонней (пределы интегрирования от $-r(k; \alpha)$ до $r(k; \alpha)$)

$$\int_{-r(k, \alpha)}^{r(k, \alpha)} f(t) dt = \alpha$$

критической области.

Распределение Стьюдента применяется для описания ошибок выборки при $k < 30$. При k , превышающем 100, данное распределение практически соответствует нормальному, для значений k из диапазона от 30 до 100 различия между распределением Стьюдента и нормальным распределением составляют несколько процентов.

Поэтому относительно оценки ошибок малыми считаются выборки объемом не более 30 единиц, большими – объемом более 100 единиц. При аппроксимации распределения Стьюдента нормальным распределением для односторонней критической области вероятность α

$$P\{t > t(k; \alpha)\} = u_{1-\alpha}(0, k/(k-2)),$$

где $u_{1-\alpha}(0, k/(k-2))$ – квантиль нормального распределения. Аналогичное соотношение можно составить и для двусторонней критической области.

Распределение Фишера

Распределению Р.А. Фишера (F -распределению Фишера – Снедекора) подчиняется случайная величина

$$X = [(y_1/k_1)/(y_2/k_2)],$$

равная отношению двух случайных величин y_1 и y_2 , имеющих хи-квадрат распределение с k_1 и k_2 степенями свободы. Область изменения аргумента x от 0 до бесконечности. Плотность распределения

$$f(x) = \left[\frac{k_1}{k_2} \right]^{k_1/2} \frac{\Gamma[(k_1 + k_2)/2]}{\Gamma(k_1/2)\Gamma(k_2/2)} x^{(k_1-2)/2} \left[1 + \frac{k_1}{k_2} x \right]^{-\frac{(k_1+k_2)}{2}} \quad (6)$$

В этом выражении k_1 обозначает число степеней свободы величины y_1 с большей дисперсией, k_2 – число степеней свободы величины y_2 с меньшей дисперсией. Плотность распределения – унимодальная, несимметричная, рис. 8.

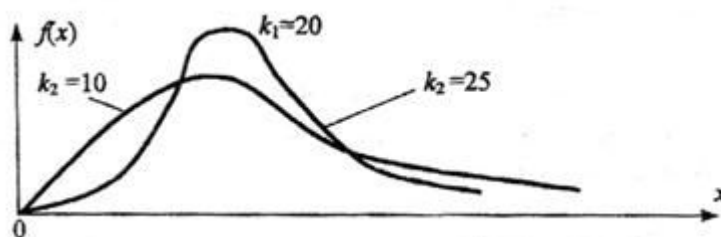


Рис. 8. Плотность распределения Фишера

Математическое ожидание случайной величины x

$$m_1 = k_2 / (k_2 - 2) \text{ при } k_2 > 2,$$

дисперсия $m_2 = [2k_2^2(k_1 + k_2 - 2)] / [k_1(k_2 - 2)^2(k_2 - 4)] \text{ при } k_2 > 4.$

При $k_1 > 30$ и $k_2 > 30$ величина x распределена приблизительно нормально: с центром распределения $(k_1 - k_2) / (2k_1 k_2)$ и дисперсией $(k_1 + k_2) / (2k_1 k_2)$.

3. Проверка гипотез о законе распределения

Обычно сущность проверки гипотезы о законе распределения ЭД заключается в следующем:

Имеется выборка ЭД фиксированного объема, выбран или известен вид закона распределения генеральной совокупности. Необходимо оценить по этой выборке параметры закона, определить степень согласованности ЭД и выбранного закона распределения, в котором параметры заменены их оценками. Пока не будем касаться способов нахождения оценок параметров распределения, а рассмотрим только вопрос проверки согласованности распределений с использованием наиболее употребительных критериев.

3.1. Критерий хи-квадрат К. Пирсона

Использование этого критерия основано на применении такой меры (статистики) расхождения между теоретическим $F(x)$ и эмпирическим распределением $F_n(x)$, ко-

торая приближенно подчиняется закону распределения χ^2 . Гипотеза H_0 о согласованности распределений проверяется путем анализа распределения этой статистики. Применение критерия требует построения *статистического* ряда.

Итак, пусть выборка представлена статистическим рядом с количеством разрядов y . Наблюдаемая частота попаданий в i -й разряд n_i . В соответствии с теоретическим законом распределения ожидаемая частота попаданий в i -й разряд составляет F_i . Разность между наблюдаемой и ожидаемой частотой составит величину $(n_i - F_i)$. Для нахождения общей степени расхождения между $F(x)$ и $F_n(x)$ необходимо подсчитать взвешенную сумму квадратов разностей по всем разрядам статистического ряда

$$\chi^2 = \sum_{i=1}^y \frac{(n_i - F_i)^2}{F_i}. \quad (7)$$

Величина χ^2 при неограниченном увеличении n имеет распределение хи-квадрат (асимптотически распределена как хи-квадрат). Это распределение зависит от числа степеней свободы k , т.е. количества независимых значений слагаемых в выражении (7). Число степеней свободы равно числу y минус число линейных связей, наложенных на выборку. Одна связь существует в силу того, что любая частота может быть вычислена по совокупности частот в оставшихся $y-1$ разрядах. Кроме того, если параметры распределения неизвестны заранее, то имеется еще одно ограничение, обусловленное подгонкой распределения к выборке. ***Если по выборке определяются f параметров распределения, то число степеней свободы составит $k = y - f - 1$.***

Очевидно, что чем меньше расхождение между теоретическими и эмпирическими частотами, тем меньше величина критерия. Область принятия гипотезы H_0 определяется условием $\chi^2 < \chi^2(k, \alpha)$, где $\chi^2(k, \alpha)$ – критическая точка распределения хи-квадрат с уровнем значимости α .

Вероятность ошибки первого рода равна α , вероятность ошибки второго рода четко определить нельзя, потому что существует бесконечно большое множество различных способов несовпадения распределений. Мощность критерия зависит от количества разрядов и объема выборки. Критерий рекомендуется применять при $n > 200$, допускается применение при $n > 40$, именно при таких условиях критерий состоятелен (как правило, отвергает неверную нулевую гипотезу).

Теоретическое значение вероятности F_i попадания случайной величины в i -й интервал определяется по формуле

$$F_i = \frac{1}{\sigma\sqrt{2\pi}} \int_{x_{i-1}}^{x_i} \exp\left(-\frac{(x - \mu_1)^2}{2\sigma^2}\right) dx. \quad (8)$$

Для нормального закона возможные значения случайной величины лежат в диапазоне от минус до плюс бесконечности, поэтому при расчетах оценок вероятностей крайний левый и крайний правый интервалы расширяются до минус и плюс бесконечности соответственно. Вычислить значения функции нормального распределения можно, воспользовавшись стандартными функциями табличного процессора или полиномом наилучшего приближения.

3.2. Критерий А.Н. Колмогорова

Для применения критерия А.Н. Колмогорова ЭД требуется представить в виде вариационного ряда (ЭД недопустимо объединять в разряды). В качестве меры расхождения между теоретической $F(x)$ и эмпирической $F_n(x)$ функциями распределения непрерывной случайной величины X используется модуль максимальной разности

$$d_n = \max |F(x) - F_n(x)| \quad (9)$$

А.Н. Колмогоров доказал, что какова бы ни была функция распределения $F(x)$ величины X при неограниченном увеличении количества наблюдений n функция распределения случайной величины $d_n \sqrt{n}$ асимптотически приближается к функции распределения

$$K(\lambda) = P(d \sqrt{n} < \lambda) = \sum_{k=-\infty}^{\infty} (-1)^k \exp(-2k^2 \lambda^2). \quad (10)$$

Иначе говоря, критерий А.Н. Колмогорова характеризует вероятность того, что величина $d_n \sqrt{n}$ не будет превосходить параметр λ для любой теоретической функции распределения. Уровень значимости α выбирается из условия

$$P(d_n \sqrt{n} > \lambda) = \alpha = 1 - K(\lambda), \quad (11)$$

в силу предположения, что почти невозможно получить это равенство, когда существует соответствие между функциями $F(x)$ и $F_n(x)$. Критерий А.Н. Колмогорова позволяет проверить согласованность распределений по малым выборкам, он проще критерия хи-квадрат, поэтому его часто применяют на практике. Но требуется учитывать два обстоятельства:

1. В соответствии с условиями его применения необходимо пользоваться следующим соотношением

$$d_n = \max(d_n^+, d_n^-),$$

где

$$d_n^+ = \max_{1 \leq j \leq n} \left| \frac{j}{n} - F(x) \right|; \quad d_n^- = \max_{1 \leq j \leq n} \left| F(x) - \frac{j-1}{n} \right|. \quad (12)$$

2. Условия применения критерия предусматривают, что теоретическая функция распределения известна полностью – известны вид функции и значения ее параметров. На практике параметры обычно неизвестны и оцениваются по ЭД. Но критерий не учитывает уменьшение числа степеней свободы при оценке параметров распределения по исходной выборке. Это приводит к завышению значения вероятности соблюдения нулевой гипотезы, т.е. повышается риск принять в качестве правдоподобной гипотезу, которая плохо согласуется с ЭД (повышается вероятность совершить ошибку второго рода).

В качестве меры противодействия такому выводу следует увеличить уровень значимости α , приняв его равным 0,1 – 0,2, что приведет к уменьшению зоны допустимых отклонений.

Пример. Проверить с помощью критерия А.Н. Колмогорова гипотезу о том, что ЭД, представленные в табл. 3, подчиняются нормальному распределению при уровне значимости $\alpha=0,1$.

Решение. Исходные данные и результаты вычислений сведены в табл. 3. Необходимые вычисления можно провести с использованием табличного процессора: значение эмпирической функции распределения $F_n(x_i)$; $i = 1, 44$; значения теоретической функции $F(x_i)$ – это значение функции нормального распределения в точке x_i .

Таблица 3

i	1	2	3	4	5	6	7	8	9	10	11
x_i	25,79	25,98	25,98	26,12	26,13	26,49	26,52	26,60	26,66	26,69	26,74
$F_n(x_i)$	0,023	0,046	0,068	0,091	0,114	0,136	0,159	0,182	0,204	0,227	0,250
$F(x_i)$	0,036	0,055	0,055	0,073	0,075	0,144	0,151	0,170	0,188	0,196	0,211
d_n^+	0,014	0,009	0,013	0,018	0,038	0,008	0,008	0,012	0,016	0,032	0,039
d_n^-	0,036	0,032	0,010	0,005	0,016	0,031	0,014	0,011	0,006	0,009	0,016
i	12	13	14	15	16	17	18	19	20	21	22
x_i	26,85	26,90	26,91	26,96	27,02	27,11	27,19	27,21	27,28	27,30	27,38
$F_n(x_i)$	0,273	0,296	0,318	0,341	0,364	0,386	0,409	0,432	0,455	0,477	0,500
$F(x_i)$	0,246	0,263	0,267	0,284	0,305	0,337	0,371	0,378	0,406	0,412	0,447
d_n^+	0,027	0,032	0,051	0,057	0,059	0,050	0,038	0,054	0,049	0,065	0,053
d_n^-	0,004	0,010	0,028	0,034	0,036	0,027	0,015	0,031	0,026	0,042	0,031
i	23	24	25	26	27	28	29	30	31	32	33
x_i	27,40	27,49	27,64	27,66	27,71	27,78	27,89	27,89	28,01	28,10	28,11
$F_n(x_i)$	0,523	0,546	0,568	0,591	0,614	0,636	0,659	0,682	0,705	0,727	0,750
$F(x_i)$	0,456	0,492	0,555	0,561	0,583	0,610	0,656	0,656	0,701	0,731	0,735
d_n^+	0,067	0,053	0,013	0,030	0,031	0,026	0,003	0,026	0,003	0,004	0,015
d_n^-	0,044	0,031	0,010	0,007	0,008	0,003	0,019	0,003	0,020	0,027	0,008
i	34	35	36	37	38	39	40	41	42	43	44
x_i	28,37	28,38	28,50	28,63	28,67	28,90	28,99	28,99	29,03	29,12	29,28
$F_n(x_i)$	0,773	0,795	0,818	0,841	0,864	0,886	0,909	0,932	0,955	0,977	1,000
$F(x_i)$	0,817	0,819	0,851	0,879	0,888	0,928	0,939	0,940	0,944	0,954	0,968
d_n^+	0,044	0,024	0,032	0,038	0,024	0,042	0,030	0,008	0,010	0,024	0,032
d_n^-	0,067	0,046	0,055	0,061	0,047	0,064	0,053	0,031	0,013	0,001	0,009

В данном примере максимальные значения d_n^+ и d_n^- одинаковы и равны 0,067. Из табл. распределения Колмогорова для $\alpha = 0,1$ найдем значение $\lambda\alpha = 1,22$. Для $n=44$ критическое значение отклонения определится из соотношения

$$d_m(\alpha = 0,1)\sqrt{n} = \lambda\alpha; \quad d_m(0,1) = \frac{1,22}{\sqrt{44}} = 0,184.$$

Поскольку величина $\max d_n = 0,067$ меньше критического значения, **гипотеза о принадлежности выборки нормальному закону не отвергается.**

Критерий Р. Мизеса

В качестве меры различия теоретической функции распределения $F(x)$ и эмпирической $F_n(x)$ по критерию Мизеса (критерию $-\omega^2$) выступает средний квадрат отклонений по всем значениям аргумента x

$$\omega_n^2 = \int_{-\infty}^{+\infty} [F_n(x) - F(x)]^2 dF(x) \quad (13)$$

Статистика критерия

$$n\omega_n^2 = \frac{1}{12n} + \sum_{i=1}^n \left[F(x_i) - \frac{i-0,5}{n} \right]^2 \quad (14)$$

При неограниченном увеличении n существует предельное распределение статистики $n\omega_n^2$. Задав значение вероятности α можно определить критические значения $n\omega_n^2(\alpha)$. Проверка гипотезы о законе распределения осуществляется обычным образом: если фактическое значение $n\omega_n^2$ окажется больше критического или равно ему, то согласно критерию Мизеса с уровнем значимости α гипотеза H_0 о том, что закон распределения генеральной совокупности соответствует $F(x)$, должна быть отвергнута.

Пример. Проверить с помощью критерия Мизеса гипотезу о том, что ЭД, представленные в предыдущем примере, подчиняются нормальному распределению при уровне значимости $\alpha = 0,1$.

Решение. Исходные данные и результаты вычислений представлены в табл. 4.

Таблица 4

i	1	2	3	4	5	6	7	8	9	10	11
x_i	25,79	25,98	25,98	26,12	26,13	26,49	26,52	26,60	26,66	26,69	26,74
$F_n(x_i)$	0,011	0,034	0,057	0,080	0,102	0,125	0,148	0,171	0,193	0,216	0,237
$F(x_i)$	0,036	0,055	0,055	0,073	0,075	0,144	0,151	0,170	0,188	0,196	0,211
$D_i \times 10^{-3}$	0,618	0,429	0,003	0,047	0,726	0,378	0,009	0,000	0,025	0,409	0,742
i	12	13	14	15	16	17	18	19	20	21	22
x_i	26,85	26,90	26,91	26,96	27,02	27,11	27,19	27,21	27,28	27,30	27,38
$F_n(x_i)$	0,261	0,284	0,307	0,330	0,352	0,375	0,398	0,421	0,443	0,466	0,489
$F(x_i)$	0,246	0,263	0,267	0,284	0,305	0,337	0,371	0,378	0,406	0,412	0,447
$D_i \times 10^{-3}$	0,231	0,439	1,572	2,071	2,243	1,467	0,717	1,790	1,391	2,866	1,755
i	23	24	25	26	27	28	29	30	31	32	33
x_i	27,40	27,49	27,64	27,66	27,71	27,78	27,89	27,89	28,01	28,10	28,11
$F_n(x_i)$	0,511	0,534	0,557	0,580	0,602	0,625	0,648	0,671	0,693	0,716	0,739
$F(x_i)$	0,456	0,492	0,555	0,561	0,583	0,610	0,656	0,656	0,701	0,731	0,735
$D_i \times 10^{-3}$	3,103	1,765	0,003	0,332	0,374	0,216	0,063	0,213	0,067	0,238	0,013
i	34	35	36	37	38	39	40	41	42	43	44
x_i	28,37	28,38	28,50	28,63	28,67	28,90	28,99	28,99	29,03	29,12	29,28
$F_n(x_i)$	0,761	0,784	0,807	0,830	0,852	0,875	0,898	0,921	0,943	0,966	0,989
$F(x_i)$	0,817	0,819	0,851	0,879	0,888	0,928	0,939	0,940	0,944	0,954	0,968
$D_i \times 10^{-3}$	3,090	1,230	1,908	2,461	1,271	2,791	1,737	0,381	0,001	0,149	0,432

В этой таблице:

$F_n(x_i) = (i-0,5)/44$ – значение эмпирической функции распределения;

$F(x_i)$ – значение теоретической функции распределения, соответствует значению функции нормального распределения в точке x_i ;

$$D_i = [F_n(x_i) - F(x_i)]^2.$$

Критическое значение статистики критерия Мизеса при заданном уровне значимости равно 0,347 (таблица распределения Мизеса). Фактическое значение статистики

$$n\omega_n^2 = \frac{1}{12 \cdot 44} + \sum_{i=1}^{44} D_i = 0,044,$$

что меньше критического значения. Следовательно, гипотеза H_0 не противоречит имеющимся данным.

Достоинством критерия Мизеса является быстрая сходимость к предельному закону, для этого достаточно не менее 40 наблюдений в области часто используемых на практике больших значений $n\omega_n^2$ (а не несколько сот, как для критерия хи-квадрат).

Сопоставляя возможности различных критериев, необходимо отметить следующие особенности:

- Критерий Пирсона устойчив к отдельным случайным ошибкам в ЭД. Однако его применение требует группирования данных по интервалам, выбор которых относительно произволен и подвержен противоречивым рекомендациям.
- Критерий Колмогорова слабо чувствителен к виду закона распределения и подвержен влиянию помех в исходной выборке, но прост в применении.
- Критерий Мизеса имеет ряд общих свойств с критерием Колмогорова: оба основаны непосредственно на результатах наблюдения и не требуют построения статистического ряда, что повышает объективность выводов; оба не учитывают уменьшение числа степеней свободы при определении параметров распределения по выборке, а это ведет к риску принятия ошибочной гипотезы. Их предпочтительно применять в тех случаях, когда параметры закона распределения известны априори, например, при проверке датчиков случайных чисел.

При проверке гипотез о законе распределения следует помнить, что *слишком хорошее совпадение с выбранным законом распределения* может быть обусловлено некачественным экспериментом («подчистка» ЭД) или предвзятой предварительной обработкой результатов (некоторые результаты отбрасываются или округляются).

Выбор критерия проверки гипотезы *относительно произволен*. Разные критерии могут давать различные выводы о справедливости гипотезы, окончательное заключение в таком случае принимается на основе неформальных соображений. Точно также *нет однозначных рекомендаций по выбору уровня значимости*.

Рассмотренный подход к проверке гипотез, основанный на применении специальных таблиц критических точек распределения, сложился в эпоху "ручной" обработки ЭД, когда наличие таких таблиц существенно снижало трудоемкость вычислений. В настоящее время математические пакеты включают процедуры вычисления стандартных функций распределений, что позволяет отказаться от использования таблиц, но может потребовать изменения правил проверки.

Лекция 7

ПРОВЕРКА СТАТИСТИЧЕСКИХ ГИПОТЕЗ

1. Сущность задачи проверки статистических гипотез

Статистическая гипотеза представляет собой некоторое предположение о законе распределения случайной величины или о параметрах этого закона, формулируемое на основе выборки.

Примерами статистических гипотез являются предположения: *генеральная совокупность распределена по экспоненциальному закону; математические ожидания двух экспоненциально распределенных выборок равны друг другу*. В первой из них высказано предположение о виде закона распределения, а во второй – о параметрах двух распределений. Гипотезы, в основе которых нет никаких допущений о конкретном виде закона распределения, называют *непараметрическими*, в противном случае – *параметрическими*.

Гипотезу, утверждающую, что различие между сравниваемыми характеристиками отсутствует, а наблюдаемые отклонения объясняются лишь случайными колебаниями в выборках, на основании которых производится сравнение, называют *нулевой* (основной) гипотезой и обозначают H_0 .

Наряду с основной гипотезой рассматривают и *альтернативную* (конкурирующую, противоречащую) ей гипотезу H_1 . И если нулевая гипотеза будет отвергнута, то будет иметь место альтернативная гипотеза.

Различают *простые и сложные гипотезы*.

Гипотезу называют *простой*, если она однозначно характеризует параметр распределения случайной величины.

Например, если l является параметром экспоненциального распределения, то гипотеза H_0 о равенстве $l=10$ – простая гипотеза.

Сложной называют гипотезу, которая состоит из конечного или бесконечного множества простых гипотез. Сложная гипотеза H_0 о неравенстве $l>10$ состоит из бесконечного множества простых гипотез H_0 о равенстве $l=b_i$, где b_i – любое число, большее 10. Гипотеза H_0 о том, что математическое ожидание нормального распределения равно двум при неизвестной дисперсии, тоже является сложной. Сложной гипотезой будет предположение о распределении случайной величины X по нормальному закону, если не фиксируются конкретные значения математического ожидания и дисперсии.

Проверка гипотезы основывается на вычислении некоторой случайной величины – критерия, точное или приближенное распределение которого известно. Обозначим эту величину через z , ее значение является функцией от элементов выборки $z=z(x_1, x_2, \dots, x_n)$.

Процедура проверки гипотезы предписывает каждому значению критерия одно из двух решений – принять или отвергнуть гипотезу. Тем самым все выборочное пространство и соответственно множество значений критерия делятся на два непересекающихся подмножества S_0 и S_1 . Если значение критерия z попадает в область S_0 , то гипотеза принимается, а если в область S_1 , то гипотеза отклоняется.

Множество S_0 называется *областью принятия гипотезы или областью допустимых значений*, а множество S_1 – *областью отклонения гипотезы или критической областью*. Выбор одной области однозначно определяет и другую область.

Принятие или отклонение гипотезы H_0 по случайной выборке соответствует истине с некоторой вероятностью и, соответственно, возможны два рода ошибок.

Ошибка первого рода возникает с вероятностью α тогда, когда отвергается верная гипотеза H_0 и принимается конкурирующая гипотеза H_1 . *Ошибка второго рода* возникает с вероятностью β в том случае, когда принимается неверная гипотеза H_0 , в то время как справедлива конкурирующая гипотеза H_1 .

Доверительная вероятность – это вероятность не совершить ошибку первого рода и принять верную гипотезу H_0 .

Вероятность отвергнуть ложную гипотезу H_0 называется **мощностью критерия**. Следовательно, при проверке гипотезы возможны четыре варианта исходов, табл. 1.

Таблица 1.

Гипотеза H_0	Решение	Вероятность	Примечание
Верна	Принимается	$1 - \alpha$	Доверительная вероятность
	Отвергается	α	Вероятность ошибки первого рода
Неверна	Принимается	β	Вероятность ошибки второго рода
	Отвергается	$1 - \beta$	Мощность критерия

Например, рассмотрим случай, когда некоторая несмещенная оценка параметра θ вычислена по выборке объема n , и эта оценка имеет плотность распределения $f(\theta)$, рис. 1.

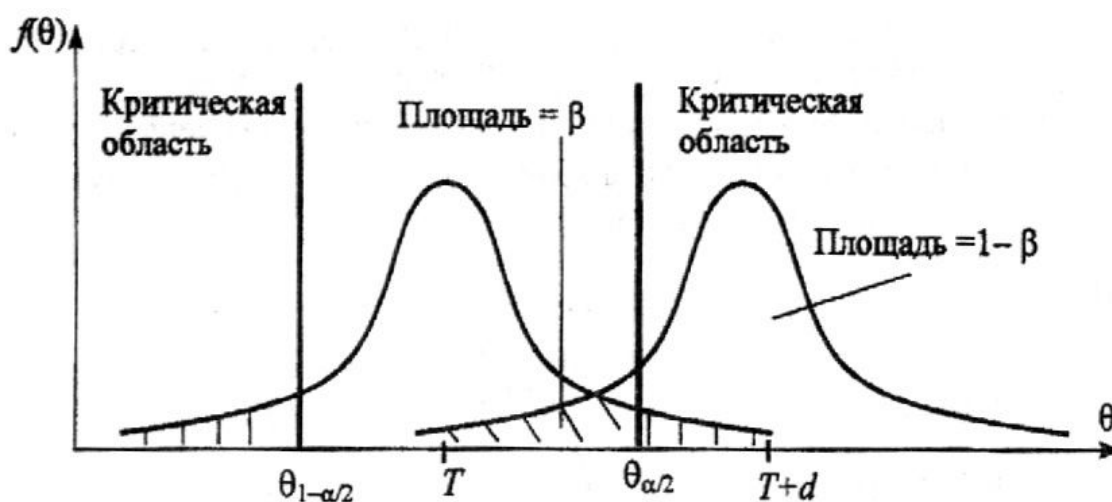


Рис. 1. Области принятия и отклонения гипотезы

Предположим, что истинное значение оцениваемого параметра равно T . Если рассматривать гипотезу H_0 о равенстве $\theta=T$, то насколько велико должно быть различие между θ и T , чтобы эту гипотезу отвергнуть. Ответить на данный вопрос можно в статистическом смысле, рассматривая вероятность достижения некоторой заданной разности между θ и T на основе выборочного распределения параметра θ .

Целесообразно полагать одинаковыми значения вероятности выхода параметра θ за нижний и верхний пределы интервала. Такое допущение во многих случаях позволяет минимизировать доверительный интервал, т.е. повысить мощность критерия проверки. Суммарная вероятность выхода параметра θ за пределы интервала с границами $\theta_{1-\alpha/2}$ и $\theta_{\alpha/2}$, составляет величину α . Эту величину следует выбирать такой, чтобы выход за пределы интервала был маловероятен. Если оценка параметра попала в заданный интервал, то нет основания подвергать сомнению проверяемую гипотезу. Гипотезу равенства $\theta=T$ можно принять. Но если после получения выборки окажется, что оценка выходит за установленные пределы, то в этом случае есть

серьезные основания отвергнуть гипотезу H_0 . Отсюда следует, что вероятность допустить ошибку *первого* рода равна α (равна уровню значимости критерия).

Если предположить, например, что истинное значение параметра в действительности равно $T+d$, то согласно гипотезе H_0 о равенстве $\theta=T$ – вероятность того, что оценка параметра θ попадет в область принятия гипотезы, составит β , рис. 1.

При заданном объеме выборки вероятность совершения ошибки первого рода можно уменьшить, снижая уровень значимости α . Однако при этом увеличивается вероятность ошибки второго рода β (снижается мощность критерия). Аналогичные рассуждения можно провести для случая, когда истинное значение параметра равно $T-d$.

Единственный способ уменьшить обе вероятности состоит в увеличении объема выборки (плотность распределения оценки параметра при этом становится более "узкой").

При выборе критической области руководствуются правилом Неймана – Пирсона: *следует так выбирать критическую область, чтобы вероятность α была мала, если гипотеза верна, и велика в противном случае*. Однако выбор конкретного значения α относительно произволен. Употребительные значения лежат в пределах от 0,001 до 0,2. В целях упрощения ручных расчетов составлены таблицы интервалов с границами $\theta_{1-\alpha/2}$ и $\theta_{\alpha/2}$ для типовых значений α и различных способов построения критерия.

При выборе уровня значимости необходимо учитывать мощность критерия при альтернативной гипотезе. Иногда большая мощность критерия оказывается существеннее малого уровня значимости, и его значение выбирают относительно большим, например 0,2. Такой выбор оправдан, если последствия ошибок второго рода более существенны, чем ошибок первого рода. Например, если отвергнуто правильное решение "продолжить работу пользователей с текущими паролями", то ошибка первого рода приведет к некоторой задержке в нормальном функционировании системы, связанной со сменой паролей. Если же принято решения не менять пароли, несмотря на опасность несанкционированного доступа посторонних лиц к информации, то эта ошибка повлечет более серьезные последствия.

В зависимости от сущности проверяемой гипотезы и используемых мер расхождения оценки характеристики от ее теоретического значения применяют различные критерии. К числу наиболее часто применяемых критериев для проверки гипотез о законах распределения относят критерии *хи-квадрат Пирсона, Колмогорова, Мизеса, Вилкоксона, о значениях параметров – критерии Фишера, Стьюдента*.

Лекция 9

МЕТОДЫ ОЦЕНКИ ПАРАМЕТРОВ РАСПРЕДЕЛЕНИЯ

1. Точечная оценка параметров распределения

Сущность задачи точечного оценивания параметров

Точечная оценка предполагает нахождение единственной числовой величины, которая и принимается за значение параметра. Такую оценку целесообразно определять в тех случаях, когда объем ЭД достаточно велик. Причем не существует единого понятия о достаточном объеме ЭД, его значение зависит от вида оцениваемого параметра (к этому вопросу вернёмся при изучении методов интервальной оценки параметров, а предварительно будем считать достаточной выборку, содержащую не менее чем 10 значений). При малом объеме ЭД точечные оценки могут значительно отличаться от истинных значений параметров, что делает их непригодными для использования.

Задача точечной оценки параметров в типовом варианте постановки состоит в следующем:

Имеется выборка наблюдений (x_1, x_2, \dots, x_n) за случайной величиной X . Выборка должна быть представительной. Объем выборки n фиксирован. Известен вид закона распределения величины X , например, в форме плотности распределения $f(T, x)$, где T – неизвестный (в общем случае векторный) параметр распределения. Параметр является неслучайной величиной.

Требуется найти оценку θ параметра T закона распределения.

Существует несколько методов решения задачи точечной оценки параметров, самыми распространёнными из которых являются методы **максимального (наибольшего) правдоподобия, моментов и квантилей.**

Метод максимального правдоподобия

Метод был предложен Р. Фишером в 1912 г. Метод основан на исследовании вероятности получения выборки наблюдений (x_1, x_2, \dots, x_n) за некоторым количественным признаком X . Значения выборки случайные, поэтому будем рассматривать их как независимые случайные величины X_1, X_2, \dots, X_n , имеющие одинаковые распределения. Если количественный признак X непрерывный, то случайные величины X_1, X_2, \dots, X_n имеют одинаковые плотности распределения вероятностей, зависящих от x и от вектора параметров распределения $\bar{\theta} = (\theta_1, \theta_2, \dots, \theta_n)$, т.е. плотности $f(x, \bar{\theta})$.

Суть метода максимального правдоподобия заключается в свойстве случайной величины реализовывать в эксперименте в основном те свои значения x_1, x_2, \dots, x_n , вероятность которых максимальна. В этом случае совместная плотность распределения вероятностей $f(x_1, x_2, \dots, x_n, \bar{\theta})$ случайных величин X_1, X_2, \dots, X_n должна быть максимальной для непрерывного признака или совместное распределение вероятностей $P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n)$ случайных величин X_1, X_2, \dots, X_n должно быть максимальным для дискретного признака.

Ввиду независимости случайных величин X_1, X_2, \dots, X_n их совместная плотность распределения будет равна произведению плотностей распределения, т.е.

$$f(x_1, x_2, \dots, x_n, \bar{\theta}) = f(x_1, \bar{\theta}) \cdot f(x_2, \bar{\theta}) \cdot \dots \cdot f(x_n, \bar{\theta}),$$

а совместная вероятность равна произведению вероятностей событий

$$P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n, \bar{\theta}) = p_{x_1}(\bar{\theta}) \cdot \dots \cdot p_{x_n}(\bar{\theta}).$$

Функцией правдоподобия $L(x_1, x_2, \dots, x_n, \bar{\theta})$ называется функция, которая в случае непрерывного количественного признака определяется по формулам: для непрерывного признака

$$L(x_1, x_2, \dots, x_n, \bar{\theta}) = f(x_1, \bar{\theta}) \cdot f(x_2, \bar{\theta}) \cdot \dots \cdot f(x_n, \bar{\theta}).$$

для дискретного признака

$$L(x_1, x_2, \dots, x_n, \bar{\theta}) = p_{x_1}(\bar{\theta}) \cdot \dots \cdot p_{x_n}(\bar{\theta}).$$

В качестве оценки $\tilde{\theta}_1, \dots, \tilde{\theta}_n$ неизвестных параметров распределения $\theta_1, \dots, \theta_n$ берут те значения, при которых функция правдоподобия достигает максимума.

Технически задача поиска максимального значения функции правдоподобия облегчается, если рассмотреть не саму функцию, а натуральный логарифм от неё, т.е. функцию $\ln L(x_1, x_2, \dots, x_n, \bar{\theta})$. За оценки неизвестного вектора параметров $\bar{\theta}$ берётся решение $\tilde{\theta}_1, \dots, \tilde{\theta}_n$ уравнения правдоподобия

$$\frac{\partial \ln L(x_1, x_2, \dots, x_n, \bar{\theta})}{\partial \theta_k} = 0; \quad k = \overline{1, n}.$$

Для проверки того, что точка оптимума соответствует максимуму функции правдоподобия, необходимо найти вторую производную от этой функции. И если вторая производная в точке оптимума отрицательна, то найденные значения параметров максимизируют функцию.

Итак, нахождение оценок максимального правдоподобия включает следующие этапы: построение функции правдоподобия (ее натурального логарифма); дифференцирование функции по искомым параметрам и составление системы уравнений; решение системы уравнений для нахождения оценок; определение второй производной функции, проверку ее знака в точке оптимума первой производной и формирование выводов.

Пример1. Время до выхода из строя прибора определяется показательным распределением

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0; \\ 0, & x < 0. \end{cases}$$

По выборке 2, 3, 5, 6, 8 найти оценку параметра λ .

Решение: Функция правдоподобия имеет вид:

$$L(2,3,5,6,8,\lambda) = \lambda e^{-\lambda 2} \cdot \lambda e^{-\lambda 3} \cdot \lambda e^{-\lambda 5} \cdot \lambda e^{-\lambda 6} \cdot \lambda e^{-\lambda 8} = \lambda^5 e^{-\lambda 24}.$$

Тогда $\ln L(2,3,5,6,8,\lambda) = 5 \ln \lambda - 24\lambda$. Уравнение правдоподобия будет иметь вид:

$$\frac{\partial \ln L}{\partial \lambda} = \frac{5}{\lambda} - 24 = 0. \text{ Отсюда получаем оценку } \tilde{\lambda} = \frac{5}{24}.$$

Метод максимального правдоподобия позволяет получить состоятельные, эффективные (если таковые существуют, то полученное решение даст эффективные оценки), достаточные, асимптотически нормально распределенные оценки. Этот метод может давать как смещенные, так и несмещенные оценки. Смещение удаётся устранить введением поправок. Метод особенно полезен при малых выборках. Если функция максимального правдоподобия имеет несколько максимумов, то из них выбирают глобальный.

Метод моментов

Метод предложен К. Пирсоном в 1894 г.

Будем считать, что вид функции распределения изучаемого количественного признака известен, но параметры этого распределения неизвестны. Нужно оценить эти параметры по выборке. Сущность метода моментов заключается в том, что по конкретной выборке x_1, x_2, \dots, x_n генеральной совокупности X , распределение которой известно с точностью до параметров $\theta_1, \dots, \theta_m$ **выбирается столько эмпирических моментов, сколько требуется оценить неизвестных параметров распределения.** Желательно применять моменты младших порядков, так как погрешности вычисления оценок резко возрастают с увеличением порядка момента.

Вычисленные по ЭД оценки моментов приравниваются к теоретическим моментам; параметры распределения определяются через моменты, и составляются уравнения, выражающие зависимость параметров от моментов, в результате получается система уравнений. Решение этой системы дает оценки параметров распределения генеральной совокупности.

Пример 2. Время до выхода из строя прибора определяется показательным распределением

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0; \\ 0, & x < 0. \end{cases}$$

По выборке x_1, x_2, \dots, x_n найдём точечную оценку неизвестного параметра распределения λ .

Так как для показательного распределения нужно оценить единственный параметр λ , то для его оценивания нужно одно уравнение.

Приравняем, например, первые начальные моменты – теоретический и выборочный.

Первый начальный теоретический момент равен

$$m_1 = \int_0^{\infty} x \cdot \lambda e^{-\lambda x} dx = \frac{1}{\lambda}.$$

Первый начальный выборочный момент вычисляется по формуле

$$m_1^B = \sum_{i=1}^n \frac{x_i}{n}.$$

Приравняв их получим $\frac{1}{\lambda} = \sum_{i=1}^n \frac{x_i}{n}$. Отсюда $\tilde{\lambda} = \frac{1}{\sum_{i=1}^n \frac{x_i}{n}}$.

2. Интервальная оценка параметров распределения

Сущность задачи интервального оценивания параметров

Интервальный метод оценивания параметров распределения случайных величин заключается в определении интервала (а не единичного значения), в котором с заданной степенью достоверности будет заключено значение оцениваемого параметра.

Интервальная оценка характеризуется двумя числами – концами интервала, внутри которого предположительно находится истинное значение параметра. Иначе говоря, вместо отдельной точки для оцениваемого параметра можно установить интервал значений, одна из точек которого является своего рода "лучшей" оценкой. *Интервальные оценки являются более полными и надежными по сравнению с точечными*, они применяются как для больших, так и для малых выборок. Совокупность методов определения промежутка, в котором лежит значение параметра T , получила название **методов интервального оценивания**. К их числу принадлежит метод Неймана.

Постановка задачи интервальной оценки параметров заключается в следующем: Имеется выборка наблюдений x_1, x_2, \dots, x_n за случайной величиной X . Объем выборки n фиксирован. Предположим, что статистическая характеристика $\tilde{\theta}$, рассчитанная по данным выборки, является выборочной оценкой неизвестного параметра θ гене-

ральной совокупности, причём θ - это постоянное число. Оценка $\tilde{\theta}$ характеризует параметр θ тем точнее, чем меньше абсолютная величина разности $|\theta - \tilde{\theta}|$.

Если $|\theta - \tilde{\theta}| < \delta$, где $\delta > 0$, то число δ называется точностью оценки $\tilde{\theta}$.

Нельзя утверждать, что оценка $\tilde{\theta}$ абсолютно удовлетворяет неравенству $|\theta - \tilde{\theta}| < \delta$. Однако можно задать вероятность γ , с которой это неравенство осуществляется.

Надёжность или доверительная вероятность оценки $\tilde{\theta}$ - это вероятность γ , с которой осуществляется неравенство $|\theta - \tilde{\theta}| < \delta$.

Принято надёжность оценки задавать перед процессом оценивания параметра генеральной совокупности. В качестве вероятности γ берут число, близкое к 1 (от 0,95 до 0,999).

Пусть $P(|\theta - \tilde{\theta}| < \delta) = \gamma$. Если неравенство $|\theta - \tilde{\theta}| < \delta$ заменить равносильным ему двойным неравенством:

$$-\delta < \theta - \tilde{\theta} < \delta, \text{ или } \tilde{\theta} - \delta < \theta < \tilde{\theta} + \delta,$$

то получим: $P(\tilde{\theta} - \delta < \theta < \tilde{\theta} + \delta) = \gamma$.

Вероятностный смысл данного отношения таков: вероятность того, что интервал $(\tilde{\theta} - \delta; \tilde{\theta} + \delta)$ покрывает неизвестный параметр θ , равна γ .

Интервал $(\tilde{\theta} - \delta; \tilde{\theta} + \delta)$, который покрывает оцениваемый неизвестный параметр θ , с заданной вероятностью γ называется доверительным интервалом.

Границы $\tilde{\theta} - \delta$ и $\tilde{\theta} + \delta$ называются **доверительными границами интервала**. Они определяются на основе выборочных данных и являются функциями от случайных величин X_1, X_2, \dots, X_n , и, следовательно, сами являются случайными величинами.

Общий метод построения доверительных интервалов

Доверительный интервал для оценки математического ожидания нормального распределения при известном среднем квадратичном отклонении

Будем считать, что количественный признак X генеральной совокупности подчиняется нормальному закону распределения с параметрами α и σ . Среднее квадратичное данного распределения σ считается известным. Нужно оценить неизвестное математическое ожидание α генеральной совокупности по выборочной средней \bar{x} , т.е. необходимо определить доверительный интервал, который покрывает параметр α с заданной надёжностью γ .

Имеет место следующее утверждение: если случайная величина X подчиняется нормальному закону распределения, то выборочная средняя \bar{X} , найденная по независимым наблюдениям, также подчиняется нормальному закону распределения с параметрами:

$$M(\bar{X}) = \alpha; \quad \sigma(\bar{X}) = \frac{\sigma}{\sqrt{n}}.$$

Для того, чтобы определить доверительный интервал, необходимо, чтобы выполнялось равенство $P(|\bar{X} - \alpha| < \delta) = \gamma$.

Для нормально распределённой случайной величины X имеет место равенство:

$$P(|\bar{X} - \alpha| < \delta) = 2\Phi\left(\frac{\delta}{\sigma(\bar{X})}\right).$$

\bar{X} имеет нормальное распределение и $\sigma(\bar{X}) = \frac{\sigma}{\sqrt{n}}$. Поэтому

$$P(|\bar{X} - \alpha| < \delta) = 2\Phi\left(\frac{\delta\sqrt{n}}{\sigma}\right) = 2\Phi(t); \quad t = \frac{\delta\sqrt{n}}{\sigma}.$$

Из равенства $t = \frac{\delta\sqrt{n}}{\sigma}$ выразим $\delta = t \frac{\sigma}{\sqrt{n}}$ и получим:

$$P\left(|\bar{X} - \alpha| < t \frac{\sigma}{\sqrt{n}}\right) = 2\Phi(t).$$

Так как вероятность P задана и равна γ , окончательно получаем:

$$P\left(\bar{x} - t \frac{\sigma}{\sqrt{n}} < \alpha < \bar{x} + t \frac{\sigma}{\sqrt{n}}\right) = 2\Phi(t) = \gamma.$$

Таким образом, с доверительной вероятностью γ можно утверждать, что доверительный интервал $\left(\bar{x} - t \frac{\sigma}{\sqrt{n}} < \alpha < \bar{x} + t \frac{\sigma}{\sqrt{n}} \right)$ покрывает неизвестный параметр α с точностью $\delta = t \frac{\sigma}{\sqrt{n}}$.

\bar{x} - находим по данным выборки, n - это объём выборки, σ - величина, известная заранее.

Число t находится из равенства $2\Phi(t) = \gamma$, или $\Phi(t) = \frac{\gamma}{2}$. По таблице функции Лапласа определяется для заданной доверительной вероятности γ величина t и затем находится величина δ .

Методика Вальда проверки гипотезы о свойствах случайной величины.

Рассматривается следующая задача: имеются экспериментальные данные о некоторой случайной величине ξ ; известно аналитическое выражение для плотности распределения этой случайной величины - некоторое выражение $f(x, \theta)$, где x - аргумент, а θ - некоторый параметр, принимающий одно из двух возможных значений - $\theta = \theta_1$ или $\theta = \theta_2$, причем $\theta_1 < \theta_2$. Какое именно из этих двух значений принимает параметр θ , - неизвестно. Именно это и надо определить по значениям случайной величины ξ , которые можно разыгрывать неограниченно много.

Примером такой плотности может служить плотность нормального распределения:

$$f(x, \theta) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\theta)^2}{2\sigma^2}},$$

где, кроме параметра θ , участвует так же еще один параметр - σ .

Опишем *методику Вальда* решения поставленной задачи.

Фиксируем достаточно малые числа α и β , например $\alpha, \beta \in (0,1)$; фиксируем, далее, числа

$$A = \frac{1-\beta}{\alpha}, \quad B = \frac{\beta}{1-\alpha};$$

при достаточно малых α, β между числами A, B имеет место неравенство: $B < A$ - именно для этого неравенства должны быть достаточно малы числа α, β .

Пусть, далее, $\xi_1, \xi_2, \dots, \xi_k, \dots$ - значения случайной величины ξ , получаемые экспериментально. Построим суммы:

$$\Lambda_t = \sum_{i=1}^t \ln \frac{f(\xi_i, \theta_2)}{f(\xi_i, \theta_1)}.$$

Основная теорема Вальда: если $\Lambda_t > \ln A$, то $\theta = \theta_2$ с вероятностью ошибки β ; если же $\Lambda_t < \ln B$, то $\theta = \theta_1$ с вероятностью ошибки α .

Принято рассматривать в описанной ситуации две величины - n и Δg_t , определяемые так:

то минимальное значение t , при котором выполняется одно из неравенств - $\Lambda_t > \ln A$ или $\Lambda_t < \ln B$ - является случайной величиной, зависящей от той или иной реализации случайной величины $\xi_1, \xi_2, \dots, \xi_k, \dots$; именно это минимальное значение t и является случайной величиной n ;

учитывая смысл только что введенного обозначения n , введем величину

$$\Delta g_t = \frac{1}{n} \sum_{i=1}^n \ln \frac{f(\xi_i, \theta_2)}{f(\xi_i, \theta_1)};$$

очевидно, и эта величина - также случайная. На практике часто имеют смысл математические ожидания последних двух случайных величин; их находят как обычные средние арифметические значения реализаций этих величин, конструируемых по различным реализациям $\xi_1, \xi_2, \dots, \xi_k, \dots$.

Лекция 10

Интервальные оценки при экспоненциальном законе распределения

В качестве примера рассмотрим определение интервальных оценок для показателей надёжности. Вероятность безотказной работы определяется формулой

$$P(t) = e^{-\lambda t}, \quad \lambda = \frac{1}{T}, \quad T = m_t.$$

Доверительный интервал для средней наработки до отказа T при равных вероятностях $\frac{\alpha}{2}$ выхода за правую и левую границы

$$\frac{2T_{\Sigma r}}{\chi_{\alpha/2}^2} < T < \frac{2T_{\Sigma r}}{\chi_{1-\alpha/2}^2},$$

где $\chi_{\alpha/2}^2, \chi_{1-\alpha/2}^2$ находятся по таблице при числе степеней свободы $k = 2r$, где r - число зарегистрированных отказов, для соответствующих вероятностей.

Когда вычисляется только нижняя граница,

$$T > \frac{2T_{\Sigma r}}{\chi_{\alpha}^2}.$$

Если $k > 100$, то доверительный интервал вычисляется по формулам для нормального усечённого распределения.

Для вероятности безотказной работы двухсторонний доверительный интервал вычисляется по формуле

$$\exp\left[-\frac{\chi_{\alpha/2}^2 t_i}{2T_{\Sigma r}}\right] < P(t_i) < \exp\left[-\frac{\chi_{1-\alpha/2}^2 t_i}{2T_{\Sigma r}}\right].$$

В случае односторонней границы

$$P(t_i) > \exp\left[-\frac{\chi_{\alpha}^2 t_i}{2T_{\Sigma r}}\right].$$

Лекция 10

10.1 Планирование испытаний и обработка экспериментальных данных

В соответствии с требованиями ГОСТ 27.002-83 планирование испытаний предусматривает ряд предварительных условий, обеспечивающих эффективность испытаний. Вводятся условные обозначения различных планов в виде совокупности трех символов, первый из которых указывает число испытываемых объектов (устройств) N , второй - наличие (R) или отсутствие (U) замены (восстановления) объектов, отказавших во время испытаний, третий - длительность испытаний (r или T). Таким образом, для испытаний N объектов без замены отказавших элементов, имеем следующие три плана:

(N, U, r) - испытания до r -го отказа, $r < N$;

(N, U, T) - испытания длительностью T ;

$[N, U, (r, T)]$ - испытание длительностью, равной $\min(t_r \text{ или } T)$, где t_r - момент r -го отказа, а T - заведомо заданный ресурс (время, или величина пробега, или число циклов и т.д.)

Аналогично вводятся обозначения для планов с заменой (восстановлением) отказавших устройств: (N, R, r) ; (N, R, T) ; $[N, R, (r, T)]$. В плане (N, R, r) в отличие от (N, U, r) число r может быть больше, чем N (где, в частности, допустимо $N = 1$). Здесь, приведено 6 наиболее распространенных типов испытаний. ГОСТ 27.001-83 предусматривает 16 планов испытаний, где учтены кроме названных условий и такие как M - восстановление объектов при испытаниях в случае их отказов; S - решение об окончании испытаний (о приемке или браковке) восстанавливаемых объектов (основывается на суммарном времени испытаний).

Таблица 1.

План испытаний	Суммарная наработка $T_{\Sigma r}$, ч
(NUr)	$T_{\Sigma r} = \sum_{j=1}^r t_j + (N - r) \cdot t_r$
(NUT)	$T_{\Sigma r} = \sum_{j=1}^r t_j + (N - r) \cdot T$
$NU(rT)$	при $t_r < T$, $T_{\Sigma r} = \sum_{j=1}^r t_j + (N - r) \cdot t_r$;

	при $t_r \geq T$, $T_{\Sigma r} = \sum_{j=1}^r t_j + (N - r) \cdot T$
(NRr)	$T_{\Sigma r} = Nt_r$
(NRT)	$T_{\Sigma r} = NT$
$NR(rT)$	при $t_r < T$, $T_{\Sigma r} = Nt_r$; при $t_r \geq T$, $T_{\Sigma r} = NT$

Результаты статистической обработки испытаний существенно зависят от вероятностных моделей, то есть от априорных (теоретических) распределений интервалов безотказной работы и восстановлений. Эти результаты могут приводить к заведомо ошибочным выводам, если модель не отражает реальные процессы возникновения отказов и механизмы восстановления. Поэтому до решения основных задач апостериорного (на основе опыта) анализа надежности целесообразно сначала проверить, с помощью статистического критерия согласия, на соответствие выбранного априорного распределения эмпирическому распределению, построенному на основании данных проведенных испытаний.

Исходными данными (случайными величинами), которые подвергаются обработке, являются время наработки на отказ, время наработки на восстановление и число отказов однотипных элементов. После того, как такой материал собран, его обработка позволяет установить законы распределения показателей надежности: вероятность безотказной работы, интенсивность отказов, среднее время наработки на отказ и др.

Знание законов распределения дает возможность определить все остальные количественные показатели надежности. Таким образом, основная задача статистической обработки состоит в определении одного из законов распределения исходных экспериментальных величин. В ряде случаев вид закона распределения известен заранее, до опыта. Например, как уже отмечалось выше, для электронной аппаратуры средств автоматики и релейной защиты справедлив экспоненциальный закон распределения показателей надежности. Это подтверждается многочисленными опытными данными, полученными в условиях эксплуатации.

При определении или подтверждении закона распределения целесообразен следующий порядок: подготовка опытных данных; построение гистограмм оцениваемого количественного показателя надежности; аппроксимация гистограмм теоретическим законом распределения и определение его параметров; проверка допустимости предполагаемого закона распределения на основе использования критериев согласия. Наиболее часто используется критерий χ^2 или критерий Колмогорова. Для получения достаточно точных результатов число наблюдений случайной величины (отказов) должно быть не менее 40-50.

10.2. Интервальная оценка показателей надежности

Количество статистических данных для оценки надежности, полученных в процессе эксплуатации, принципиально ограничено. Полученные по ограниченному объему информации точечные оценки могут оказаться весьма приблизительными. Причем отклонения этих оценок от истинного значе-

ния оцениваемого параметра являются величинами случайными. Очевидно, что с увеличением числа наблюдений (отказов) случайная ошибка оценки показателей уменьшается. На основе опытных данных используется специальная методика оценки показателей надежности в определенном интервале возможных их значений. Предположим, что истинное значение средней наработки до отказа составляет T_0 , а средняя наработка до отказа определена по полученным отказам:

$$\hat{T} = \frac{1}{n} \sum_{i=1}^n t_i,$$

где n - количество отказов за время испытаний, t_i - наработка до i -го отказа. Чем меньше n тем больше расхождение между T_0 и \hat{T} , то есть существует интервал расхождения. Найти точные границы, в пределах которых находится истинное значение искомой величины, не представляется возможным. Однако можно определить интервал ее возможных значений с некоторой доверительной вероятностью β . При этом, чем больше доверительная вероятность β , тем шире границы интервала и наоборот. В общем виде, эта зависимость имеет запись

$$\beta = P(T_H \leq T_0 \leq T_B) \quad (1)$$

где T_H и T_B - соответственно нижняя и верхняя границы средней наработки до отказа, где лежат \hat{T} и T_0 .

Вероятность того, что значение T_0 выйдет за заданный интервал называется **уровнем значимости**:

$$\alpha = P(T_H > T_0 > T_B) = 1 - \beta. \quad (2)$$

Значения доверительных вероятностей β обычно принимают равными 0,9; 0,95; 0,99. Соответствующие им уровни значимости составят 0,1; 0,05; 0,01. Доверительная вероятность β , определяемая выражением (1), характеризует степень достоверности результатов двусторонней (то есть с определением верхней и нижней границ) оценки.

Доверительный интервал для средней наработки до отказа при равных вероятностях $\frac{\alpha}{2}$ выхода за правую (верхнюю) и левую (нижнюю) границы для экспоненциального распределения определяется по выражению

$$T_H = \frac{2T_{\Sigma r}}{\chi_{\frac{\alpha}{2}, 2r}^2} < \hat{T} < \frac{2T_{\Sigma r}}{\chi_{1-\frac{\alpha}{2}, 2r}^2} = T_B, \quad (3)$$

где $\chi_{\frac{\alpha}{2}, 2r}^2$ и $\chi_{1-\frac{\alpha}{2}, 2r}^2$ - значения χ^2 (хи-квадрат) при параметрах $\frac{\alpha}{2}$ и $1 - \frac{\alpha}{2}$; $2r = k$ - число степеней свободы, для вероятностей $P = \frac{\alpha}{2}$; $P = 1 - \frac{\alpha}{2}$ соответственно.

Когда вычисляется только нижняя граница, то

$$\hat{T} > \frac{2T_{\Sigma r}}{\chi_{\alpha, 2r}^2} \quad (4)$$

В выражениях (3) и (4) $T_{\Sigma r}$ - суммарная наработка до отказа по отказам, зафиксированным во время эксперимента. Значения $\chi_{\frac{\alpha}{2}, 2r}^2$ и $\chi_{\frac{1-\alpha}{2}, 2r}^2$ определяются по таблице П-1 квантилей распределения χ^2 (хи-квадрат).

Таким образом, для заданных уровней значимости α и числа степеней свободы k по таблице находят соответствующие значения χ^2 , подставляют в выражение (3) и находят T_n и T_0 . Величина α задается в зависимости от требований, предъявляемых к анализируемой системе.

10.3. Статистические оценки генеральной совокупности. Задача об оценке качества по выборке

Пусть ξ - случайная величина и ξ_1, \dots, ξ_n - ее значения, совокупность которых принято называть *генеральной совокупностью*. Напомним две основные статистические оценки генеральной совокупности: *математическое ожидание* - число $M_\xi = \frac{1}{n} \sum_{i=1}^n \xi_i$, и *дисперсия* - число

$$D_\xi = \frac{1}{n} \sum_{i=1}^n (\xi_i - M_\xi)^2.$$

Часто вместо символов M_ξ, D_ξ пишут $M(\xi), D(\xi)$; кроме того, во многих случаях применяется обозначение $\sigma(\xi)$ для числа $\sqrt{D(\xi)}$.

Если наряду со случайной величиной ξ имеется вторая случайная величина η и набор ее значений того же объема - η_1, \dots, η_n , то рассматривают *корреляционный момент* пары случайных величин ξ, η - число

$$k_{\xi\eta} = \frac{1}{n} \sum_{i=1}^n (\xi_i - M(\xi))(\eta_i - M(\eta)),$$

а также *коэффициент корреляции*

$$r_{\xi\eta} = \frac{k_{\xi\eta}}{\sigma(\xi)\sigma(\eta)}.$$

Если о паре случайных величин ξ, η известно, что между ними существует некоторая зависимость $\eta = \eta(\xi)$, то имеется методика приближенного отыскания этой зависимости, основанная на методе наименьших квадратов. Если требуется среди многочленов степени, не превосходящей некоторого m , найти тот, который имитирует эту зависимость лучше всего, то достаточно решить несложную систему линейных алгебраических уравнений и при $m < n$ получить определенный многочлен $Q(x) = q_0 + q_1x + \dots + q_mx^m$; доказывается, что при наличии зависимости $\eta = \eta(\xi)$ дисперсия $D(\eta)$ не превосходит числа

$$\frac{1}{n-m} \sum_{i=1}^n (\eta_i - Q(\xi_i))^2,$$

которое принято называть *оценкой дисперсии по методу наименьших квадратов*.

Рассмотрим классическую задачу о *статистической оценке качества*.

Имеется партия некоторого товара (например, партия автомобилей). Требуется оценить качество этой партии, сделав некоторую выборку из нее (например, отобрав несколько автомобилей, оценить их и по этим оценкам сделать выводы обо всей партии). В такой постановке речь идет об *оценке качества генеральной совокупности по выборке*.

Излагаемая ниже методика решения этой задачи предложена А. Вальдом.

Введем какой-либо числовой параметр, характеризующий качество партии. Например, если речь идет о партии автомобилей, то таким параметром может быть доля (процент) автомобилей бракованных. Если этот параметр меньше некоторого числа, то партия будет считаться хорошей, а если этот параметр больше некоторого другого числа, то партия будет считаться плохой.

Пусть α - вероятность забраковать *хорошую* партию (т.е. вероятность того, что анализ выборки покажет, что партия *плохая* в то время, как на самом деле, партия *хорошая*). Пусть β - вероятность принять *плохую* партию (т.е. вероятность того, что анализ выборки покажет, что партия *хорошая* в то время, как в действительности - партия *плохая*).

Пусть N - объем партии и n_0 - объем выборки. Пусть, далее, ν - *приемочное число*, а m - количество дефектных изделий в выборке, т.е. если в выборке $m \leq \nu$, то партия принимается (или партия - хорошая), а если в выборке $m > \nu$, то вся партия не принимается (или партия - плохая). Наконец, обозначим через L количество дефектных изделий во всей партии и пусть $l_0, l_1, l_0 < l_1$ - границы контролируемого параметра L , т.е. если $L \leq l_0$, то партия - хорошая, а если $L \geq l_1$, то партия - плохая.

В этих условиях вероятности α, β можно вычислить явно с помощью комбинаторных формул. Напишем лишь окончательный результат, используя стандартное обозначение $P(a)$ для вероятности наступления события a :

$$\alpha = P(M > \nu | L = l_0) = 1 - \sum_{m=0}^{\nu} \frac{C_{l_0}^m C_{N-l_0}^{n_0-m}}{C_N^{n_0}},$$

$$\beta = P(M \leq \nu | L = l_1) = \sum_{m=0}^{\nu} \frac{C_{l_1}^m C_{N-l_1}^{n_0-m}}{C_N^{n_0}}.$$

Если $n_0 \leq \frac{N}{10}$, то α, β приближенно равны:

$$\alpha \cong 1 - \sum_{m=0}^{\nu} C_{n_0}^m p_0^m (1-p_0)^{n_0-m}, \beta \cong \sum_{m=0}^{\nu} C_{n_0}^m p_1^m (1-p_1)^{n_0-m},$$

где $p_0 = \frac{l_0}{N}$, $p_1 = \frac{l_1}{N}$. Если $p_0, p_1 < 0.1$, то возможны дальнейшие упрощения формул.

Числа α, β называют, соответственно, *риском поставщика* и *риском потребителя*. При составлении плана контроля партии по выборке стремятся так выбрать границы $l_0 < l_1$, чтобы эти риски были, по возможности, одинаковыми.

Примечание 1. Практическое вычисление выражений типа

$$\alpha = 1 - \sum_{m=0}^{\nu} \frac{C_{l_0}^m C_{N-l_0}^{n_0-m}}{C_N^{n_0}}, \beta = \sum_{m=0}^{\nu} \frac{C_{l_1}^m C_{N-l_1}^{n_0-m}}{C_N^{n_0}}$$

требует предварительного преобразования их к виду, удобному для реализации вычислений на компьютере; без такой подготовки непосредственное вычисление факториалов создает ситуацию «переполнения» памяти. Вот основное в этой связи наблюдение.

Выражение C_n^k представляет собой следующее произведение

$$\frac{n}{k} \cdot \frac{n-1}{k-1} \cdot \frac{n-2}{k-2} \cdot \dots \cdot \frac{n-k+1}{1}.$$

Здесь ровно k сомножителей. Теперь рассмотрим выражение

$$\frac{C_{l_0}^m C_{N-l_0}^{n_0-m}}{C_N^{n_0}};$$

распишем его в соответствии с указанным только что правилом:

$$\frac{\frac{l_0}{m} \cdot \frac{l_0-1}{m-1} \cdot \dots \cdot \frac{l_0-m+1}{1} \cdot \frac{N-l_0}{n_0-m} \cdot \frac{N-l_0-1}{n_0-m-1} \cdot \dots \cdot \frac{N-l_0-n_0+m+1}{1}}{\frac{N}{n_0} \cdot \frac{N-1}{n_0-1} \cdot \dots \cdot \frac{N-n_0+1}{1}}.$$

В знаменателе этого выражения указано n_0 дробей; в числителе количество дробей есть $m + (n_0 - m)$, то есть – их столько же, сколько в знаменателе.

Примечание 2. В практических задачах, связанных с статистической оценкой генеральной совокупности, потребуется следующая справочная информация (приводится здесь исключительно для удобства читателя).

1) Случайная величина с равномерным распределением в интервале (a, b) имеет плотность

$$f(x) = \begin{cases} 0, & x < a, \\ \frac{1}{b-a}, & a \leq x \leq b, \\ 0, & x > b. \end{cases}$$

Ее математическое ожидание $M = \frac{a+b}{2}$ и ее дисперсия $D = \frac{(b-a)^2}{12}$.

Для имитации на компьютере случайной величины, равномерно распределенной в интервале $(0, 1)$, в каждом языке программирования предусмотрена специальная команда; будем условно обозначать ее RND (то есть символ RND обозначает число из совокупности равномерно распределенных чисел в интервале $(0, 1)$). Таким образом, для имитации чисел, равномерно распределенных в интервале (a, b) , нужно применять команду $(b-a)RND$.

2) Случайная величина с экспоненциальным распределением имеет плотность

$$f(x) = \begin{cases} 0, & x < 0, \\ \lambda e^{-\lambda x}, & x \geq 0. \end{cases}$$

Ее математическое ожидание $M = \frac{1}{\lambda}$, дисперсия $D = \frac{1}{\lambda^2}$. Можно доказать, что имитация экспоненциально распределенной случайной величины осуществляется по схеме $x = -\frac{1}{\lambda} \ln(1 - RND)$.

3) Случайная величина с нормальным распределением имеет плотность

$$f(x) = \frac{1}{\sqrt{2\pi b}} \exp\left(-\frac{(x-a)^2}{2b}\right).$$

Ее математическое ожидание $M = a$, дисперсия $D = b$. Можно доказать, что имитация нормально распределенной случайной величины осуществляется

по схеме $x = \sqrt{b} \left(\sum_{i=1}^{12} RND - 6 \right) + a$.

4) Случайная величина с логнормальным распределением имеет плот-

ность $f(x) = \frac{1}{x\sqrt{2\pi b}} \exp\left(-\frac{(\ln(x) - a)^2}{2b}\right)$. Ее математическое ожидание $M = \exp\left(\frac{1}{2}b + a\right)$, дисперсия $D = (\exp(b + 2a))(\exp(b) - 1)$. Для имитации логнормально-распределенной случайной величины с параметрами a, b , указанными в вышеприведенной формуле для плотности, следует имитировать нормально распределенную случайную величину с математическим ожиданием a и дисперсией b , а затем взять экспоненциальную функцию от каждого из этих чисел.

10.4. Случайные процессы как модели порождения данных. Линейные модели случайных процессов.

Существует много задач, данные для которых возникают случайно. Например, задача о прогнозировании землетрясений опирается на данные, собираемые специальными датчиками и рассматриваемые современной наукой как случайные числа. При исследовании случайных данных всегда возникает необходимость в их имитации на уровне уже распознанных закономерностей - тогда исследования можно проводить без дальнейшего сбора информации на уровне датчиков. В этой связи к настоящему времени сложилось несколько *моделей порождения данных*. Одной из таких моделей является *случайный процесс*.

Случайным процессом называется функция вещественной переменной, значениями которой являются случайные величины. Если $z(t)$ - случайный процесс, то аргумент t удобно рассматривать как время. Мы будем предполагать, что для всех рассматриваемых ниже случайных процессов $z(t)$ при $t < 0$ обязательно $z(t) = 0$.

Поскольку для каждого t величина $z(t)$ - случайная, постольку для каждого t имеется функция (или плотность) распределения этой случайной величины. В теории случайных процессов принято особым образом выделять такой случайный процесс $z(t)$, для которого каждая случайная величина $z(t)$ является нормально распределенной с математическим ожиданием 0 и дисперсией 1 (для всех $t \geq 0$); этот случайный процесс называется *белым шумом* и всюду ниже будет обозначаться символом $a(t)$. По традиции, аргумент случайного процесса записывают в виде индекса, т.е. вместо $z(t)$ пишут z_t , вместо $a(t)$ пишут a_t и т.д.

Условимся, что всюду в дальнейшем время будет меняться дискретно, только по целым числам, т.е. $t = 0, \pm 1, \pm 2, \dots$.

Среди случайных процессов особое место занимают процессы линейные, которые определяются так:

случайный процесс z_t называется *линейным*, если имеет место равенство: $z_t = a_t + \pi_1 a_{t-1} + \pi_2 a_{t-2} + \dots + \pi_s a_{t-s} + \dots$, где $\pi_1, \pi_2, \dots, \pi_s, \dots$ - числовая последовательность, а суммирование в правой части последнего равенства происходит фактически на конечном числе слагаемых, так как, напомним, для всех рассматриваемых случайных процессов $z(t)$ при $t < 0$ обязательно $z(t) = 0$.

При описании линейных случайных процессов удобно использовать технику формальных степенных рядов по следующей схеме.

Будем обозначать символом B так называемый *оператор сдвига*, т.е. символ, позволяющий формально сделать запись: $z_{t-1} = Bz_t$.

Такая запись означает, что $z_{t-2} = Bz_{t-1}$, а вместе эти две записи означают, что

$z_{t-2} = Bz_{t-1} = B^2 z_t$. Продолжая этот формализм, получаем:

$z_{t-k} = B^k z_t$.

Поэтому выражение $z_t = a_t + \pi_1 a_{t-1} + \pi_2 a_{t-2} + \dots + \pi_s a_{t-s} + \dots$ переписывается так:

$$z_t = a_t + \pi_1 B a_t + \pi_2 B^2 a_t + \dots + \pi_k B^k a_t + \dots;$$

последнее выражение можно переписать в виде

$$z_t = \pi(B) a_t,$$

где $\pi(B) = 1 + \pi_1 B + \pi_2 B^2 + \dots + \pi_k B^k + \dots$ - формальный степенной ряд. Учитывая, что у этого формального степенного ряда свободный член отличен от нуля, можно найти обратный к нему - такой ряд $\psi(B)$, что $\psi(B)\pi(B) = 1$; при этом $\psi(B)$ будет иметь следующий вид:

$$\psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \dots + \psi_k B^k + \dots,$$

где $\psi_1, \psi_2, \dots, \psi_k, \dots$ - числовая последовательность. С помощью последнего формального степенного ряда линейный случайный процесс переписывается в следующем виде:

$$z_t = a_t - \psi_1 z_{t-1} - \psi_2 z_{t-2} - \dots - \psi_k z_{t-k} - \dots$$

Таким образом, всякий линейный случайный процесс возникает по указанному выше правилу из белого шума и может задаваться через самого себя рекуррентно.

Остановимся на вопросе компьютерной имитации линейного случайного процесса.

Начнем с имитации белого шума. Из математики известно, что если взять двенадцать случайных чисел, распределенных равномерно в интервале $(0,1)$ - s_1, \dots, s_{12} - и с их помощью построить число $\sigma = \left(\sum_{i=1}^{12} s_i \right) - 6$, то числа σ , многократно повторенные с помощью такой процедуры, будут нормально распределены с математическим ожиданием 0 и дисперсией 1.

Отсюда ясно, как осуществить имитацию белого шума: если требуется воспроизвести значения a_0, a_1, \dots, a_n , то нужно построить с помощью датчика случайных чисел, равномерно распределенных в интервале $(0,1)$, наборы $s_{i,1}, s_{i,2}, \dots, s_{i,12}$, $i = 0, 1, \dots, n$, а затем положить

$$a_i = \left(\sum_{k=1}^{12} s_{i,k} \right) - 6, \quad i = 0, 1, \dots, n.$$

Заметим, что если бы требовалось имитировать нормально распределенные случайные числа, но не с нулевым математическим ожиданием и единичной дисперсией, а с произвольно заданными математическим ожиданием m и дисперсией D , то несложное обобщение вышеупомянутой математической теоремы привело бы нас к следующей формуле:

$$a_i = \sqrt{D} \left(\left(\sum_{k=1}^{12} s_{i,k} \right) - 6 \right) + m, i = 0, 1, \dots, n.$$

В случае же произвольного линейного случайного процесса

$z_t = a_t + \pi_1 a_{t-1} + \pi_2 a_{t-2} + \dots + \pi_s a_{t-s} + \dots$ последовательность действий для его имитации такова:

1) фиксируется число n , для которого будут имитироваться числа

$$z_0, z_1, \dots, z_n;$$

2) формируется имитация белого шума a_0, a_1, \dots, a_n в соответствии с

описанным выше правилом;

3) формируются числа

$$z_0 = a_0;$$

$$z_1 = a_1 + \pi_1 a_0;$$

$$z_2 = a_2 + \pi_1 a_1 + \pi_2 a_0;$$

.....

$$z_n = a_n + \pi_1 a_{n-1} + \pi_2 a_{n-2} + \dots + \pi_n a_0.$$

