

БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ХИМИЧЕСКИЙ ФАКУЛЬТЕТ
Кафедра физической химии

А. В. Блохин

ТЕОРИЯ ЭКСПЕРИМЕНТА

Курс лекций

В двух частях

Часть 1

МИНСК
2002

Автор-составитель Блохин А.В., кандидат химических наук.

Рецензенты:

кандидат химических наук ***Н.Н. Горошко;***
Л.М. Володкович.

Утверждено на заседании
Ученого совета химического факультета
29 марта 2002 г., протокол № 5.

ПРЕДИСЛОВИЕ

Учебное пособие представляет собой лекции по курсу «Теория эксперимента» для студентов IV курса химического факультета, специализирующихся на кафедре физической химии, и содержит основы современных методологических подходов к постановке и обработке результатов физико-химических исследований и математических методов, применяемых при планировании и оптимизации эксперимента.

В первой части пособия введено понятие о результатах эксперимента как случайных величинах, информация о которых содержится в законах распределения. Рассмотрен нормальный закон распределения вероятностей для непрерывных величин. Во многих прикладных задачах нет необходимости использовать законы распределения в полном виде, вместо них можно воспользоваться числовыми характеристиками случайной величины, в сжатой форме выражающими наиболее существенные особенности ее распределения. Введено понятие и рассмотрены свойства наиболее часто применяемых моментов распределения — математического ожидания и дисперсии. Рассмотрены основные понятия математической статистики: генеральная совокупность и случайная выборка, оценки генеральных параметров и их свойства, методы проверки статистических гипотез и построение доверительных интервалов для генерального среднего и дисперсии. Для получения оценок генеральных параметров используется метод максимального правдоподобия. Указаны способы оценки случайной и суммарной погрешности косвенных измерений. Представлены методы проверки однородности двух и более выборочных дисперсий, сравнения средних и расчета средневзвешенного значения величины.

Во второй части пособия рассмотрены основные методы корреляционного и регрессионного анализов, широко применяемых при обработке результатов физико-химических измерений. Введено понятие о стохастической связи между случайными величинами и коэффициенте корреляции, характеризующем тесноту линейной зависимости между ними. Коэффициенты полиномиальных зависимостей определяются методом наименьших квадратов, который обосновывается как частный случай метода максимального правдоподобия при нормальном распределении случайных величин. Использование полиномиальных моделей позволяет улучшать аппроксимацию экспериментальных данных, повышая порядок полиномов. Представлены основы дисперсионного анализа, использующего свойство аддитивности дисперсии изучаемой случайной величины, что дает возможность разло-

жить ее на отдельные составляющие, обусловленные влиянием независимых факторов или их взаимодействий. Рассмотрены основные положения и методы обработки результатов для однофакторного и двухфакторного дисперсионного анализов; метод планирования эксперимента по схеме латинского квадрата для трехфакторного анализа.

Изложены методы планирования эксперимента с использованием полиномиальных моделей, направленные на поиск оптимальных условий при неизвестном механизме протекания процессов. Показано, что выбор плана эксперимента определяется задачей исследования. Линейные модели используются в методе крутого восхождения по поверхности отклика. Для достижения экстремума может быть также использован метод симплекс-планирования. Для описания области, близкой к экстремуму, применяются композиционные планы второго порядка.

Лекции основаны на материале, представленном в следующих учебных пособиях:

1. *Ахназарова С.Л., Кафаров В.В.* Методы оптимизации эксперимента в химической технологии. М.: Высш. шк., 1985. 327 с.
2. *Спиридонов В.В., Лопаткин А.А.* Математическая обработка физико-химических данных. М.: МГУ, 1970. 221 с.
3. *Тейлор Дж.* Введение в теорию ошибок. М.: Мир, 1985. 272 с.

Изложенный в пособии материал условно систематизирован по разделам-лекциям и представляет собой *теоретическую основу* для рассмотрения практических вопросов и задач, возникающих при постановке, планировании и обработке физико-химических экспериментов. Многие положения и правила даны без математических доказательств, рассмотрение которых не является целью курса. Проведение с помощью этого пособия лекций-консультаций позволит, во-первых, высвободить дополнительное время для решения практических заданий в рамках отведенных на курс учебных часов (традиционно 24 лекционных часа и 10 часов семинарских занятий) и, во-вторых, активизировать самостоятельную работу студентов. На каждом занятии после обсуждения теоретических вопросов студентам будут предложены практические задачи, основанные на экспериментальных исследованиях, выполненных сотрудниками кафедры физической химии. Последние, после их апробации, составят в будущем третью часть данного пособия.

ВВЕДЕНИЕ

Задачей большинства физико-химических экспериментов является количественное изучение каких-либо свойств вещества. Для этого проводятся измерения одной или нескольких физических величин с последующей обработкой полученных данных. Экспериментальные результаты всегда содержат погрешности, связанные с тем, что любые измерения сопровождаются действием и взаимодействием большого числа разнообразных и трудноучитываемых факторов. Конечной целью любого исследования является не только представление наилучшей, по мнению экспериментатора, оценки измеряемой величины, но и максимально достоверной оценки погрешности измерений.

Любой прибор или устройство для измерения физических величин можно рассматривать в виде объекта (рис. 1), для которого x_1, \dots, x_k — входные измеряемые и регулируемые параметры; w_1, \dots, w_l — неконтролируемые, случайным образом изменяющиеся параметры («шум» объекта); y_1, \dots, y_m — выходные параметры. Комплекс параметров x_1, \dots, x_k называют *основным*, поскольку он определяет условия эксперимента. Результат опыта зависит не только от основных параметров, но и от «шума» объекта, влияние которого носит случайный характер. Поэтому естественно рассматривать и результат эксперимента, и ошибку измерения как случайные величины, управляемые вероятностными законами, и применять для учета действия случайных факторов теорию вероятностей. Тогда влияние случайных ошибок на результат измерения можно количественно оценить при помощи математической статистики — науки, занимающейся применением вероятностных методов к решению задач в различных областях наук, в частности в задаче обработки результатов наблюдений.

Современная химическая промышленность выпускает несколько десятков тысяч наименований продуктов, в лабораториях разрабатываются сотни новых технологических процессов. Экспериментальное изучение механизмов протекания всех этих процессов нереально, между тем задачи оптимизации и управления этими процессами необходимо решать. Для этих целей успешно применяются экспериментально-статистические методы, с помощью которых составляется *математическая модель* объекта и при неизвестном механизме протекающих в объекте процессов изучается зависимость *отклика* системы на изменения основных параметров.

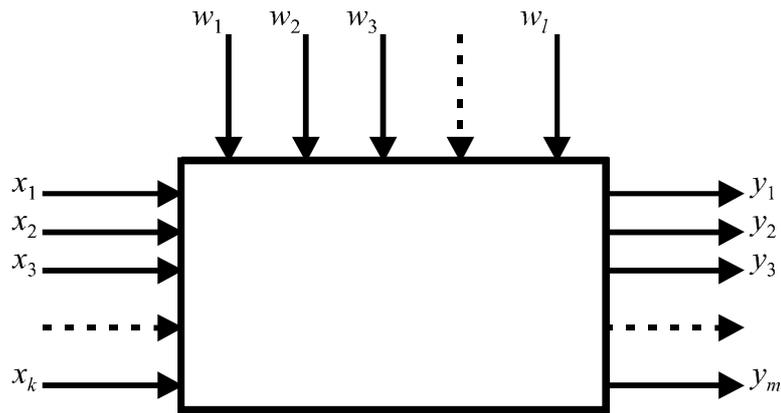


Рис. 1. Схема объекта.

Математической моделью объекта служит *функция отклика*, связывающая выходной параметр, характеризующий результаты эксперимента, с переменными, которые варьируют при проведении опытов:

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

Независимые переменные x_1, x_2, \dots, x_k называют *факторами*, пространство с координатами x_1, x_2, \dots, x_k — *факторным пространством*, а геометрическое изображение функции отклика в факторном пространстве — *поверхностью отклика*.

Эффективность экспериментов в большой степени зависит от методов их проведения. *Пассивный* эксперимент является традиционным методом, когда ставится большая серия опытов с поочередным варьированием каждой из переменных. Обработка опытных данных проводится статистическими методами, позволяющими оптимизировать процедуру обработки и анализа эксперимента. Используя *активный* (спланированный) эксперимент, можно достичь существенно большего — оптимизировать и стадию постановки эксперимента. Под *планированием эксперимента* понимают оптимальное управление экспериментом в условиях неполной информации о механизме процесса. Развитие этой концепции связано с работами Р. Фишера, главная идея которых состоит в раздельной оценке эффектов в многофакторной ситуации. Широко применяемое планирование эксперимента при поиске оптимальных условий процесса связано с работами Бокса и Уилсона. В настоящее время методы планирования и оптимизации эксперимента широко применяются при изучении процессов в лабораторных и полужаводских условиях и несколько реже в промышленности.

ЛЕКЦИЯ 1

Случайные величины. Классификация ошибок измерений. Абсолютная и относительная погрешность. Прямые и косвенные измерения. Оценка погрешностей функций приближенных аргументов. Распределение случайных величин. Функция распределения и плотность распределения.

1.1. Случайные величины.

Классификация ошибок измерений.

Абсолютная и относительная погрешность.

Под *случайной величиной* понимают величину, принимающую в результате испытания значение, которое принципиально нельзя предсказать, исходя из условий опыта. Случайная величина обладает целым набором допустимых значений, но в результате каждого отдельного опыта принимает лишь какое-то одно из них. В отличие от неслучайных величин, изменяющих свое значение только при изменении условий опыта, случайная величина может принимать различные значения даже при неизменном комплексе основных факторов.

Различают *дискретные* и *непрерывные* случайные величины. Возможные значения дискретных величин можно заранее перечислить. Значения непрерывной случайной величины не могут быть заранее перечислены, они заполняют собой некоторый интервал. Набор допустимых значений сам по себе слабо характеризует случайную величину. Чтобы ее полностью охарактеризовать, необходимо не только указать, какие значения она может принимать, но и как часто.

Каждый результат измерения — случайная величина. Отклонение результата реального измерения от истинного значения величины называется *ошибкой измерения*. («Ошибка» в научном смысле означает неизбежную погрешность, которая сопутствует всем измерениям). Ни одну физическую величину (длину, время, температуру и т.д.) невозможно измерить с полной определенностью. Лучшее, на что можно рассчитывать, — это свести ошибки к возможному минимуму и надежно рассчитать их величины.

Различают ошибки измерений трех видов:

1. *Грубые ошибки* возникают вследствие нарушения основных условий измерения. Результат, содержащий грубую ошибку, резко отличается по величине от остальных измерений, на чем основаны некоторые критерии исключения грубых ошибок.

2. *Систематические ошибки* постоянны во всей серии измерений или изменяются по определенному закону. Выявление их требует специальных исследований, их всегда стремятся свести к минимуму, а при необходимости они обычно учитываются введением соответствующих поправок в результаты измерения.
3. *Случайные ошибки* — ошибки измерения, остающиеся после устранения всех выявленных грубых и систематических ошибок. Они вызываются большим количеством таких факторов, эффекты действия которых столь незначительны, что их нельзя выделить в отдельности (при данном уровне техники измерения). При этом распределение случайных ошибок обычно симметрично относительно нуля: ошибки, противоположные по знаку, но равные по абсолютной величине, встречаются довольно часто.

Корректный способ представления результатов любого измерения состоит в том, что экспериментатор указывает свою наилучшую оценку измеряемой величины и интервал, в котором, как он уверен, она лежит. Чтобы охарактеризовать отклонение приближенного значения некоторой величины от ее истинного значения, вводят понятия абсолютной и относительной погрешностей, отвлекаясь от конкретного источника погрешностей.

Пусть A — точное значение исследуемой величины, a — ее наилучшая экспериментальная оценка (обычно среднее арифметическое серии измерений). Под *абсолютной ошибкой* (или *погрешностью*) величины a понимают абсолютное значение разности между этими значениями:

$$\varepsilon = |A - a| = |\Delta a|, \quad (1.1)$$

или

$$A = a \pm \varepsilon. \quad (1.2)$$

Предельная абсолютная погрешность определяется как

$$\varepsilon_{\text{пр.}} \geq |A - a|, \quad (1.3)$$

или

$$\varepsilon_{\text{пр.}} \geq \varepsilon, \quad (1.4)$$

при этом

$$(a + \varepsilon_{\text{пр.}}) \geq A \text{ и } A \geq (a - \varepsilon_{\text{пр.}}), \quad (1.5)$$

т. е. истинное значение искомой величины заведомо лежит в пределах

$$a - \varepsilon_{\text{пр.}} \leq A \leq a + \varepsilon_{\text{пр.}}. \quad (1.6)$$

Для характеристики относительной точности измерений, зависящей от значения измеряемой величины, вводится *относительная погрешность*:

$$\delta = \frac{\varepsilon}{|A|}, \quad (1.7)$$

$$\delta |A| = \varepsilon. \quad (1.8)$$

По аналогии с абсолютной погрешностью вводится также понятие *предельной относительной погрешности*:

$$\delta_{\text{пр.}} \geq \frac{\varepsilon}{|A|}, \quad (1.9)$$

или

$$\delta_{\text{пр.}} |A| \geq \varepsilon. \quad (1.10)$$

Тогда

$$\delta_{\text{пр.}} = \frac{\varepsilon_{\text{пр.}}}{|A|}, \quad (1.11)$$

или

$$\delta_{\text{пр.}} |A| = \varepsilon_{\text{пр.}}. \quad (1.12)$$

В вышеприведенные формулы входит неизвестная величина A , что делает невозможным численное определение погрешности. Практически поступают следующим образом: так как в большинстве случаев абсолютная погрешность много меньше самой измеряемой величины, т. е. $\varepsilon \ll |A|$, $\varepsilon \ll |a|$ или $A \approx a$, то для таких достаточно точных измерений можно записать:

$$\delta |a| \approx \varepsilon \text{ и } \delta_{\text{пр.}} |a| \approx \varepsilon_{\text{пр.}}$$

Тогда с учетом определений абсолютной и относительной погрешностей получаем

$$A = a \pm \varepsilon = a \left(1 \pm \frac{\varepsilon}{a} \right) \approx a(1 \pm \delta) \text{ или } A \approx a(1 \pm \delta_{\text{пр.}}). \quad (1.13)$$

Относительная погрешность в отличие от абсолютной является величиной безразмерной и для большинства измерений представляет собой малое число, поэтому ее часто умножают на 100 и приводят в процентах.

1.2. Оценка погрешностей функций приближенных аргументов.

Измерения делят на *прямые* и *косвенные*. В первом случае непосредственно измеряется определяемая величина, при косвенных измерениях она задается некоторой функцией от непосредственно измеряемых величин. Подавляющее большинство физико-химических свойств веществ и параметров процессов определяются в результате косвенных измерений, погрешность которых зависит от погрешностей непосредственно измеряемых величин, использованных в расчетах.

Предположим, что некоторые величины X_1, X_2, \dots, X_n измерены с абсолютными погрешностями $\Delta x_1, \Delta x_2, \dots, \Delta x_n$ и что измеренные значения используются для вычисления функции

$$Z = f(X_1, X_2, \dots, X_n). \quad (1.14)$$

Очевидно, что погрешности приближенных аргументов должны привести к погрешности в значении искомой функции, что можно записать в следующем виде:

$$Z + \Delta z = f(X_1 + \Delta x_1, X_2 + \Delta x_2, \dots, X_n + \Delta x_n), \quad (1.15)$$

где Δz — абсолютная погрешность функции Z .

Разложим правую часть равенства (1.15) в ряд Тейлора:

$$\begin{aligned} Z + \Delta z = & f(X_1, X_2, \dots, X_n) + \\ & + \sum_{i=1}^n \left(\frac{\partial f}{\partial X_i} \right) \Delta x_i + \sum_{i=1}^n \frac{1}{2} \left(\frac{\partial^2 f}{\partial X_i^2} \right) \Delta x_i^2 + \dots \end{aligned} \quad (1.16)$$

Если предположить, что измерения достаточно точны, так что величины Δx_i малы по сравнению со значениями аргументов X_i , то в выражении (1.16) можно отбросить все члены, содержащие абсолютные погрешности аргументов во второй и высшей степенях. Тогда

$$Z + \Delta z \approx f(X_1, X_2, \dots, X_n) + \sum_{i=1}^n \left(\frac{\partial f}{\partial X_i} \right) \Delta x_i, \quad (1.17)$$

откуда с учетом (1.14) получаем

$$\Delta z \approx \sum_{i=1}^n \left(\frac{\partial f}{\partial X_i} \right) \Delta x_i. \quad (1.18)$$

Выражение для предельной абсолютной погрешности функции n переменных запишется в следующем виде:

$$\varepsilon_{\text{пр}} \approx \sum_{i=1}^n \left| \frac{\partial f}{\partial X_i} \right| |\Delta x_i| = \sum_{i=1}^n \left| \frac{\partial f}{\partial X_i} \right| \varepsilon_i, \quad (1.19)$$

т.е. предельная абсолютная погрешность функции независимых переменных равна сумме частных производных этой функции, умноженных на соответствующие абсолютные погрешности аргументов. В практических расчетах значения частных производных берутся в точках, соответствующих измеренным значениям x_i или средним арифметическим \bar{x}_i , если проводились серии измерений.

В математической статистике также доказывается, что если абсолютные погрешности аргументов независимы и случайны, то наилучшей оценкой погрешности функции (1.14) будет *квадратичная сумма* ее частных производных, умноженных на соответствующие погрешности аргументов:

$$\Delta z \approx \sqrt{\sum_{i=1}^n \left(\frac{\partial f}{\partial X_i} \Delta x_i \right)^2}. \quad (1.20)$$

Формулы (1.19) и (1.20) являются основными при практических расчетах. Из них можно вывести формулы для расчетов погрешностей косвенных измерений для некоторых частных случаев, использование которых на практике бывает более удобным:

1. *Измеренная величина умножается на точное число.* Если величина X измерена с погрешностью Δx и используется для вычисления

$$Z = BX,$$

в котором B — точное число, то абсолютная погрешность в Z равна

$$|\Delta z| = |B| \cdot |\Delta x|. \quad (1.21)$$

2. *Погрешность в суммах и разностях.* Если величины X_1, X_2, \dots, X_n измерены с малыми погрешностями $\Delta x_1, \Delta x_2, \dots, \Delta x_n$ и измеренные значения используются для вычисления функции

$$Z = (X_1 + \dots + X_m) - (X_k + \dots + X_n),$$

а погрешности аргументов независимы и случайны, то погрешность в Z равна квадратичной сумме исходных погрешностей:

$$|\Delta z| = \sqrt{(\Delta x_1)^2 + \dots + (\Delta x_m)^2 + (\Delta x_k)^2 + \dots + (\Delta x_n)^2}; \quad (1.22)$$

в любом случае она никогда не больше, чем их обычная сумма

$$|\Delta z| \leq |\Delta x_1| + \dots + |\Delta x_m| + |\Delta x_k| + \dots + |\Delta x_n|. \quad (1.23)$$

3. *Погрешности в произведениях и частных.* Если величины X_1, X_2, \dots, X_n измерены с малыми погрешностями $\Delta x_1, \Delta x_2, \dots, \Delta x_n$ и измеренные значения используются для вычисления функции

$$Z = \frac{X_1 \times \dots \times X_m}{X_k \times \dots \times X_n},$$

а погрешности аргументов независимы и случайны, то относительная погрешность в Z равна квадратичной сумме исходных относительных погрешностей:

$$\frac{|\Delta z|}{|Z|} = \sqrt{\left(\frac{\Delta x_1}{|X_1|}\right)^2 + \dots + \left(\frac{\Delta x_m}{|X_m|}\right)^2 + \left(\frac{\Delta x_k}{|X_k|}\right)^2 + \dots + \left(\frac{\Delta x_n}{|X_n|}\right)^2}; \quad (1.24)$$

в любом случае она никогда не больше, чем их обычная сумма

$$\frac{|\Delta z|}{|Z|} \leq \frac{|\Delta x_1|}{|X_1|} + \dots + \frac{|\Delta x_m|}{|X_m|} + \frac{|\Delta x_k|}{|X_k|} + \dots + \frac{|\Delta x_n|}{|X_n|}. \quad (1.25)$$

4. *Погрешность в произвольной функции одной переменной.* Если величина X измерена с погрешностью Δx и используется для вычисления функции $Z = f(X)$, то абсолютная погрешность в Z равна

$$|\Delta z| = \left| \frac{\partial Z}{\partial X} \right| |\Delta x|. \quad (1.26)$$

5. *Погрешность в степенной функции.* Если величина X измерена с погрешностью Δx и используется для вычисления степенной функции $Z = X^m$ (где m — фиксированное известное число), относительная погрешность в Z в $|m|$ раз больше, чем в X :

$$\frac{|\Delta z|}{|Z|} = |m| \cdot \frac{\Delta x}{|X|}. \quad (1.27)$$

Пользуясь формулами (1.21) - (1.27), можно справиться практически с любой задачей вычисления ошибок в случае косвенных измерений. Любой расчет может быть представлен как последовательность определенных шагов, каждый из которых включает один из следующих видов операций: 1) нахождение сумм и разностей, 2) расчет произведений и частных, 3) вычисление функции одного переменного

(данный метод называют «шаг за шагом»). Однако в случае когда выражение для вычисления функции Z включает одну и ту же величину более чем один раз (например, дважды X_1), то некоторые из ошибок могут взаимно компенсироваться и в результате расчет ошибки методом «шаг за шагом» может привести к переоценке конечной погрешности. Поэтому в подобных случаях рекомендуется пользоваться общими формулами (1.19) и (1.20).

1.3. Распределение случайных величин. Функция распределения и плотность распределения случайной величины.

Пусть дискретная физическая величина X может принимать в результате опыта значения x_1, x_2, \dots, x_n . Отношение числа опытов m_i , в результате которых величина X принимает значение x_i , к общему числу проведенных опытов n называется *частотой появления события* $X = x_i$. Частота (m_i/n) является случайной величиной и меняется в зависимости от количества проведенных опытов. Однако при большом количестве опытов (в пределе $n \rightarrow \infty$) она стабилизируется около некоторого значения p_i , называемого *вероятностью события* $X = x_i$ (статистическое определение):

$$p_i = P(X = x_i) \approx (m_i/n). \quad (1.28)$$

Очевидно, что сумма вероятностей реализации всех возможных значений случайной величины равна единице:

$$\sum_{i=1}^n p_i = 1. \quad (1.29)$$

Дискретную случайную величину можно полностью задать *вероятностным рядом*, указав вероятность p_i для каждого значения x_i :

x_1	x_2	x_3	...	x_n
p_1	p_2	p_3	...	p_n

Законом распределения случайной величины называют любое соотношение, устанавливающее связь между возможными значениями случайной величины и соответствующими им вероятностями. Вероятностный ряд является одним из видов законов распределения случайной величины.

Распределение непрерывной случайной величины нельзя задать вероятностным рядом, поскольку число значений, которое она может принимать, так велико, что для большинства из них вероятность при-

нять эти значения равна нулю. Поэтому для непрерывных физических величин изучается вероятность того, что в результате опыта значение случайной величины попадет в некоторый интервал. Удобно пользоваться вероятностью события $X \leq x$, где x — произвольное действительное число. Эта вероятность

$$P(X \leq x) = F(x) \quad (1.30)$$

является функцией от x и называется *функцией распределения* (*предельной функцией распределения*, *функцией распределения генеральной совокупности*) случайной величины. В виде функции распределения можно задать распределение как непрерывной, так и дискретной случайной величины (рис. 2 и 3). $F(x)$ является неубывающей функцией, т.е. если $x_1 \leq x_2$, то $F(x_1) \leq F(x_2)$ (рис. 3).

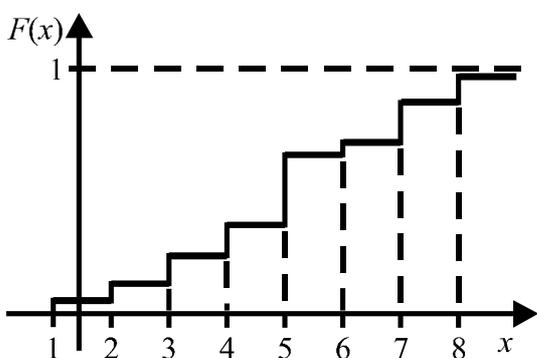


Рис. 2. Функция распределения дискретной случайной величины.

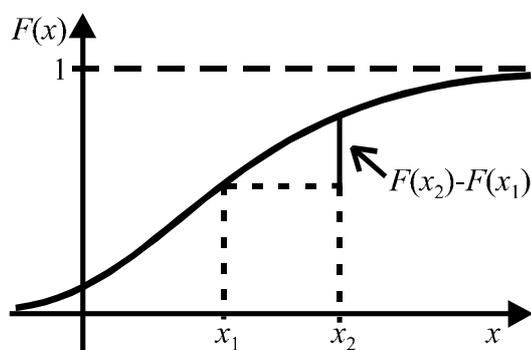


Рис. 3. Функция распределения непрерывной случайной величины.

Ордината кривой $F(x)$, соответствующая точке x_i , представляет собой вероятность того, что случайная величина X при испытании окажется $\leq x_i$. Тогда вероятность того, что значения случайной величины будут лежать в интервале от x_1 до x_2 , равна

$$P(x_1 \leq X \leq x_2) = F(x_2) - F(x_1). \quad (1.31)$$

Значения $F(x)$ при предельных значениях аргумента равны: $F(-\infty) = 0$, $F(+\infty) = 1$. Следует отметить, что функция распределения дискретной случайной величины всегда есть разрывная функция. Скачки происходят в точках, соответствующих возможным значениям этой величины, и равны вероятностям этих значений (рис. 2).

Для непрерывной случайной величины наиболее часто используется производная функции распределения — *плотность распределения* случайной величины X .

Если $F(x)$ непрерывна и дифференцируема, то

$$f(x) = \frac{dF(x)}{dx}. \quad (1.32)$$

Задание $f(x)$ также полностью определяет случайную величину. Плотность распределения является неотрицательной функцией (рис. 4).

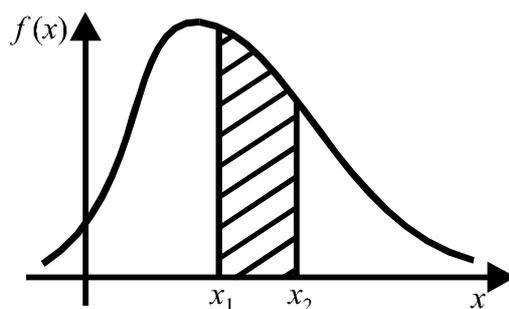


Рис. 4. Плотность распределения непрерывной случайной величины.

Площадь, ограниченная осью x , прямыми $x = x_1$ и $x = x_2$ и кривой плотности распределения, равна вероятности того, что случайная величина примет значения из интервала $x_1 \div x_2$:

$$P(x_1 \leq X \leq x_2) = \int_{x_1}^{x_2} f(x) dx = F(x_2) - F(x_1). \quad (1.33)$$

Тогда

$$F(x) = P(-\infty \leq X \leq x) = \int_{-\infty}^x f(x) dx. \quad (1.34)$$

Поскольку попадание случайной величины в интервал $-\infty < X < +\infty$ есть достоверное событие, то

$$\int_{-\infty}^{+\infty} f(x) dx = 1. \quad (1.35)$$

ЛЕКЦИЯ 2

Числовые характеристики случайной величины. Свойства математического ожидания и дисперсии. Нормированная случайная величина. Квантили. Нормальное и стандартное распределения случайной величины. Функция Лапласа. Задача об абсолютном отклонении.

2.1. Числовые характеристики случайной величины. Свойства математического ожидания и дисперсии. Нормированная случайная величина.

Вместо полного определения случайной величины в виде законов распределения вероятностей в прикладных задачах ее часто определяют при помощи *числовых характеристик* — чисел (вещественных), выражающих характерные особенности случайной величины, называемых *моментами случайной величины*.

Наиболее часто в приложениях математической статистики используют *математическое ожидание* (характеристику положения значений случайной величины на числовой оси) и *дисперсию* (или *среднее квадратичное отклонение*), определяющую характер разброса значений случайной величины.

Математическое ожидание (генеральное среднее) случайной величины (*начальный момент первого порядка*) принято обозначать $M[X]$, m_x или m . Оно определяется для дискретной и непрерывной случайной величины соответственно как

$$m = M[X] = \sum_{i=1}^n x_i p_i, \quad (2.1)$$

$$m_x = M[X] = \int_{-\infty}^{+\infty} x f(x) dx. \quad (2.2)$$

Для случайных величин математическое ожидание является теоретической величиной, к которой приближается среднее значение \bar{x} случайной величины X при большом количестве испытаний.

Свойства математического ожидания:

1. Если c — постоянное число (неслучайная величина), то

$$M[c] = c, \quad (2.3)$$

$$M[cX] = c M[X]. \quad (2.4)$$

2. Математическое ожидание суммы случайных величин равно сумме математических ожиданий этих случайных величин:

$$M[X_1 + X_2 + \dots + X_n] = M[X_1] + M[X_2] + \dots + M[X_n]. \quad (2.5)$$

3. Математическое ожидание произведения независимых случайных величин равно произведению математических ожиданий сомножителей:

$$M[X_1 \cdot X_2 \cdot X_3 \cdot \dots \cdot X_n] = M[X_1] \cdot M[X_2] \cdot M[X_3] \cdot \dots \cdot M[X_n]. \quad (2.6)$$

Случайные величины называются *независимыми*, если каждая из них имеет самостоятельное распределение, не зависящее от возможных значений других величин.

4. Если случайная величина Z является некоторой нелинейной функцией n независимых случайных величин

$$Z = f(X_1, X_2, \dots, X_n),$$

которая мало меняется в небольших интервалах изменения аргументов, то

$$M[Z] = f(M[X_1], M[X_2], \dots, M[X_n]). \quad (2.7)$$

Дисперсией (вторым центральным моментом) случайной величины называется математическое ожидание квадрата отклонения случайной величины от ее математического ожидания, т. е.

$$D[X] = M[(X - m_x)^2]. \quad (2.8)$$

Для дискретной и непрерывной случайных величин дисперсия определяется следующим образом соответственно:

$$D[X] = \sum_{i=1}^n (x_i - m_x)^2 p_i, \quad (2.9)$$

$$D[X] = \int_{-\infty}^{+\infty} (x - m_x)^2 f(x) dx. \quad (2.10)$$

Другие обозначения для дисперсии: D_x , σ_x^2 , $\sigma^2(X)$.

Дисперсия играет важную роль при статистических расчетах и является мерой рассеяния значений x около их математического ожидания. Корень квадратный из второго центрального момента называется *средним квадратичным отклонением (стандартным отклонением, или стандартом)*:

$$\sigma_x = \sigma = \sqrt{D[X]}. \quad (2.11)$$

Свойства дисперсии:

1. Если c — постоянное число (неслучайная величина), то

$$\sigma^2(c) = 0, \quad (2.12)$$

$$\sigma^2(cX) = c^2 \sigma^2(X). \quad (2.13)$$

2. Дисперсия случайной величины равна математическому ожиданию квадрата случайной величины минус квадрат ее математического ожидания:

$$\sigma^2(X) = M[X^2] - m_x^2. \quad (2.14)$$

3. Дисперсия суммы независимых случайных величин равна сумме дисперсий этих величин:

$$\sigma^2(X_1 + X_2 + \dots + X_n) = \sigma^2(X_1) + \sigma^2(X_2) + \dots + \sigma^2(X_n). \quad (2.15)$$

Выражение (2.15) называют *законом сложения дисперсий*. Следует отметить, что закон сложения справедлив для дисперсий случайных величин (σ^2), а не среднеквадратичных отклонений (σ).

4. Если случайная величина Z является нелинейной функцией n независимых случайных величин

$$Z = f(X_1, X_2, \dots, X_n),$$

которая мало меняется в небольших интервалах изменения аргументов, то ее дисперсия приближенно равна

$$\begin{aligned} \sigma^2(Z) = & \left(\frac{\partial f}{\partial X_1} \right)^2 \sigma^2(X_1) + \left(\frac{\partial f}{\partial X_2} \right)^2 \sigma^2(X_2) + \\ & + \dots + \left(\frac{\partial f}{\partial X_n} \right)^2 \sigma^2(X_n). \end{aligned} \quad (2.16)$$

Выражение (2.16) называют *законом накопления ошибок*, и он часто используется в теории ошибок для определения случайной ошибки функции по значениям случайных ошибок аргументов.

Третий центральный момент, разделенный на σ_x^3 , называется *коэффициентом асимметрии* плотности распределения:

$$\gamma = \left(\int_{-\infty}^{+\infty} (x - m_x)^3 f(x) dx \right) / \sigma_x^3. \quad (2.17)$$

На рис. 5 приведены примеры плотностей распределения с одинаковыми математическим ожиданием и дисперсией, но с разными коэффициентами асимметрии.

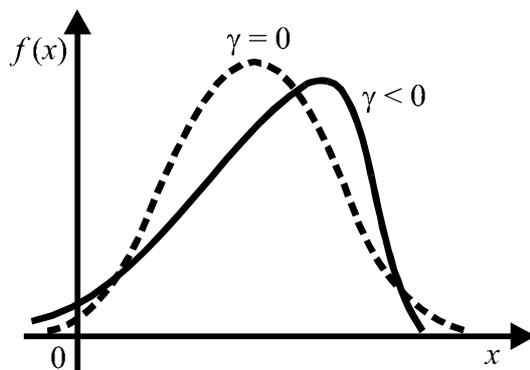


Рис. 5. Плотности распределения с нулевым и ненулевым коэффициентами асимметрии.

Если у случайной величины X существуют первый и второй моменты, то можно построить *нормированную случайную величину*

$$X_0 = \frac{X - m_x}{\sigma_x}, \quad (2.18)$$

для которой

$$M[X_0] = 0, D[X_0] = 1. \quad (2.19)$$

Докажем, что для нормированной случайной величины справедливы утверждения (2.19):

$$M[X_0] = M\left[\frac{X - m_x}{\sigma_x}\right] = \frac{1}{\sigma_x} M[X - m_x] = \frac{1}{\sigma_x} [M(X) - m_x] = 0,$$

$$D[X_0] = D\left[\frac{X - m_x}{\sigma_x}\right] = \frac{1}{\sigma_x^2} D(X - m_x) = \frac{1}{\sigma_x^2} [D(X) - 0] = \frac{D[X]}{\sigma_x^2} = 1.$$

Существуют следующие соотношения между функциями распределения, соответствующими нормированной X_0 и ненормированной X величинам:

$$f(x) = \frac{1}{\sigma_x} f_1(x_0) = \frac{1}{\sigma_x} f_1\left(\frac{x - m_x}{\sigma_x}\right), \quad (2.20)$$

$$f_1(x_0) = \sigma_x f(x) = \sigma_x f(m_x + \sigma_x x_0), \quad (2.21)$$

$$F(x) = F_1(x_0) = F_1\left(\frac{x - m_x}{\sigma_x}\right), \quad (2.22)$$

$$F_1(x_0) = F(x) = F(m_x + \sigma_x x_0). \quad (2.23)$$

Рассмотренные выше моменты являются общими (интегральными) характеристиками распределения случайной величины. Вторая группа параметров характеризует отдельные значения функции распределения. К ним относятся *квантили*. *Квантилем* x_β распределения случайной величины X с функцией распределения $F(x)$ называется решение уравнения $F(x_\beta) = \beta$, т. е. такое значение случайной величины, что $P(X \leq x_\beta) = \beta$. Наиболее важное значение имеет квантиль $x_{1/2}$, называемый *медианой распределения* (рис. 6).

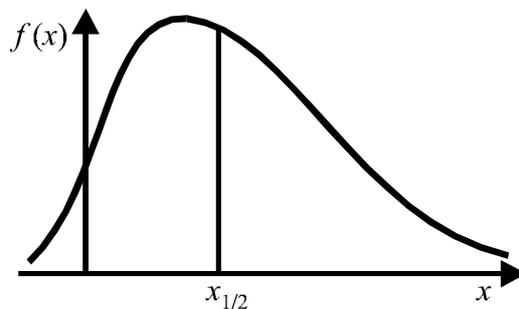


Рис. 6. Медиана распределения.

Ордината медианы пополам рассекает площадь между кривой плотности вероятности и осью абсцисс. Если распределение симметрично, то $x_{1/2} = m_x$.

2.2. Нормальное и стандартное распределения случайной величины. Функция Лапласа. Задача об абсолютном отклонении.

Непрерывная случайная величина X называется распределенной по *нормальному закону*, если ее плотность распределения имеет вид

$$f(x) = \frac{1}{\sqrt{2\pi} \sigma_x} \exp\left(-\frac{(x - m_x)^2}{2\sigma_x^2}\right), \quad (-\infty < x < +\infty), \quad (2.24)$$

где m_x и σ_x^2 — математическое ожидание и дисперсия случайной величины X .

Функция распределения равна

$$F(x) = \frac{1}{\sqrt{2\pi}\sigma_x} \int_{-\infty}^x \exp\left(-\frac{(x-m_x)^2}{2\sigma_x^2}\right) dx. \quad (2.25)$$

Нормальное распределение наиболее часто встречается на практике и теоретически наиболее полно разработано. Множество событий происходит случайно вследствие воздействия на них большого числа независимых (или слабо зависимых) возмущений, и у таких явлений закон распределения близок к нормальному. Установлено, что *нормальное распределение содержит минимум информации о случайной величине* по сравнению с любыми распределениями с той же дисперсией. Следовательно, замена некоторого распределения эквивалентным нормальным не может привести к переоценке точности наблюдений, что широко используется на практике.

График плотности нормального распределения называется *нормальной кривой*, или *кривой Гаусса* (рис. 7).

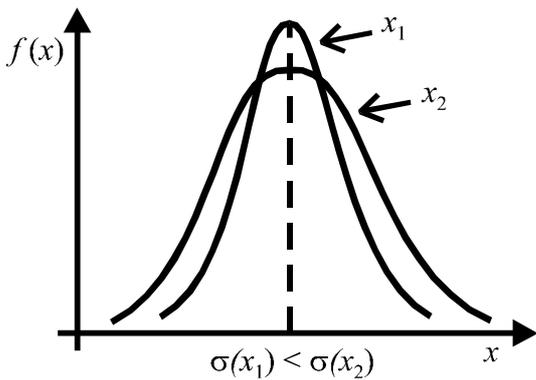


Рис. 7. Кривая Гаусса.

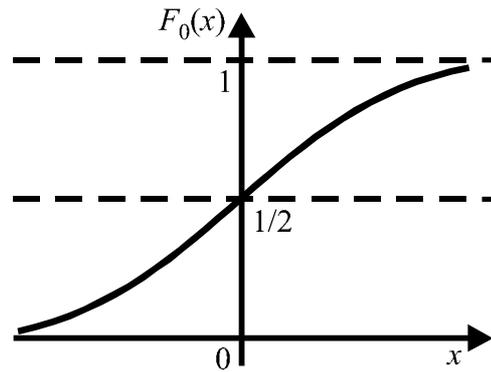


Рис. 8. График функции $F_0(x)$ стандартного распределения.

Нормальное распределение нормированной случайной величины называется *стандартным*. Его функция распределения имеет вид

$$F_0(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp(-x^2/2) dx, \quad (2.26)$$

а график этой функции представлен на рис. 8.

Вероятность того, что значения нормированной случайной величины будут лежать в интервале от x_{01} до x_{02} , равна

$$P(x_{01} \leq X_0 \leq x_{02}) = F_0(x_{02}) - F_0(x_{01}). \quad (2.27)$$

Функция

$$\Phi(X) = F_0(x) - 1/2 \quad (2.28)$$

называется *функцией Лапласа*

$$\Phi(X) = F_0(x) - 1/2 = F_0(x) - F_0(0) = \frac{1}{\sqrt{2\pi}} \int_0^x \exp(-x^2/2) dx. \quad (2.29)$$

Значения функции Лапласа табулированы (приложение 1). Так как она является нечетной функцией, т. е. $\Phi(-x) = -\Phi(x)$, то таблицы значений $\Phi(x)$ составлены лишь для $x > 0$.

Для нормированной случайной величины с учетом (2.27) и (2.28) имеем:

$$\begin{aligned} P(x_{01} \leq X_0 \leq x_{02}) &= F_0(x_{02}) - F_0(x_{01}) = \\ &= \Phi(x_{02}) + 1/2 - \Phi(x_{01}) - 1/2 = \Phi(x_{02}) - \Phi(x_{01}). \end{aligned} \quad (2.30)$$

Тогда в общем случае

$$\begin{aligned} P(x_1 \leq X \leq x_2) &= P\left(\frac{x_1 - m_x}{\sigma_x} \leq X_0 \leq \frac{x_2 - m_x}{\sigma_x}\right) = \\ &= \Phi\left(\frac{x_2 - m_x}{\sigma_x}\right) - \Phi\left(\frac{x_1 - m_x}{\sigma_x}\right). \end{aligned} \quad (2.31)$$

Во многих практических задачах x_1 и x_2 симметричны относительно математического ожидания, в частности *в задаче об абсолютном отклонении*. Абсолютным отклонением является величина

$$|\Delta x| = |X - m_x|. \quad (2.32)$$

Требуется найти вероятность того, что абсолютное отклонение случайной величины не превзойдет некоторого заданного числа ε :

$$P(|\Delta x| \leq \varepsilon) = P(m_x - \varepsilon \leq X \leq m_x + \varepsilon). \quad (2.33)$$

В частности, для нормированной случайной величины

$$P(|\Delta x_0| \leq \varepsilon) = P(-\varepsilon \leq X_0 \leq +\varepsilon) = \Phi(\varepsilon) - \Phi(-\varepsilon) = 2\Phi(\varepsilon). \quad (2.34)$$

Тогда для нормально распределенной случайной величины с параметрами m_x и σ_x справедливо

$$P(|\Delta x| \leq \varepsilon) = P\left(|\Delta x_0| \leq \frac{\varepsilon}{\sigma_x}\right) = 2\Phi\left(\frac{\varepsilon}{\sigma_x}\right). \quad (2.35)$$

Обозначив $\varepsilon/\sigma_x = k$, из (2.35) получаем

$$P(|\Delta x| \leq k\sigma_x) = 2\Phi(k), \quad (2.36)$$

откуда

$$P(|\Delta x| \leq \sigma_x) = 2\Phi(1) = 0.6826,$$

$$P(|\Delta x| \leq 2\sigma_x) = 2\Phi(2) = 0.9544,$$

$$P(|\Delta x| \leq 3\sigma_x) = 2\Phi(3) = 0.9973.$$

Таким образом, отклонения больше, чем утроенный стандарт (утроенное стандартное отклонение), практически невозможны. На практике часто величины $2\sigma_x$ (или $3\sigma_x$) считают максимально допустимой ошибкой и отбрасывают результаты измерений, для которых величина отклонения превышает это значение, как содержащие грубые ошибки.

Нормальное распределение обладает также свойством *линейности*: если независимые случайные величины X_1 и X_2 имеют нормальные распределения, то для произвольных чисел α и β величина

$$Y = \alpha X_1 + \beta X_2$$

также имеет нормальное распределение, причем из свойств математического ожидания и дисперсии следует, что

$$M[Y] = \alpha M[X_1] + \beta M[X_2], \quad (2.37)$$

$$\sigma[Y] = \sqrt{\alpha^2 \sigma^2[X_1] + \beta^2 \sigma^2[X_2]}. \quad (2.38)$$

ЛЕКЦИЯ 3

Генеральная совокупность и случайная выборка. Выборочная функция распределения. Гистограммы. Понятие об оценках параметров генерального распределения. Метод максимального правдоподобия. Оценка математического ожидания и дисперсии нормально распределенной случайной величины. Дисперсия среднего серии измерений.

3.1. Генеральная совокупность и случайная выборка.

Выборочная функция распределения. Гистограммы.

Понятие об оценках параметров генерального распределения.

Явление статистической устойчивости результатов наблюдений имеет место лишь при большом (в пределе — бесконечно большом) числе измерений. В подавляющем же числе экспериментов исследователю приходится иметь дело лишь с ограниченным, обычно небольшим, числом наблюдений. В силу закона случая какие-то величины, определенные по малому числу наблюдений, в общем случае могут не совпадать с теми же величинами, вычисленными по большому числу наблюдений, выполненных в тех же условиях. Поэтому в математической статистике вводят понятие абстрактной *генеральной совокупности*, состоящей из всех допустимых значений случайной величины, и *выборки*, представляющей собой совокупность ограниченного числа значений, полученных в результате опытов. В соответствии с этим различают *выборочные* характеристики случайной величины, найденные по ограниченному числу наблюдений и зависящие от этого числа, и соответствующие им характеристики генеральной совокупности. При этом выборочные характеристики рассматриваются как оценки соответствующих характеристик генеральной совокупности.

Выборка называется *репрезентативной* (представительной), если она дает достаточное представление об особенностях генеральной совокупности. Однако из случайного характера выборок следует, что любое суждение о генеральной совокупности само случайно. Предположим, что в результате эксперимента получена выборка из x_1, x_2, \dots, x_n значений случайной величины X . Обозначим через n_x число выборочных значений, расположенных левее x — некоторой точки числовой оси X . Отношение (n_x/n) есть частота появления значений X , меньших x , и является функцией от x . Эта функция, получаемая по выборке, называется *эмпирической*, или *выборочной функцией распре-*

деления (в отличие от распределения генеральной совокупности) и обозначается как

$$F_n(x) = n_x / n. \quad (3.1)$$

Можно доказать, что с вероятностью, равной 1, при $n \rightarrow \infty$ максимальная разность между функциями распределения случайной величины $F_n(x)$ и $F(x)$ стремится к нулю. На практике это означает, что при достаточно большой выборке функцию распределения генеральной совокупности приближенно можно заменять выборочной функцией распределения. Пусть $x_1 < x_2 < \dots < x_n$ (упорядоченная по величине выборка, или *вариационный ряд*). Все элементы выборки имеют одинаковую вероятность, равную $1/n$. Поэтому

$$\begin{aligned} F_n(x) &= 0 && \text{при } x < x_1, \\ F_n(x) &= k/n && \text{при } x_k \leq x < x_{k+1}, \text{ где } k = 1, 2, \dots, n-1, \\ F_n(x) &= 1 && \text{при } x \geq x_n. \end{aligned}$$

График $F_n(x)$ представлен на рис. 9. Все элементы выборки оказываются точками разрыва этой функции. В точке разрыва $x = x_k$ функция скачком переходит от значения $(k-1)/n$ к значению k/n , которое и удерживает в следующем интервале.

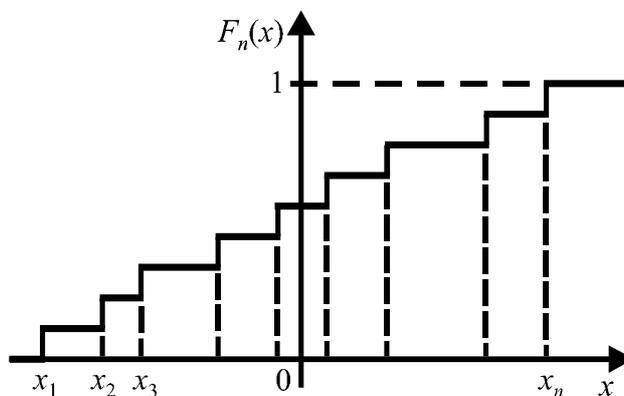


Рис. 9. Выборочная функция распределения.

При обработке выборок обычно используют метод «сгруппированных данных»: выборка объема n преобразуется в статистический ряд. Весь диапазон значений случайной величины от x_{\min} до x_{\max} делится на k равных интервалов ($j = 1, 2, \dots, k$). Число интервалов можно выбрать произвольно или по эмпирическим формулам, например:

$$k = 1 + 1.39 \ln n \quad (3.2)$$

с округлением до ближайшего целого. Длина интервала равна

$$h = (x_{\max} - x_{\min}) / k. \quad (3.3)$$

Число элементов выборки, попавших в j -интервал, обозначим через n_j . Величина

$$p_j^* = n_j / n \quad (3.4)$$

определяет относительную частоту попадания случайной величины в j -интервал. Все точки, попавшие в j -интервал, относят к его середине:

$$x_j^* = (x_{j-1} + x_j) / 2. \quad (3.5)$$

Статистический ряд записывается в виде табл. 1.

Таблица 1

Статистический ряд.

Интервал	Длина интервала	Середина интервала	Число точек в интервале	Относительная частота
1	(x_{\min}, x_1)	x_1^*	n_1	p_1^*
2	(x_1, x_2)	x_2^*	n_2	p_2^*
...
k	(x_{k-1}, x_{\max})	x_k^*	n_k	p_k^*
Σ			n	1

График, построенный по данным табл. 1, называется *гистограммой* эмпирического, или выборочного, распределения (рис. 10). На рис. 11 приведен график функции $F_n(x)$, построенный по сгруппированным данным.

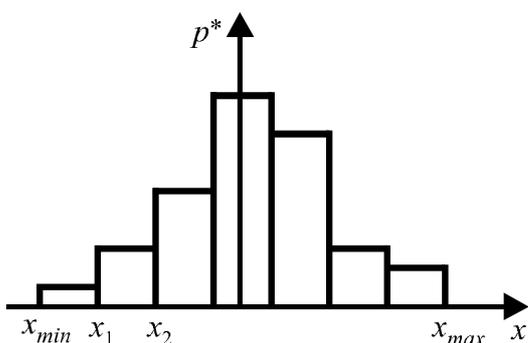


Рис. 10. Гистограмма распределения.

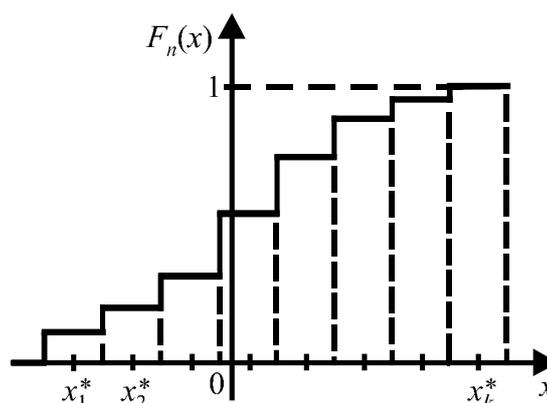


Рис. 11. График функции $F_n(x)$, построенный по сгруппированным данным.

При обработке результатов наблюдений обычно не удается получить эмпирическую функцию распределения. Однако даже простейший анализ условий опыта позволяет с достаточной уверенностью определять тип неизвестной функции распределения. Окончательное уточнение неизвестной функции распределения сводится к определению некоторых числовых параметров распределения. По выборкам могут быть рассчитаны выборочные статистические характеристики (выборочное среднее, дисперсия и т.д.), которые являются *оценками* соответствующих генеральных параметров.

Оценка $a^*(x_1, x_2, \dots, x_n)$ называется *состоятельной*, если с увеличением объема выборки n она стремится (по вероятности) к оцениваемому параметру a . *Эмпирические (выборочные) моменты являются состоятельными оценками теоретических моментов.*

Оценка $a^*(x_1, x_2, \dots, x_n)$ называется *несмещенной*, если ее математическое ожидание *при любом объеме выборки* равно оцениваемому параметру a , т. е. $M[a^*] = a$.

Важной характеристикой оценок генеральных параметров является также их *эффективность*, которая для различных несмещенных оценок одного и того же параметра при фиксированном объеме выборок обратно пропорциональна дисперсиям этих оценок.

3.2. Метод максимального правдоподобия.

Для получения точечных оценок используют различные методы. Широко применяется *метод максимального правдоподобия*. Сущность метода заключается в нахождении таких оценок неизвестных параметров, для которых функция правдоподобия при случайной выборке объема n будет иметь максимальное значение.

Пусть плотность распределения случайной величины X задается функцией $f(x, a)$, где a — неизвестный параметр, входящий в выражение закона распределения. На опыте получена выборка значений x_1, x_2, \dots, x_n . Окружим каждую точку x_i окрестностью длины δ . Тогда вероятность попадания в интервал с границами $(x_i - \delta/2), (x_i + \delta/2)$ приближенно равна $f(x, a) \delta$. Если произведено n наблюдений, то вероятность того, что одновременно первое наблюдение попадет в первый интервал, второе — во второй и т.д., есть вероятность совместного осуществления всех этих независимых событий и равна

$$\begin{aligned} P(x, a) &= f(x_1, a) \cdot f(x_2, a) \cdot \dots \cdot f(x_n, a) \cdot \delta^n = \\ &= f(x_1) \cdot f(x_2) \cdot \dots \cdot f(x_n) \cdot \delta^n. \end{aligned} \quad (3.6)$$

Так как событие с вероятностью P осуществилось на самом деле при первом же испытании, то естественно предположить, что ему соответствует максимальная вероятность. Поэтому в качестве оценки следует взять то значение a^* из области допустимых значений параметра a , для которого эта вероятность принимает наибольшее возможное значение, т.е. корень уравнения

$$\left. \frac{\partial P(x, a)}{\partial a} \right|_{a=a^*} = 0. \quad (3.7)$$

Достаточным условием максимума при этом является выполнение неравенства

$$\frac{\partial^2 P(x, a)}{\partial a^2} < 0. \quad (3.8)$$

Решение проще получить, если перейти к функции

$$L(x, a) = \ln \frac{P(x, a)}{\delta^n} = \sum_{i=1}^n \ln f(x_i, a), \quad (3.9)$$

которая называется *функцией правдоподобия*.

Вероятность P и функция L имеют максимумы при одних и тех же значениях определяемых параметров, так как

$$\frac{\partial L}{\partial a} = \frac{\partial}{\partial a} \ln P = \frac{1}{P} \frac{\partial P}{\partial a}, \quad P > 0. \quad (3.10)$$

В общем случае, когда требуется оценить одновременно несколько параметров одномерного или многомерного распределения, формулировка принципа максимального правдоподобия сохранится: надо найти такую совокупность допустимых значений параметров a_1^* , a_2^* , ..., a_k^* , которая обращает функцию правдоподобия в максимум.

Найдем методом максимального правдоподобия оценку параметра λ показательного распределения с плотностью

$$f(x) = \lambda \exp(-\lambda x), \quad 0 \leq x < \infty \quad (3.11)$$

по выборке x_1, x_2, \dots, x_n .

Функция правдоподобия примет следующий вид:

$$L = \sum_{i=1}^n \ln(\lambda \exp(-\lambda x_i)) = n \ln \lambda - \sum_{i=1}^n \lambda x_i. \quad (3.12)$$

Тогда

$$\frac{\partial L}{\partial \lambda} = \frac{n}{\lambda} - \sum_{i=1}^n x_i = 0, \quad (3.13)$$

$$\lambda^* = \frac{n}{\sum_{i=1}^n x_i} = \frac{1}{\bar{x}}, \quad (3.14)$$

где \bar{x} — среднее выборки.

3.3. Оценка математического ожидания и дисперсии нормально распределенной случайной величины. Дисперсия среднего серии измерений.

Пусть распределение случайной величины X подчинено нормальному закону

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right).$$

Тогда вероятность совместного осуществления n независимых событий $X = x_i$ ($i = 1, 2, \dots, n$) равна

$$P(x, m, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - m)^2\right) \cdot \delta^n \quad (3.15)$$

и функция правдоподобия

$$L(x, m, \sigma^2) = \ln \frac{P}{\delta^n} = -\frac{n}{2} \ln 2\pi - \frac{n}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - m)^2. \quad (3.16)$$

Продифференцируем (3.16) по m

$$\frac{\partial L}{\partial m} = \frac{2}{2\sigma^2} \sum_{i=1}^n (x_i - m) = 0. \quad (3.17)$$

Поскольку $1/\sigma^2 \neq 0$, то

$$\sum_{i=1}^n (x_i - m) = 0.$$

Тогда оценка для математического ожидания равна

$$m^* = \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad (3.18)$$

где \bar{x} — среднее арифметическое выборки (серии измерений). Отметим, что для выборочного среднего сохраняются все свойства математического ожидания. Например, если Z является нелинейной функцией n независимых случайных величин

$$Z = f(X_1, X_2, \dots, X_n),$$

то ее выборочное среднее приближенно выражается формулой

$$\bar{z} = f(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n).$$

Дифференцируя функцию правдоподобия (3.16) по σ^2 , получаем

$$\frac{\partial L}{\partial \sigma^2} = -\frac{n}{2} \frac{1}{\sigma^2} + \frac{1}{2(\sigma^2)^2} \sum_{i=1}^n (x_i - m)^2 = 0, \quad (3.19)$$

$$-\frac{1}{2\sigma^2} \left[n - \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - m)^2 \right] = 0. \quad (3.20)$$

Поскольку $1/(2\sigma^2) \neq 0$, то

$$n - \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - m)^2 = 0, \quad (3.21)$$

откуда находим оценку s_1^2 для дисперсии случайной величины:

$$(\sigma^2)^* = s_1^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2. \quad (3.22)$$

Метод максимального правдоподобия всегда приводит к состоятельным, хотя иногда и смещенным оценкам, имеющим наименьшую возможную дисперсию при неограниченном возрастании объема выборки. Так, выборочная дисперсия s_1^2 оказывается смещенной оценкой генеральной дисперсии

$$M[s_1^2] = \frac{n-1}{n} \sigma^2. \quad (3.23)$$

Для получения несмещенной оценки дисперсию s_1^2 надо умножить на величину $n/(n-1)$

$$s^2 = s_1^2 \frac{n}{n-1} = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}. \quad (3.24)$$

Уменьшение знаменателя в (3.24) на единицу непосредственно связано с тем, что величина \bar{x} , относительно которой берутся отклонения, сама зависит от элементов выборки. Каждая величина, зависящая от элементов выборки и входящая в формулу выборочной дисперсии, называется *связью*. Можно доказать, что знаменатель выборочной дисперсии всегда равен разности между объемом выборки n и числом связей l , наложенных на эту выборку. Эта разность

$$f = n - l \quad (3.25)$$

называется *числом степеней свободы* выборки.

В практических вычислениях для выборочной дисперсии s^2 часто более удобна следующая формула, получаемая из (3.24) путем арифметических преобразований:

$$s^2 = \frac{\sum_{i=1}^n x_i^2 - \frac{\left(\sum_{i=1}^n x_i\right)^2}{n}}{n-1}. \quad (3.26)$$

Итак, для нормально распределенной случайной величины получают по выборке следующие оценки генеральных параметров распределения: среднее арифметическое \bar{x} для математического ожидания m и выборочную дисперсию s^2 для генеральной дисперсии σ^2 .

Определим дисперсию среднего арифметического через дисперсию единичного наблюдения, воспользовавшись свойствами дисперсии. Если X_1, X_2, \dots, X_n — независимые случайные величины, a_1, a_2, \dots, a_n — неслучайные величины, а функция Z равна

$$Z = a_1 X_1 + a_2 X_2 + \dots + a_n X_n, \quad (3.27)$$

то дисперсия Z определяется следующим образом:

$$\sigma^2(Z) = a_1^2 \sigma^2(X_1) + a_2^2 \sigma^2(X_2) + \dots + a_n^2 \sigma^2(X_n). \quad (3.28)$$

Пусть в результате одной серии опытов получена выборка x_1, x_2, \dots, x_n . Если провести несколько серий подобных наблюдений, то в общем случае будут получены другие совокупности значений случай-

ной величины X : x_1', x_2', \dots, x_n' ; $x_1'', x_2'', \dots, x_n''$ и т.д. Поэтому значения x_1, x_2, \dots, x_n в серии из n наблюдений можно рассматривать как случайные величины с некоторыми дисперсиями $\sigma^2(x_1), \sigma^2(x_2), \dots, \sigma^2(x_n)$. Поскольку эти случайные величины возникают при измерении одной и той же случайной величины X , то дисперсии их естественно считать одинаковыми:

$$\sigma^2(x_1) = \sigma^2(x_2) = \dots = \sigma^2(x_n) = \sigma^2. \quad (3.29)$$

Применим теперь (3.28) для случая, когда Z является средним арифметическим (в этом случае $a_1 = a_2 = \dots = a_n = 1/n$):

$$\sigma^2(\bar{x}) = \frac{1}{n^2} [\sigma^2(x_1) + \sigma^2(x_2) + \dots + \sigma^2(x_n)] = \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n}. \quad (3.30)$$

Из (3.30) следует, что дисперсия среднего в n раз меньше дисперсии единичного измерения, поэтому для стандартного отклонения

$$\sigma(\bar{x}) = \frac{\sigma}{\sqrt{n}}. \quad (3.31)$$

Если принять $\sigma(\bar{x})$ в качестве меры случайной ошибки среднего выборки, то *увеличение числа параллельных определений одной и той же величины снижает величину случайной ошибки*. Это свойство случайной величины используют на практике для повышения точности результатов измерений.

Так как свойства генеральных дисперсий сохраняются и для их оценок — выборочных дисперсий, то

$$s^2(Z) = a_1^2 s^2(X_1) + a_2^2 s^2(X_2) + \dots + a_n^2 s^2(X_n), \quad (3.32)$$

$$s^2(\bar{x}) = \frac{s^2(X)}{n}, \quad s(\bar{x}) = \frac{s(X)}{\sqrt{n}}, \quad (3.33)$$

где s^2 — выборочные дисперсии, s — выборочное отклонение.

ЛЕКЦИЯ 4

Доверительные интервалы и доверительная вероятность, уровень значимости. Проверка статистических гипотез, критерии значимости, ошибки первого и второго рода. Построение доверительного интервала для математического ожидания непосредственно измеряемой величины. Распределение Стьюдента.

4.1. Доверительные интервалы и доверительная вероятность, уровень значимости.

Выборочные параметры распределения, определяемые по серии измерений, являются случайными величинами, следовательно, и их отклонения от генеральных параметров также будут случайными. Оценка этих отклонений носит вероятностный характер — при статистическом анализе можно лишь указать вероятность той или иной погрешности.

Пусть для генерального параметра a получена из опыта несмещенная оценка a^* . Назначим достаточно большую вероятность β (такую, что событие с вероятностью β можно считать практически достоверным) и найдем такое значение $\varepsilon_\beta = f(\beta)$, для которого

$$P(|a^* - a| \leq \varepsilon_\beta) = \beta. \quad (4.1)$$

Диапазон практически возможных значений ошибки, возникающей при замене a на a^* , будет $\pm \varepsilon_\beta$. Большие по абсолютной величине ошибки будут появляться только с малой вероятностью

$$p = 1 - \beta, \quad (4.2)$$

называемой *уровнем значимости*. Иначе выражение (4.1) можно интерпретировать как вероятность того, что истинное значение параметра a лежит в пределах

$$a^* - \varepsilon_\beta \leq a \leq a^* + \varepsilon_\beta. \quad (4.3)$$

Вероятность β называется *доверительной вероятностью* и характеризует надежность полученной оценки. Интервал $I_\beta = a^* \pm \varepsilon_\beta$ называется *доверительным интервалом*. Границы интервала $a' = a^* - \varepsilon_\beta$ и $a'' = a^* + \varepsilon_\beta$ называются *доверительными границами*. Доверительный интервал при данной доверительной вероятности определяет точность оценки. Величина доверительного интервала зависит от доверительной вероятности, с которой гарантируется нахождение параметра a

внутри доверительного интервала: чем больше величина β , тем больше интервал I_β (и величина ε_β). Увеличение числа опытов проявляется в сокращении доверительного интервала при постоянной доверительной вероятности или в повышении доверительной вероятности при сохранении доверительного интервала.

На практике обычно фиксируют значение доверительной вероятности (0,9; 0,95 или 0,99) и затем определяют доверительный интервал результата I_β . При построении доверительного интервала решается задача об абсолютном отклонении:

$$P\left(|a^* - a| \leq \varepsilon_\beta\right) = P(|\Delta a| \leq \varepsilon_\beta) = F(\varepsilon_\beta) - F(-\varepsilon_\beta) = \int_{-\varepsilon_\beta}^{\varepsilon_\beta} f(a) da = \beta. \quad (4.4)$$

Таким образом, если бы был известен закон распределения оценки a^* , задача определения доверительного интервала решалась бы просто. Рассмотрим построение доверительного интервала для математического ожидания нормально распределенной случайной величины X с известным генеральным стандартом σ по выборке объемом n . Наилучшей оценкой для математического ожидания m является среднее выборки \bar{x} со стандартным отклонением среднего

$$\sigma(\bar{x}) = \sigma / \sqrt{n}.$$

Используя функцию Лапласа, получаем

$$P\left(|\bar{x} - m_x| \leq \varepsilon_\beta\right) = \beta = 2\Phi\left(\frac{\varepsilon_\beta}{\sigma(\bar{x})}\right). \quad (4.5)$$

Задавшись доверительной вероятностью β , определим по таблице функции Лапласа (приложение 1) величину $k_\beta = \varepsilon_\beta / \sigma(\bar{x})$. Тогда доверительный интервал для математического ожидания принимает вид

$$\bar{x} - k_\beta \sigma(\bar{x}) \leq m_x \leq \bar{x} + k_\beta \sigma(\bar{x}), \quad (4.6)$$

или

$$\bar{x} - k_\beta \frac{\sigma}{\sqrt{n}} \leq m_x \leq \bar{x} + k_\beta \frac{\sigma}{\sqrt{n}}. \quad (4.7)$$

Из (4.7) видно, что уменьшение доверительного интервала обратно пропорционально корню квадратному из числа опытов.

Знание генеральной дисперсии позволяет оценивать математическое ожидание даже по одному наблюдению. Если для нормально

распределенной случайной величины X в результате эксперимента получено значение x_1 , то доверительный интервал для математического ожидания при выбранной β имеет вид

$$x_1 - \sigma U_{1-p/2} \leq m_x \leq x_1 + \sigma U_{1-p/2}, \quad (4.8)$$

где $U_{1-p/2}$ — квантиль стандартного нормального распределения (приложение 2).

Закон распределения оценки a^* зависит от закона распределения величины X и, в частности, от самого параметра a . Чтобы обойти это затруднение, в математической статистике применяют два метода:

1) приближенный — при $n \geq 50$ заменяют в выражении для ε_β неизвестные параметры их оценками, например:

$$k_\beta = \varepsilon_\beta / \sigma(\bar{x}) \approx \varepsilon_\beta / s(\bar{x});$$

2) от случайной величины a^* переходят к другой случайной величине Θ^* , закон распределения которой не зависит от оцениваемого параметра a , а зависит только от объема выборки n и от вида закона распределения величины X . Такого рода величины наиболее подробно изучены для нормального распределения случайных величин. В качестве доверительных границ Θ' и Θ'' обычно используются симметричные квантили

$$\Theta_{(1-\beta)/2} \leq \Theta^* \leq \Theta_{(1+\beta)/2}, \quad (4.9)$$

или с учетом (4.2)

$$\Theta_{p/2} \leq \Theta^* \leq \Theta_{1-p/2}. \quad (4.10)$$

4.2. Проверка статистических гипотез, критерии значимости, ошибки первого и второго рода.

Под *статистическими гипотезами* понимаются некоторые предположения относительно распределений генеральной совокупности той или иной случайной величины. Под проверкой гипотезы понимают сопоставление некоторых статистических показателей, *критериев проверки (критериев значимости)*, вычисляемых по выборке, с их значениями, определенными в предположении, что данная гипотеза верна. При проверке гипотез обычно подвергается испытанию некоторая гипотеза H_0 в сравнении с альтернативной гипотезой H_1 .

Чтобы решить вопрос о принятии или непринятии гипотезы, задаются уровнем значимости p . Наиболее часто используются уровни

значимости, равные 0.10, 0.05 и 0.01. По этой вероятности, используя гипотезу о распределении оценки Θ^* (критерия значимости), находят квантильные доверительные границы, как правило, симметричные $\Theta_{p/2}$ и $\Theta_{1-p/2}$. Числа $\Theta_{p/2}$ и $\Theta_{1-p/2}$ называются *критическими значениями гипотезы*; значения $\Theta^* < \Theta_{p/2}$ и $\Theta^* > \Theta_{1-p/2}$ образуют критическую область гипотезы (или область непринятия гипотезы) (рис. 12).

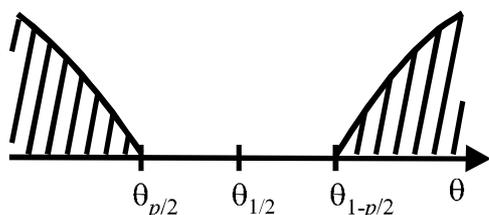


Рис. 12. Критическая область гипотезы.

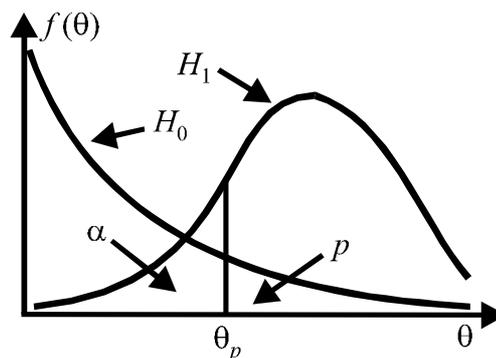


Рис. 13. Проверка статистических гипотез.

Если найденное по выборке Θ_0 попадает между $\Theta_{p/2}$ и $\Theta_{1-p/2}$, то гипотеза допускает такое значение в качестве случайного и поэтому нет оснований ее отвергать. Если же значение Θ_0 попадает в критическую область, то по данной гипотезе оно является практически невозможным. Но поскольку оно появилось, то отвергается сама гипотеза.

При проверке гипотез можно совершить ошибки двух типов. *Ошибка первого рода* состоит в том, что *отвергается гипотеза, которая на самом деле верна*. Вероятность такой ошибки не больше принятого уровня значимости. *Ошибка второго рода* состоит в том, что *гипотеза принимается, а на самом деле она неверна*. Вероятность этой ошибки тем меньше, чем выше уровень значимости, так как при этом увеличивается число отвергаемых гипотез. Если вероятность ошибки второго рода равна α , то величину $(1 - \alpha)$ называют *мощностью критерия*.

На рис. 13 приведены две кривые плотности распределения случайной величины Θ , соответствующие двум гипотезам H_0 и H_1 . Если из опыта получается значение $\Theta > \Theta_p$, то отвергается гипотеза H_0 и принимается гипотеза H_1 , и наоборот, если $\Theta < \Theta_p$.

Площадь под кривой плотности вероятности, соответствующей справедливости гипотезы H_0 вправо от значения Θ_p , равна уровню

значимости p , т. е. вероятности ошибки первого рода. Площадь под кривой плотности вероятности, соответствующей справедливости гипотезы H_1 влево от Θ_p , равна вероятности ошибки второго рода α , а вправо от Θ_p — мощности критерия $(1 - \alpha)$. Таким образом, чем больше p , тем больше $(1 - \alpha)$. При проверке гипотезы стремятся из всех возможных критериев выбрать тот, у которого при заданном уровне значимости меньше вероятность ошибки второго рода.

Обычно в качестве оптимального уровня значимости при проверке гипотез используют $p = 0,05$, так как если проверяемая гипотеза принимается с данным уровнем значимости, то гипотезу, безусловно, следует признать согласующейся с экспериментальными данными; с другой стороны, использование данного уровня значимости не дает оснований для отбрасывания гипотезы.

Например, найдены два значения a_1^* и a_2^* некоторого выборочного параметра, которые можно рассматривать как оценки генеральных параметров a_1 и a_2 . Высказывается гипотеза, что различие между a_1^* и a_2^* случайное и что генеральные параметры a_1 и a_2 равны между собой, т. е. $a_1 = a_2$. Такая гипотеза называется *нулевой*, или *нуль-гипотезой*. Для ее проверки нужно выяснить, значимо ли расхождение между a_1^* и a_2^* в условиях нулевой гипотезы. Для этого обычно исследуют случайную величину $\Delta a^* = a_1^* - a_2^*$ и проверяют, значимо ли ее отличие от нуля. Иногда удобнее рассматривать величину a_1^*/a_2^* , сравнивая ее с единицей.

Отвергая нулевую гипотезу, тем самым принимают альтернативную, которая распадается на две: $a_1^* > a_2^*$ и $a_1^* < a_2^*$. Если одно из этих равенств заведомо невозможно, то альтернативная гипотеза называется *односторонней*, и для ее проверки применяют *односторонние* критерии значимости (в отличие от обычных, *двусторонних*). При этом необходимо рассматривать лишь одну из половин критической области (рис. 12).

Например, $p = 0,05$ при двустороннем критерии соответствуют критические значения $\Theta_{0,025}$ и $\Theta_{0,975}$, т. е. значимыми (неслучайными) считаются Θ^* , принявшие значения $\Theta^* < \Theta_{0,025}$ и $\Theta^* > \Theta_{0,975}$. При одностороннем критерии одно из этих неравенств заведомо невозможно (например, $\Theta^* < \Theta_{0,025}$) и значимыми будут лишь $\Theta^* > \Theta_{0,975}$. Вероятность последнего неравенства равна 0,025, и, следовательно, уровень значимости будет равен 0,025. Таким образом, если при односторон-

нем критерии значимости использовать те же критические числа, что и при двустороннем, этим значениям будет соответствовать вдвое меньший уровень значимости.

Обычно для одностороннего критерия берут тот же уровень значимости, что и для двустороннего, так как при этих условиях оба критерия обеспечивают одинаковую ошибку первого рода. *Для этого односторонний критерий надо выводить из двустороннего, соответствующего вдвое большему уровню значимости, чем тот, что принят.* Чтобы сохранить для одностороннего критерия уровень значимости $p = 0,05$, для двустороннего необходимо взять $p = 0,10$, что дает критические значения $\Theta_{0,05}$ и $\Theta_{0,95}$. Из них для одностороннего критерия останется какое-нибудь одно, например, $\Theta_{0,95}$. Уровень значимости для одностороннего критерия равен при этом 0.05. Этому же уровню значимости для двустороннего критерия соответствует критическое значение $\Theta_{0,975}$. Но $\Theta_{0,95} < \Theta_{0,975}$, значит, при одностороннем критерии большее число гипотез будет отвергнуто и, следовательно, меньше будет ошибка второго рода.

4.3. Построение доверительного интервала для математического ожидания непосредственно измеряемой величины.

Распределение Стьюдента.

При отсутствии грубых и систематических ошибок математическое ожидание случайной величины совпадает с истинным результатом наблюдений. Легче всего оценить математическое ожидание при известной дисперсии генеральной совокупности (выражения 4.6 – 4.8). Однако значение σ^2 нельзя получить из наблюдений, ее можно только оценить при помощи выборочной дисперсии s^2 . Ошибка от этой замены будет тем меньше, чем больше объем выборки n . На практике эту погрешность не учитывают при $n \leq 50$ и в формуле (4.7) для доверительного интервала генеральный параметр σ заменяют выборочным стандартом. В дальнейшем примем, что наблюдаемая случайная величина имеет нормальное распределение.

При небольших объемах выборок для построения доверительного интервала математического ожидания используют *распределение Стьюдента, или t-распределение*. Распределение Стьюдента имеет величина t

$$t = \frac{\bar{x} - m_x}{s_x} \sqrt{n} \quad (4.11)$$

с плотностью вероятности

$$\varphi(t) = \frac{1}{\sqrt{\pi f}} \frac{\Gamma\left(\frac{f+1}{2}\right)}{\Gamma\left(\frac{f}{2}\right)} \left(1 + \frac{t^2}{f}\right)^{-\left(\frac{f+1}{2}\right)}, \quad -\infty < t < +\infty, \quad (4.12)$$

где $\Gamma(f)$ — гамма-функция Эйлера:

$$\Gamma(z) = \int_0^{\infty} e^{-y} y^{z-1} dy; \quad (4.13)$$

f — число степеней свободы выборки. Если дисперсия s^2 и среднее \bar{x} определяются по одной и той же выборке, то $f = n - 1$.

Распределение Стьюдента зависит только от числа степеней свободы f , с которым определена выборочная дисперсия. На рис. 14 приведены графики плотности t -распределения для нескольких чисел свободы f и нормальная кривая.

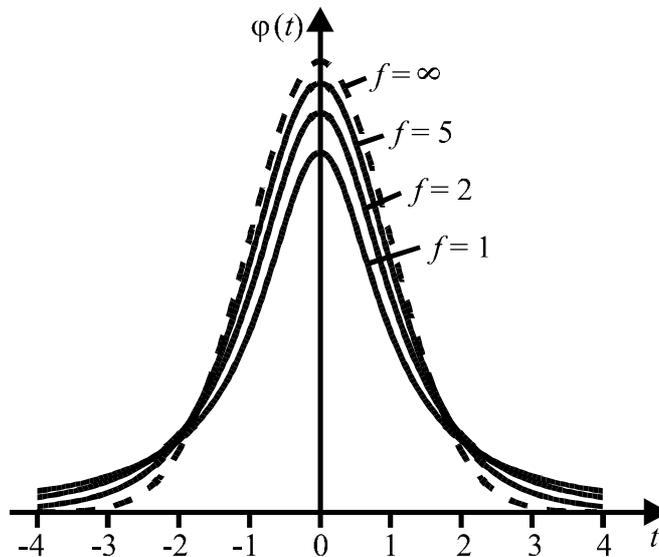


Рис. 14. Плотность распределения Стьюдента.

Кривые t -распределения по своей форме напоминают нормальную кривую, но при малых f они медленнее сближаются с осью абсцисс при $|t| \rightarrow \infty$. При $f \rightarrow \infty$ $s^2 \rightarrow \sigma^2$, поэтому распределение Стьюдента сближается (в пределе соответствует) с нормальным распределением.

Вероятность того, что случайная величина попадет в интервал $(t_{p/2}; t_{1-p/2})$, определяется выражением

$$P(t_{p/2} \leq t \leq t_{1-p/2}) = 1 - p = \beta. \quad (4.14)$$

Распределение Стьюдента симметрично относительно нуля, поэтому

$$t_{p/2} = -t_{1-p/2}. \quad (4.15)$$

Учитывая симметрию t -распределения, часто пользуются обозначением $t_p(f)$, где f — число степеней свободы, p — уровень значимости, т. е. вероятность того, что t находится за пределами интервала $(t_{p/2}; t_{1-p/2})$. Подставляя в (4.14) выражение для t (4.11) с учетом (4.15), получаем неравенство

$$-t_{1-p/2} \leq \frac{\bar{x} - m_x}{S_x} \sqrt{n} \leq t_{1-p/2}, \quad (4.16)$$

и после преобразований имеем

$$\bar{x} - \frac{S_x}{\sqrt{n}} t_{1-p/2} \leq m_x \leq \bar{x} + \frac{S_x}{\sqrt{n}} t_{1-p/2}. \quad (4.17)$$

Значения квантилей $t_{1-p/2}$ для различных чисел степеней свободы f и уровней значимости p приведены в приложении 3. Выражение (4.17) означает, что интервал с доверительными границами

$$(\bar{x} - s(\bar{x}) t_{1-p/2}) \div (\bar{x} + s(\bar{x}) t_{1-p/2}) \quad (4.18)$$

накрывает с вероятностью β генеральное среднее измеряемой величины. Величина доверительного интервала (4.18) определяет надежность среднего выборки. Величину

$$s(\bar{x}) t_{1-p/2} = \frac{S_x}{\sqrt{n}} t_{1-p/2} = \varepsilon_{\text{случ}}, \quad (4.19)$$

т. е. половину доверительного интервала, называют *случайной ошибкой*. С учетом только случайной ошибки результат измерений некоторой величины следует записывать так:

$$X = \bar{x} \pm \varepsilon_{\text{случ}} = \bar{x} \pm \frac{S_x}{\sqrt{n}} t_{1-p/2}. \quad (4.20)$$

ЛЕКЦИЯ 5

Оценка случайной и суммарной ошибки косвенных измерений. Оценка дисперсии нормально распределенной случайной величины; распределение Пирсона. Сравнение двух дисперсий, распределение Фишера.

5.1. Оценка случайной и суммарной ошибки косвенных измерений.

В самом общем виде пример косвенных измерений формулируется следующим образом: имеется известная функция нескольких аргументов

$$Z = f(X_1, X_2, \dots, X_k),$$

причем на опыте непосредственно измеряются случайные величины X_1, X_2, \dots, X_k . При строгом статистическом анализе случайной ошибки Z необходимо найти закон распределения функции по известным законам распределения аргументов, что связано с большими вычислительными трудностями. Из-за этого строгая оценка ошибки косвенных измерений трудно выполнима и практически нецелесообразна. Поэтому используются упрощенные подходы, значительно облегчающие расчеты и вместе с тем дающие удовлетворительные для практических целей результаты.

Рассмотрим вначале случай, когда Z является известной функцией только одного параметра X : $Z = f(X)$. Введем допущение о том, что в небольших интервалах изменения нормально распределенного аргумента функция этого аргумента также подчиняется нормальному закону распределения. Пусть x_1, x_2, \dots, x_n — результаты n измерений величины X . Для каждого из x_i можно найти соответствующее значение z_i , затем вычислить среднее \bar{z} и выборочную дисперсию $s^2(Z)$ с числом степеней свободы $f = n - 1$. Тогда согласно изложенному в предыдущем разделе имеем

$$\varepsilon_{\text{случ}}(Z) = s(\bar{z})t_{1-p/2} = \frac{s(Z)}{\sqrt{n}}t_{1-p/2}. \quad (5.1)$$

При учете только случайной ошибки результат измерений функции следует записать так:

$$Z = \bar{z} \pm \varepsilon_{\text{случ}}(Z) = \bar{z} \pm \frac{s(Z)}{\sqrt{n}}t_{1-p/2}. \quad (5.2)$$

Если $f(x)$ является достаточно сложной функцией и каждый раз вычисление величины z_i по значению x_i трудоемко, то можно определить сначала величины \bar{x} и $s(X)$, а затем пересчитать их в соответствующие величины \bar{z} и $s(Z)$ при помощи приближенных формул:

$$\bar{z} = f(\bar{x}), \quad (5.3)$$

$$s(Z) = \left| \frac{\partial f}{\partial X} \right|_{X=\bar{x}} s(X). \quad (5.4)$$

Для случая, когда Z является известной функцией нескольких аргументов, используем следующие допущения:

- 1) Случайные величины X_1, X_2, \dots, X_k независимы.
- 2) В небольших интервалах изменения аргументов функция Z распределена нормально.
- 3) Выборочная дисперсия величины \bar{z} равна соответствующей генеральной

$$s^2(\bar{z}) = \sigma^2(\bar{z}). \quad (5.5)$$

Оценка случайной ошибки функции проводится в следующем порядке. Находим среднее функции:

$$\bar{z} = f(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_k), \quad (5.6)$$

где $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_k$ — средние по выборкам соответствующих аргументов. Затем по закону накопления ошибок оцениваем выборочную дисперсию

$$s^2(\bar{z}) = \sum_{j=1}^k \left(\frac{\partial f}{\partial X_j} \right)_{X_j=\bar{x}_j}^2 s^2(\bar{x}_j). \quad (5.7)$$

Тогда величина случайной ошибки функции определяется следующим образом:

$$\varepsilon_{\text{случ}}(Z) = U_{1-p/2} s(\bar{z}), \quad (5.8)$$

где $U_{1-p/2}$ — квантиль стандартного нормального распределения (приложение 2), равный 1,96 для уровня значимости $p = 0,05$.

При учете только случайной ошибки для доверительной вероятности $\beta = 0,95$ результат измерений функции нескольких аргументов следует записать так:

$$Z = \bar{z} \pm \varepsilon_{\text{случ}}(Z) = \bar{z} \pm U_{1-p/2} s(\bar{z}) = \bar{z} \pm 1.96 \cdot s(\bar{z}) \approx \bar{z} \pm 2 s(\bar{z}). \quad (5.9)$$

Случайную ошибку косвенных измерений можно оценить также, воспользовавшись формулами расчета погрешностей функций приближенных аргументов (лекция 1) для случая, когда погрешности аргументов независимы и случайны:

$$\varepsilon_{\text{случ}}(Z) \approx 2s(\bar{z}) = \sqrt{\sum_{j=1}^k \left(\frac{\partial f}{\partial X_j} \Delta x_j \right)^2} = \sqrt{\sum_{j=1}^k \left(\frac{\partial f}{\partial X_j} \right)^2 (2s(\bar{x}_j))^2}, \quad (5.10)$$

при этом в качестве абсолютной погрешности аргументов следует использовать удвоенное значение среднеквадратичных отклонений их средних

$$\Delta x_j = 2s(\bar{x}_j). \quad (5.11)$$

В общем случае при представлении результатов измерений следует учитывать не только случайную, но и систематическую ошибку методики или прибора. Предполагая, что эти два типа ошибки взаимонезависимы, суммарная ошибка измерений равна:

$$\varepsilon_{\text{сумм}} = \varepsilon_{\text{сист}} + \varepsilon_{\text{случ}}. \quad (5.12)$$

Систематические ошибки являются величинами, не зависящими от числа измерений, и определяются спецификой используемой аппаратуры и методом измерений. Так, например, с помощью ртутного термометра нельзя измерить температуру с точностью, большей $0,01^\circ\text{C}$ (редко $0,005^\circ\text{C}$); значение эталонного сопротивления может быть известно с точностью $0,1\%$ или $0,01\%$; и т. д. Если и источники, и величины систематических ошибок определены, то их влияние на окончательный результат косвенных измерений для функции нескольких аргументов можно оценить как предельную абсолютную погрешность по формуле (лекция 1)

$$\varepsilon_{\text{сист}} = \varepsilon_{\text{пр}} \approx \sum_{j=1}^k \left| \frac{\partial f}{\partial X_j} \right| |\Delta x_j|. \quad (5.13)$$

Величина систематической ошибки ограничивает число верных значащих цифр при представлении результатов эксперимента. С учетом систематической ошибки результат любого измерения следует записывать следующим образом:

$$Z = \bar{z} \pm \varepsilon_{\text{сумм}} = \bar{z} \pm (\varepsilon_{\text{сист}} + \varepsilon_{\text{случ}}). \quad (5.14)$$

5.2. Оценка дисперсии нормально распределенной случайной величины.

Дисперсию генеральной совокупности σ^2 нормальной распределенной случайной величины можно оценить, если известно распределение ее оценки — выборочной дисперсии s^2 . Распределение выборочной дисперсии можно получить при помощи распределения Пирсона или χ^2 -распределения.

Пусть имеется выборка n независимых наблюдений x_1, x_2, \dots, x_n над нормально распределенной случайной величиной. Можно показать, что сумма

$$\chi^2 = \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{\sigma} \right)^2 \quad (5.15)$$

имеет распределение с $f = n - 1$ степенями свободы. Плотность χ^2 распределения зависит только от числа степеней свободы f :

$$\varphi(\chi^2) = \frac{1}{2^{f/2} \Gamma(f/2)} (\chi^2)^{f/2 - 1} e^{-\chi^2/2}, \quad 0 \leq \chi^2 < \infty, \quad (5.16)$$

где $\Gamma(f)$ — гамма-функция. На рис. 15 приведены кривые плотности вероятности χ^2 распределения при некоторых значениях f . Кривые асимметричны, степень асимметрии уменьшается с увеличением f .

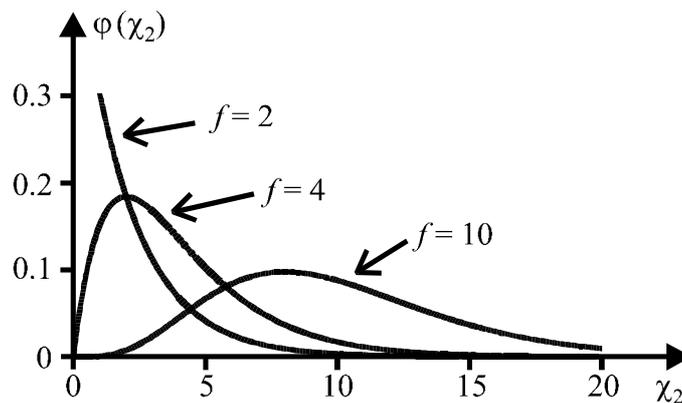


Рис. 15. Плотность χ^2 -распределения.

При доверительной вероятности $\beta = 1 - p$ двусторонняя доверительная оценка для χ^2 имеет вид

$$\chi_{p/2}^2 \leq \chi^2 \leq \chi_{1-p/2}^2, \quad (5.17)$$

односторонние оценки имеют вид

$$\chi^2 \leq \chi_{1-p}^2, \quad \chi^2 \geq \chi_p^2. \quad (5.18)$$

Квантили χ_{1-p}^2 при различных p и f приведены в приложении 4.

Поскольку выборочная дисперсия определяется по формуле

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{f},$$

то с учетом (5.15) имеем:

$$\chi^2 = f s^2 / \sigma^2. \quad (5.19)$$

Подставляя (5.19) в (5.17) и решая полученное неравенство относительно σ^2 , получим доверительные двусторонние границы для генеральной дисперсии:

$$\chi_{p/2}^2 \leq f s^2 / \sigma^2 \leq \chi_{1-p/2}^2, \quad (5.20)$$

$$\frac{f s^2}{\chi_{1-p/2}^2} \leq \sigma^2 \leq \frac{f s^2}{\chi_{p/2}^2}. \quad (5.21)$$

Аналогично получаются односторонние доверительные оценки:

$$\sigma^2 \leq f s^2 / \chi_p^2, \quad \sigma^2 \geq f s^2 / \chi_{1-p}^2. \quad (5.22)$$

С ростом числа степеней свободы асимметрия кривых χ^2 -распределения уменьшается, соответственно уменьшается и асимметрия доверительных границ. Можно показать, что при $n \geq 30$ выборочный стандарт s распределен приблизительно нормально с математическим ожиданием $m_s = \sigma$ и среднеквадратичной ошибкой

$$\sigma_s = \sigma / \sqrt{2f}. \quad (5.23)$$

Неизвестный генеральный стандарт в (5.23) при $n \geq 30$ заменяют выборочным

$$\sigma_s \approx s / \sqrt{2f}. \quad (5.24)$$

Тогда по уравнению (4.8) (лекция 4) доверительные границы для генерального стандарта определяются неравенством

$$s - (s / \sqrt{2f}) U_{1-p/2} \leq \sigma \leq s + (s / \sqrt{2f}) U_{1-p/2}. \quad (5.25)$$

5.3. Сравнение двух дисперсий. Распределение Фишера.

При обработке результатов измерений часто бывает необходимым сравнить две или несколько выборочных дисперсий. Основная гипотеза, которая при этом проверяется, следующая: можно ли считать сравниваемые выборочные дисперсии оценками одной и той же генеральной дисперсии? Рассмотрим две выборки

$$x_1', x_2', \dots, x_{n_1}' \text{ и } x_1'', x_2'', \dots, x_{n_2}'',$$

средние значения которых равны \bar{x}_1 и \bar{x}_2 . Выборочные дисперсии определяются со степенями свободы $f_1 = n_1 - 1$ и $f_2 = n_2 - 1$:

$$s_1^2 = \frac{\sum_{i=1}^{n_1} (x_i' - \bar{x}_1)^2}{f_1}; \quad s_2^2 = \frac{\sum_{i=1}^{n_2} (x_i'' - \bar{x}_2)^2}{f_2}. \quad (5.26)$$

Требуется выяснить, являются ли выборочные дисперсии s_1^2 и s_2^2 значимо различными или же полученные выборки можно рассматривать как взятые из генеральных совокупностей с равными дисперсиями.

Допустим, что первая выборка была взята из генеральной совокупности с дисперсией σ_1^2 , а вторая — из генеральной совокупности с дисперсией σ_2^2 . Проверяется *нулевая гипотеза* о равенстве генеральных дисперсий $H_0: \sigma_1^2 = \sigma_2^2$. Чтобы отвергнуть эту гипотезу, нужно доказать значимость различия между s_1^2 и s_2^2 при выбранном уровне значимости p .

В качестве критерия значимости обычно используется *критерий Фишера*. *Распределением Фишера* (F -распределением, ν^2 -распределением) называется распределение случайной величины

$$F = \frac{(s_1^2 / \sigma_1^2)}{(s_2^2 / \sigma_2^2)}. \quad (5.27)$$

Плотность F -распределения определяется выражением

$$\varphi(F) = \frac{\Gamma\left(\frac{f_1 + f_2}{2}\right)}{\Gamma\left(\frac{f_1}{2}\right)\Gamma\left(\frac{f_2}{2}\right)} \left(\frac{f_1}{f_2}\right)^{f_1/2} \frac{F^{(f_1-2)/2}}{\left(1 + \frac{f_1}{f_2}F\right)^{(f_1+f_2)/2}}, \quad 0 \leq F \leq \infty, \quad (5.28)$$

где $\Gamma(f)$ — гамма-функция. Распределение Фишера зависит только от числа степеней свободы f_1 и f_2 . На рис. 16 приведены кривые плотности вероятности F -распределения для некоторых значений f_1 и f_2 . Кривые имеют асимметричную форму.

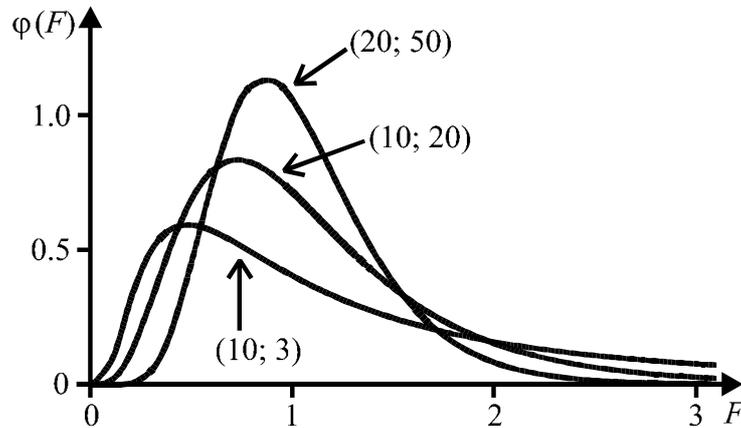


Рис. 16. Плотность F -распределения.

В приложении 5 приведены квантили F_{1-p} (критерии Фишера) для уровня значимости $p = 0,05$. Для определения квантилей F_p используется соотношение

$$F_p(f_1, f_2) = \frac{1}{F_{1-p}(f_2, f_1)}. \quad (5.29)$$

В условиях нулевой гипотезы $\sigma_1^2 = \sigma_2^2$ и $\sigma_1^2/\sigma_2^2 = 1$ и, следовательно, F -распределение может быть непосредственно использовано для оценки отношения s_1^2/s_2^2 . При доверительной вероятности $(1 - p)$ двусторонняя оценка величины F имеет вид

$$F_{p/2}(f_1, f_2) \leq F \leq F_{1-p/2}(f_1, f_2) \quad (5.30)$$

или с учетом (5.29)

$$\frac{1}{F_{1-p/2}(f_2, f_1)} \leq F \leq F_{1-p/2}(f_1, f_2). \quad (5.31)$$

В условиях нулевой гипотезы $F = s_1^2/s_2^2$ и, следовательно, с вероятностью $(1 - p)$ должно выполняться двустороннее неравенство

$$\frac{1}{F_{1-p/2}(f_2, f_1)} \leq \frac{s_1^2}{s_2^2} \leq F_{1-p/2}(f_1, f_2) \quad (5.32)$$

или одно из односторонних неравенств, например, для оценки сверху:

$$\frac{s_1^2}{s_2^2} \leq F_{1-p}(f_1, f_2). \quad (5.33)$$

Вероятность неравенств, противоположных (5.32) – (5.33), равна уровню значимости p ; они образуют критическую область для нулевой гипотезы. Если полученное дисперсионное отношение попадает в критическую область, то различие между дисперсиями значимо. Для удобства будем обозначать большую дисперсию через s_1^2 .

При проверке нулевой гипотезы $\sigma_1^2 = \sigma_2^2$ односторонний критерий применяется, если альтернативной гипотезой является $\sigma_1^2 > \sigma_2^2$, т. е. что большей выборочной дисперсии заведомо не может соответствовать меньшая генеральная. При этом различие между дисперсиями согласно (5.33) следует считать значимым, если

$$\frac{s_1^2}{s_2^2} > F_{1-p}(f_1, f_2). \quad (5.34)$$

Значения $F_{1-p}(f_1, f_2)$ для $p = 0,05$ приведены в приложении 5.

Двусторонний критерий значимости применяется для альтернативной гипотезы $\sigma_1^2 \neq \sigma_2^2$, т. е. когда соотношение между генеральными дисперсиями неизвестно. При этом в неравенстве (5.32) необходимо проверять только правую часть, так как левая часть всегда выполняется по условию

$$\frac{s_1^2}{s_2^2} > 1, \text{ а } \frac{1}{F_{1-p/2}(f_2, f_1)} < 1$$

при небольших p . При этом различие между дисперсиями следует считать значимым, если

$$\frac{s_1^2}{s_2^2} > F_{1-p/2}(f_1, f_2). \quad (5.35)$$

Критерий Фишера используется для сравнения дисперсий и в том случае, когда одна из дисперсий является генеральной (ее число степеней свободы считается равным ∞).

ЛЕКЦИЯ 6

Определение дисперсии по текущим измерениям. Сравнение нескольких дисперсий; критерии Бартлетта, Кохрена. Сравнение двух средних; расчет средневзвешенного значения. Проверка однородности результатов измерений. Сравнение выборочного распределения и распределения генеральной совокупности; критерии согласия Пирсона, Колмогорова.

6.1. Определение дисперсии по текущим измерениям. Сравнение нескольких дисперсий.

Математическое ожидание и дисперсия генеральной совокупности оцениваются средним и дисперсией выборки тем точнее, чем больше объем выборки. При этом среднее характеризует результат измерений, а дисперсия — точность этого результата (*дисперсия воспроизводимости*). Если проделано n параллельных опытов (опытов, проведенных при неизменном комплексе основных факторов) и получена выборка y_1, y_2, \dots, y_n значений измеряемой величины, то дисперсия воспроизводимости равна

$$s_{\text{воспр.}}^2 = \frac{\sum_{u=1}^n (y_u - \bar{y})^2}{n-1}, \text{ где } \bar{y} = \frac{\sum_{u=1}^n y_u}{n}, \quad (6.1)$$

и ошибка опыта (*ошибка воспроизводимости*)

$$s_{\text{воспр.}} = \sqrt{s_{\text{воспр.}}^2}. \quad (6.2)$$

Для оценки точности применяемой методики можно поставить специальную серию опытов, многократно повторяя измерение для одного и того же образца. Однако более надежным методом является определение ошибки воспроизводимости по текущим измерениям. Предположим, что выполняются измерения некоторой физической характеристики для k образцов, при этом для каждого образца делается различное число параллельных опытов: n_1, n_2, \dots, n_k . Частные дисперсии для каждой выборки обозначим как $s_1^2, s_2^2, \dots, s_k^2$. Числа степеней свободы частных дисперсий равны: $f_1 = n_1 - 1, f_2 = n_2 - 1, \dots, f_k = n_k - 1$. Общая дисперсия воспроизводимости всех опытов будет равна *средневзвешенному значению частных дисперсий* (в качестве весов берутся степени свободы):

$$s_{\text{воспр.}}^2 = \frac{f_1 s_1^2 + f_2 s_2^2 + \dots + f_k s_k^2}{f_1 + f_2 + \dots + f_k} =$$

$$= \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + \dots + (n_k - 1)s_k^2}{n_1 + n_2 + \dots + n_k - k}. \quad (6.3)$$

Число степеней свободы общей дисперсии равно общему числу измерений минус число связей, использованных для определения k средних:

$$f_{\text{воспр.}} = n_1 + n_2 + \dots + n_k - k = \sum_{j=1}^k n_j - k. \quad (6.4)$$

Если число опытов для каждого образца одинаково ($n_1 = n_2 = \dots = n_k = n$), то

$$s_{\text{воспр.}}^2 = \frac{(n-1)(s_1^2 + s_2^2 + \dots + s_k^2)}{nk - k} = \frac{(n-1) \sum_{j=1}^k s_j^2}{k(n-1)} = \frac{\sum_{j=1}^k s_j^2}{k}, \quad (6.5)$$

т. е. при равном числе параллельных опытов общая дисперсия воспроизводимости равна среднеарифметическому значению частных дисперсий. Число степеней свободы равно $f_{\text{воспр.}} = k(n-1)$. Число степеней свободы у общей дисперсии воспроизводимости гораздо больше, чем у каждой дисперсии в отдельности. Поэтому общая дисперсия воспроизводимости намного точнее оценивает дисперсию генеральной совокупности.

При вычислении дисперсии воспроизводимости по текущим измерениям можно объединять между собой только те результаты, которые можно рассматривать как выборки из генеральных совокупностей с равными дисперсиями.

Итак, при определении оценки дисперсии по текущим измерениям

$$s_{\text{воспр.}}^2 = \frac{f_1 s_1^2 + f_2 s_2^2 + \dots + f_k s_k^2}{f_1 + f_2 + \dots + f_k} = \frac{\sum_{j=1}^k f_j s_j^2}{f_{\text{воспр.}}} \quad (6.6)$$

принимается нулевая гипотеза равенства соответствующих генеральных дисперсий. Проверить эту гипотезу для выборок разного объема можно по *критерию Бартлета*. Бартлет показал, что в условиях нулевой гипотезы отношение V/C , где

$$B = f_{\text{воспр}} \ln s_{\text{воспр}}^2 - \sum_{j=1}^k f_j \ln s_j^2, \quad (6.7)$$

$$C = 1 + \frac{1}{3(k-1)} \left(\sum_{j=1}^k \frac{1}{f_j} - \frac{1}{f_{\text{воспр}}} \right), \quad (6.8)$$

распределено приближенно как χ^2 с $k - 1$ степенями свободы, если все $f_j > 2$. Гипотеза равенства генеральных дисперсий принимается, если при выбранном уровне значимости p

$$B/C \leq \chi_{1-p}^2. \quad (6.9)$$

Различие между выборочными дисперсиями можно считать незначимым, а сами выборочные дисперсии — однородными. Так как всегда $C > 1$, то при $B \leq \chi_{1-p}^2$ нулевую гипотезу следует принять; если же $B > \chi_{1-p}^2$, то критерий Бартлета вычисляются полностью.

Если выборочные дисперсии получены по выборкам одинаковых объемов ($n_1 = n_2 = \dots = n_k = n$), то для их сравнения используют более удобный и точный критерий Кохрена. Кохрен исследовал распределение отношения максимальной выборочной дисперсии к сумме всех дисперсий

$$G = \frac{s_{\max}^2}{\sum_{j=1}^k s_j^2} \quad (6.10)$$

Распределение случайной величины G зависит только от числа суммируемых дисперсий k и числа степеней свободы $f = n - 1$, с которым определена каждая дисперсия.

В приложении 6 приведены квантили G_{1-p} для уровня значимости $p = 0,05$. Если найденное по выборочным дисперсиям значение критерия Кохрена окажется меньше табличного

$$G < G_{1-p}(k, f), \quad (6.11)$$

то расхождение между дисперсиями следует считать случайным при выбранном уровне значимости. Если при этом определяется оценка для дисперсии воспроизводимости, то однородные дисперсии можно усреднить.

6.2. Сравнение двух средних. Расчет средневзвешенного значения.

Для сравнения между собой двух средних, полученных по выборкам из нормально распределенных генеральных совокупностей, применяется критерий Стьюдента.

Пусть заданы две случайные выборки объемами n_1 и n_2 . Первая выборка взята из нормально распределенной совокупности с параметрами m_x и σ_x^2 , вторая — из совокупности с параметрами m_y и σ_y^2 . По выборкам получены оценки для этих параметров: \bar{x} , s_x^2 и \bar{y} , s_y^2 . Требуется проверить нулевую гипотезу: $m_x = m_y$ при условии $\sigma_x^2 = \sigma_y^2 = \sigma^2$. Однородность дисперсий s_x^2 и s_y^2 проверяется по критерию Фишера. Рассмотрим случайную величину

$$z = \bar{x} - \bar{y}. \quad (6.12)$$

По свойству линейности (уравнения 2.37 – 2.38) величина z распределена нормально с параметрами

$$m_z = m_x - m_y, \quad (6.13)$$

$$\sigma_z^2 = \sigma_x^2 + \sigma_y^2 = \frac{\sigma_x^2}{n_1} + \frac{\sigma_y^2}{n_2} = \sigma^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right). \quad (6.14)$$

Составим нормированную случайную величину

$$\frac{z - m_z}{\sigma_z} = \frac{(\bar{x} - \bar{y}) - (m_x - m_y)}{\sigma \sqrt{(1/n_1 + 1/n_2)}}. \quad (6.15)$$

При замене генерального стандарта выборочным получается величина, имеющая распределение Стьюдента:

$$t = \frac{(\bar{x} - \bar{y}) - (m_x - m_y)}{s \sqrt{(1/n_1 + 1/n_2)}}, \quad (6.16)$$

с числом степеней свободы $f = n_1 + n_2 - 2$. В качестве выборочного стандарта используется ошибка опыта, равная

$$s = \sqrt{\frac{f_1 s_1^2 + f_2 s_2^2}{f_1 + f_2}} = \sqrt{\frac{(n_1 - 1) s_1^2 + (n_2 - 1) s_2^2}{n_1 + n_2 - 2}}. \quad (6.17)$$

При доверительной вероятности $\beta = 1 - p$ получаем двустороннюю оценку для разности $(m_x - m_y)$

$$\begin{aligned} \bar{x} - \bar{y} - t_{1-p/2} s \sqrt{(1/n_1 + 1/n_2)} &\leq m_x - m_y \leq \\ &\leq \bar{x} - \bar{y} + t_{1-p/2} s \sqrt{(1/n_1 + 1/n_2)} \end{aligned} \quad (6.18)$$

или односторонние оценки

$$m_x - m_y \leq \bar{x} - \bar{y} + t_{1-p} s \sqrt{(1/n_1 + 1/n_2)}, \quad (6.19)$$

$$m_x - m_y \geq \bar{x} - \bar{y} - t_{1-p} s \sqrt{(1/n_1 + 1/n_2)}. \quad (6.20)$$

В условиях нулевой гипотезы $m_x = m_y$ и неравенства (6.18) – (6.20) дают критерий проверки этой гипотезы. Нулевая гипотеза отвергается при двустороннем критерии, если

$$|\bar{x} - \bar{y}| > t_{1-p/2} s \sqrt{(1/n_1 + 1/n_2)}, \quad (6.21)$$

и при одностороннем критерии, если

$$|\bar{x} - \bar{y}| > t_{1-p} s \sqrt{(1/n_1 + 1/n_2)}. \quad (6.22)$$

В том случае если выборочные средние являются оценками одного и того же математического ожидания и выборочные дисперсии однородны, то полученные выборки можно объединить в одну серию и рассчитать для нее общие среднее и дисперсию.

Приведенными критериями нельзя пользоваться, если выборочные дисперсии неоднородны (т. е. $\sigma_x^2 \neq \sigma_y^2$). Для этого случая существует несколько приближенных критериев для сравнения двух средних. При $n_1 = n_2 = n$ можно воспользоваться приближенным t -критерием:

$$t \cong \frac{(\bar{x} - \bar{y})\sqrt{n}}{\sqrt{s_x^2 + s_y^2}} \quad (6.23)$$

с числом степеней свободы $f = \frac{n-1}{c^2 - (1-c)^2}$, где $c = \frac{s_x^2}{s_x^2 + s_y^2}$.

Если число степеней свободы дисперсии s_x^2 равно $f_1 = n_1 - 1$, дисперсии s_y^2 — $f_2 = n_2 - 1$, можно использовать другой приближенный критерий. Вычислим величину

$$T = \frac{v_1 t_{1-p/2}(f_1) + v_2 t_{1-p/2}(f_2)}{\sqrt{v_1 + v_2}}, \quad (6.24)$$

где $v_1 = s_x^2 / n_1$ и $v_2 = s_y^2 / n_2$. Нулевая гипотеза отвергается, если

$$|\bar{x} - \bar{y}| > T.$$

Сформулированный критерий является двусторонним, он превращается в односторонний при замене $p/2$ на p .

При сравнении нескольких средних можно использовать t -критерий, проводя сравнение попарно. Если выборочные средние оценивают одно и то же математическое ожидание, то в качестве единственной наилучшей оценки обычно используется *средневзвешенное значение*. Пусть независимым образом получено k оценок ($j = 1, 2, \dots, k$) некоторой величины X :

$$\bar{x}_j \pm \frac{s(x_j)}{\sqrt{n_j}} t_{1-p/2}(f_j) = \bar{x}_j \pm \Delta x_j. \quad (6.25)$$

Определим вес результата, полученного в каждой серии опытов:

$$w_j = 1 / \Delta x_j^2. \quad (6.26)$$

Тогда средневзвешенное значение (наилучшая оценка для X) равно

$$\bar{X} = \frac{\sum_{j=1}^k w_j \bar{x}_j}{\sum_{j=1}^k w_j}, \quad (6.27)$$

а его погрешность определяется формулой

$$\Delta \bar{X} = \sqrt{1 / \sum_{j=1}^k w_j}. \quad (6.28)$$

6.3. Проверка однородности результатов измерений.

Грубые измерения являются результатом поломки прибора или недосмотра экспериментатора, и результат, содержащий грубую ошибку, сильно отличается по величине. На этом основаны статистические критерии оценки и исключения грубых измерений. Наличие грубой ошибки в выборке нарушает характер распределения случайной величины, изменяет его параметры, т.е. нарушается однородность наблюдений. Следовательно, выявление грубых ошибок можно трактовать как проверку однородности наблюдений, т.е. проверку гипотезы о том, что все элементы выборки получены из одной и той же генеральной совокупности.

Пусть имеется выборка x_1, x_2, \dots, x_n значений нормально распределенной случайной величины X . Обозначим через x_{\max} (x_{\min}) наибольший (наименьший) результат измерений. Величины

$$v = \frac{x_{\max} - \bar{x}}{s \cdot \sqrt{\frac{n-1}{n}}}, \quad (6.29)$$

$$v' = \frac{\bar{x} - x_{\min}}{s \cdot \sqrt{\frac{n-1}{n}}} \quad (6.30)$$

имеют специальное распределение, зависящее только от числа степеней свободы $f = n - 2$. В приложении 7 приведены значения v (v') для $p = 0,10; 0,05; 0,025$ и $0,01$ при числе степеней свободы от 1 до 23. Величина x_{\max} (x_{\min}) исключается из выборки как грубое измерение (на уровне значимости p), если определенное по формулам (6.29) и (6.30) значение v или v' окажется больше табличного.

Если сомнение вызывают два или три элемента выборки, поступают следующим образом. Для всех сомнительных элементов вычисляют v (v'), и исследование начинается с элемента, имеющего наименьшее значение v (v'). Остальные сомнительные элементы из выборки исключаются. Для этой уменьшенной выборки вычисляют \bar{x} , s и новое значение v (v') для исследуемого элемента. Если исследуемый элемент является грубым измерением, еще с большим основанием можно считать грубыми ранее исключенные элементы. Если исследуемый элемент не является грубым измерением, его присоединяют к выборке и начинают исследовать следующий по величине v (v') элемент выборки, при этом снова вычисляя новые значения \bar{x} , s , и т.д.

6.4. Сравнение выборочного распределения и распределения генеральной совокупности. Критерии согласия Пирсона и Колмогорова.

Гипотезу о нормальности изучаемого распределения в математической статистике называют *основной гипотезой*. Проверку этой гипотезы по выборке проводят при помощи *критериев согласия*. Критерии согласия позволяют определить вероятность того, что при гипотетическом законе распределения наблюдающееся в рассматриваемой выборке отклонение вызывается случайными причинами, а не ошибкой в

гипотезе. Если эта вероятность велика, то отклонение от гипотетического закона распределения следует признать случайным и считать, что гипотеза о предполагаемом законе распределения не опровергается. Критерий согласия позволяет лишь утверждать, что гипотеза не противоречит опытным данным, если вероятность наблюдаемого отклонения от гипотетического закона велика. Чаще всего используется один из двух критериев согласия: *критерий Пирсона* (критерий χ^2) и *критерий Колмогорова*.

Критерий согласия Пирсона. Для применения критерия χ^2 весь диапазон изменения случайной величины в выборке объема n разбивается на k интервалов (от 8 до 20). Число элементов выборки, попавших в i -интервал, обозначим через n_i . Построенная по этим данным гистограмма выборочного распределения служит основанием для выбора типа закона распределения.

Параметры этого распределения могут быть найдены или из теоретических соображений, или нахождением их оценок по выборке. На основании принятого закона распределения вычисляются вероятности p_i попадания случайной величины X в i -интервал. Величина, характеризующая отклонение выборочного распределения от предполагаемого, определяется формулой

$$\chi^2 = \sum_{i=1}^k \frac{(n_i - np_i)^2}{np_i}, \quad (6.31)$$

где k — число интервалов; n — объем выборки.

Сумма (6.31) имеет приближенно χ^2 -распределение с $f = (k - c - 1)$ степенями свободы, где c — число параметров гипотетического закона распределения, определяемых по выборке. Для нормального распределения $c = 2$, если \bar{x} и s определяются по данной выборке.

Гипотеза о принятом типе закона распределения принимается на выбранном уровне значимости p , если $\chi^2 \leq \chi_{1-p}^2$, где χ_{1-p}^2 — квантиль распределения Пирсона для данного p и числа степеней свободы f (приложение 4). В противном случае делается вывод о том, что гипотеза не согласуется с выборочным распределением.

При использовании критерия χ^2 желательно, чтобы объем выборки был достаточно велик: $n \geq 50 \div 150$, а количество элементов $n_i \geq 5 \div 8$. Вероятности p_i для нормального закона распределения можно определить по формуле

$$P(a \leq X \leq b) = \Phi\left(\frac{b - \bar{x}}{s}\right) - \Phi\left(\frac{a - \bar{x}}{s}\right).$$

При подсчете теоретических вероятностей p_i считается, что крайний левый интервал простирается до $-\infty$; крайний правый — до $+\infty$.

Критерий согласия Колмогорова. Для применения этого критерия необходимо определить наибольшее абсолютное отклонение выборочной функции распределения $F_n(x)$ от генеральной $F(x)$:

$$D = \max|F_n(x) - F(x)|, \quad (6.32)$$

затем вычислить величину λ :

$$\lambda = \sqrt{n} D. \quad (6.33)$$

Квантили λ_{1-p} распределения Колмогорова приведены в приложении 8. Если $\lambda < \lambda_{1-p}$, то гипотеза о совпадении теоретического закона распределения $F(x)$ с выборочным $F_n(x)$ не отвергается. При $\lambda \geq \lambda_{1-p}$ гипотеза отклоняется (или считается сомнительной). Уровень значимости при применении критерия Колмогорова выбирают обычно равным $0.2 \div 0.3$.

Для нормального распределения $F(x)$ определяется по формуле

$$F(x) = \frac{1}{2} + \Phi\left(\frac{x - \bar{x}}{s}\right).$$

В случае выборок небольшого объема ($n < 20$) для проверки гипотезы о законе распределения можно использовать простые критерии, основанные на сравнении генеральных параметров распределения и их оценок, полученных по выборке. В качестве оцениваемых параметров удобнее всего брать моменты.

ПРИЛОЖЕНИЯ

Приложение 1

Значения функции Лапласа $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x \exp(-x^2 / 2) dx$.

x	$\Phi(x)$
0.00	0.0000
0.01	0.0040
0.02	0.0080
0.03	0.0120
0.04	0.0160
0.05	0.0199
0.06	0.0239
0.07	0.0279
0.08	0.0319
0.09	0.0359
0.10	0.0398
0.11	0.0438
0.12	0.0478
0.13	0.0517
0.14	0.0557
0.15	0.0596
0.16	0.0636
0.17	0.0675
0.18	0.0714
0.19	0.0753
0.20	0.0793
0.21	0.0832
0.22	0.0871
0.23	0.0910
0.24	0.0948
0.25	0.0987
0.26	0.1026
0.27	0.1064
0.28	0.1103

x	$\Phi(x)$
0.29	0.1141
0.30	0.1179
0.31	0.1217
0.32	0.1255
0.33	0.1293
0.34	0.1331
0.35	0.1368
0.36	0.1406
0.37	0.1443
0.38	0.1480
0.39	0.1517
0.40	0.1554
0.41	0.1591
0.42	0.1628
0.43	0.1664
0.44	0.1700
0.45	0.1736
0.46	0.1772
0.47	0.1808
0.48	0.1844
0.49	0.1879
0.50	0.1915
0.51	0.1950
0.52	0.1985
0.53	0.2019
0.54	0.2054
0.55	0.2088
0.56	0.2123
0.57	0.2157

x	$\Phi(x)$
0.58	0.2190
0.59	0.2224
0.60	0.2257
0.61	0.2291
0.62	0.2324
0.63	0.2357
0.64	0.2389
0.65	0.2422
0.66	0.2454
0.67	0.2486
0.68	0.2517
0.69	0.2549
0.70	0.2580
0.71	0.2611
0.72	0.2642
0.73	0.2673
0.74	0.2703
0.75	0.2734
0.76	0.2764
0.77	0.2794
0.78	0.2823
0.79	0.2852
0.80	0.2881
0.81	0.2910
0.82	0.2939
0.83	0.2967
0.84	0.2995
0.85	0.3023
0.86	0.3051

Продолжение приложения 1

x	$\Phi(x)$
0.87	0.3078
0.88	0.3106
0.89	0.3133
0.90	0.3159
0.91	0.3186
0.92	0.3212
0.93	0.3238
0.94	0.3264
0.95	0.3289
0.96	0.3315
0.97	0.3340
0.98	0.3365
0.99	0.3389
1.00	0.3413
1.01	0.3438
1.02	0.3461
1.03	0.3485
1.04	0.3508
1.05	0.3531
1.06	0.3554
1.07	0.3577
1.08	0.3599
1.09	0.3621
1.10	0.3643
1.11	0.3665
1.12	0.3686
1.13	0.3708
1.14	0.3729
1.15	0.3749
1.16	0.3770
1.17	0.3790
1.18	0.3810
1.19	0.3830
1.20	0.3849
1.21	0.3869

x	$\Phi(x)$
1.22	0.3883
1.23	0.3907
1.24	0.3925
1.25	0.3944
1.26	0.3962
1.27	0.3980
1.28	0.3997
1.29	0.4015
1.30	0.4032
1.31	0.4049
1.32	0.4066
1.33	0.4082
1.34	0.4099
1.35	0.4115
1.36	0.4131
1.37	0.4147
1.38	0.4162
1.39	0.4177
1.40	0.4192
1.41	0.4207
1.42	0.4222
1.43	0.4236
1.44	0.4251
1.45	0.4265
1.46	0.4279
1.47	0.4292
1.48	0.4306
1.49	0.4319
1.50	0.4332
1.51	0.4345
1.52	0.4357
1.53	0.4370
1.54	0.4382
1.55	0.4394
1.56	0.4406

x	$\Phi(x)$
1.57	0.4418
1.58	0.4429
1.59	0.4441
1.60	0.4452
1.61	0.4463
1.62	0.4474
1.63	0.4484
1.64	0.4495
1.65	0.4505
1.66	0.4515
1.67	0.4525
1.68	0.4535
1.69	0.4545
1.70	0.4554
1.71	0.4564
1.72	0.4573
1.73	0.4582
1.74	0.4591
1.75	0.4599
1.76	0.4608
1.77	0.4616
1.78	0.4625
1.79	0.4633
1.80	0.4641
1.81	0.4649
1.82	0.4656
1.83	0.4664
1.84	0.4671
1.85	0.4678
1.86	0.4686
1.87	0.4693
1.88	0.4699
1.89	0.4706
1.90	0.4713
1.91	0.4719

Продолжение приложения 1

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
1.92	0.4726	2.28	0.4887	2.72	0.4967
1.93	0.4732	2.30	0.4893	2.74	0.4969
1.94	0.4738	2.32	0.4898	2.76	0.4971
1.95	0.4744	2.34	0.4904	2.78	0.4973
1.96	0.4750	2.36	0.4909	2.80	0.4974
1.97	0.4756	2.38	0.4913	2.82	0.4976
1.98	0.4761	2.40	0.4918	2.84	0.4977
1.99	0.4767	2.42	0.4922	2.86	0.4979
2.00	0.4772	2.44	0.4927	2.88	0.4980
2.02	0.4783	2.46	0.4931	2.90	0.4981
2.04	0.4793	2.48	0.4934	2.92	0.4982
2.06	0.4803	2.50	0.4938	2.94	0.4984
2.08	0.4812	2.52	0.4941	2.96	0.49846
2.10	0.4821	2.54	0.4945	2.98	0.49856
2.12	0.4830	2.56	0.4948	3.00	0.49865
2.14	0.4838	2.58	0.4951	3.20	0.49931
2.16	0.4846	2.60	0.4953	3.40	0.49966
2.18	0.4854	2.62	0.4956	3.60	0.49984
2.20	0.4861	2.64	0.4959	3.80	0.499928
2.22	0.4868	2.66	0.4961	4.00	0.499968
2.24	0.4875	2.68	0.4963	5.00	0.499997
2.26	0.4881	2.70	0.4965		

Приложение 2

Квантили нормального распределения.

p	$1-p/2$	$U_{1-p/2}$	p	$1-p/2$	$U_{1-p/2}$
0.80	0.60	0.25	0.05	0.975	1.96
0.50	0.75	0.67	0.04	0.980	2.05
0.40	0.80	0.84	0.02	0.990	2.33
0.30	0.85	1.04	0.01	0.995	2.58
0.25	0.875	1.15	0.005	0.9975	2.81
0.20	0.90	1.28	0.002	0.999	3.09
0.15	0.925	1.44	0.001	0.9995	3.29
0.10	0.95	1.64	0.0001	0.99995	3.89

Приложение 3

Квантили распределения Стьюдента $t_{1-p/2} (t_p(f))$.

Число степеней свободы f	Уровни значимости p						
	0.20	0.10	0.05	0.02	0.01	0.005	0.001
1	3.08	6.31	12.71	31.82	63.66	127.32	636.32
2	1.89	2.92	4.30	6.97	9.93	14.09	31.60
3	1.64	2.35	3.18	4.54	5.84	7.45	12.94
4	1.53	2.13	2.78	3.75	4.60	5.60	8.61
5	1.48	2.02	2.57	3.37	4.03	4.77	6.86
6	1.44	1.94	2.45	3.14	3.71	4.32	5.96
7	1.42	1.90	2.37	3.00	3.50	4.03	5.41
8	1.40	1.86	2.31	2.90	3.36	3.83	5.04
9	1.38	1.83	2.26	2.82	3.25	3.69	4.78
10	1.37	1.81	2.23	2.76	3.17	3.58	4.59
11	1.36	1.80	2.20	2.72	3.11	3.50	4.44
12	1.36	1.78	2.18	2.68	3.06	3.43	4.32
13	1.35	1.77	2.16	2.65	3.01	3.37	4.22
14	1.34	1.76	2.15	2.62	2.98	3.33	4.14
15	1.34	1.75	2.13	2.60	2.95	3.29	4.07
16	1.34	1.75	2.12	2.58	2.92	3.25	4.02
17	1.33	1.74	2.11	2.57	2.90	3.22	3.97
18	1.33	1.73	2.10	2.55	2.88	3.20	3.92
19	1.33	1.73	2.09	2.54	2.86	3.17	3.88
20	1.33	1.73	2.09	2.53	2.85	3.15	3.85
22	1.32	1.72	2.07	2.51	2.82	3.12	3.79
24	1.32	1.71	2.06	2.49	2.80	3.09	3.75
26	1.32	1.71	2.06	2.48	2.78	3.07	3.71
28	1.31	1.70	2.05	2.47	2.76	3.05	3.67
30	1.31	1.70	2.04	2.46	2.75	3.03	3.65
40	1.30	1.68	2.02	2.42	2.70	2.97	3.55
60	1.30	1.67	2.00	2.39	2.66	2.91	3.46
120	1.29	1.66	1.98	2.36	2.62	2.86	3.37
∞	1.28	1.64	1.96	2.33	2.58	2.81	3.29

Приложение 4

Квантили распределения Пирсона χ^2_{1-p} .

Число степеней свободы f	Уровни значимости p							
	0.99	0.98	0.95	0.90	0.80	0.70	0.50	0.30
1	0.00016	0.0006	0.0039	0.016	0.064	0.148	0.455	1.07
2	0.020	0.040	0.103	0.211	0.446	0.713	1.386	2.41
3	0.115	0.185	0.352	0.584	1.005	1.424	2.336	3.66
4	0.30	0.43	0.71	1.06	1.65	2.19	3.36	4.9
5	0.55	0.75	1.14	1.61	2.34	3.00	4.35	6.1
6	0.87	1.13	1.63	2.2	3.07	3.83	5.35	7.2
7	1.24	1.56	2.17	2.83	3.82	4.67	6.35	8.4
8	1.65	2.03	2.73	3.49	4.59	5.53	7.34	9.5
9	2.09	2.53	3.32	4.17	5.38	6.39	8.34	10.7
10	2.56	3.06	3.94	4.86	6.18	7.27	9.34	11.8
11	3.1	3.6	4.6	5.6	7.0	8.1	10.3	12.9
12	3.6	4.2	5.2	6.3	7.8	9.0	11.3	14.0
13	4.1	4.8	5.9	7.0	8.6	9.9	12.3	15.1
14	4.7	5.4	6.6	7.8	9.5	10.8	13.3	16.2
15	5.2	6.0	7.3	8.5	10.3	11.7	14.3	17.3
16	5.8	6.6	8.0	9.3	11.2	12.6	15.3	18.4
17	6.4	7.3	8.7	10.1	12.0	13.5	16.3	19.5
18	7.0	7.9	9.4	10.9	12.9	14.4	17.3	20.6
19	7.6	8.6	10.1	11.7	13.7	15.4	18.3	21.7
20	8.3	9.2	10.9	12.4	14.6	16.3	19.3	22.8
21	8.9	9.9	11.6	13.2	15.4	17.2	20.3	23.9
22	9.5	10.6	12.3	14.0	16.3	18.1	21.3	24.9
23	10.2	11.3	13.1	14.8	17.2	19.0	22.3	26.0
24	10.9	12.0	13.8	15.7	18.1	19.9	23.3	27.1
25	11.5	12.7	14.6	16.5	18.9	20.9	24.3	28.2
26	12.2	13.4	15.4	17.3	19.8	21.8	25.3	29.3
27	12.9	14.1	16.2	18.1	20.7	22.7	26.3	30.3
28	13.6	14.8	16.9	18.9	21.6	23.6	27.3	31.4
29	14.3	15.6	17.7	19.8	22.4	24.6	28.3	32.5
30	15.0	16.3	18.5	20.6	23.4	25.5	29.3	33.5

Число степеней свободы f	Уровни значимости p							
	0.20	0.10	0.05	0.02	0.01	0.005	0.002	0.001
1	1.64	2.7	3.8	5.4	6.6	7.9	9.5	10.8
2	3.22	4.6	6	7.8	9.2	10.6	12.4	13.8
3	4.64	6.3	7.8	9.8	11.3	12.8	14.8	16.3
4	6.0	7.8	9.5	11.7	13.3	14.9	16.9	18.5
5	7.3	9.2	11.1	13.4	15.1	16.3	18.9	20.5
6	8.6	10.6	12.6	15.0	16.8	18.6	20.7	22.5
7	9.8	12.0	14.1	16.6	18.5	20.3	22.6	24.3
8	11.0	13.4	15.5	18.2	20.1	21.9	24.3	26.1
9	12.2	14.7	16.9	19.7	21.7	23.6	26.1	27.9
10	13.4	16.0	18.3	21.2	23.2	25.2	27.7	29.6
11	14.6	17.3	19.7	22.6	24.7	26.8	29.4	31.3
12	15.8	18.5	21.0	24.1	26.2	28.3	31	32.9
13	17.0	19.8	22.4	25.5	27.7	29.8	32.5	34.5
14	18.2	21.1	23.7	26.9	29.1	31.3	34	36.1
15	19.3	22.3	25.0	28.3	30.6	32.8	35.5	37.7
16	20.5	23.5	26.3	29.6	32.0	34.3	37	39.2
17	21.6	24.8	27.6	31.0	33.4	35.7	38.5	40.8
18	22.8	26.0	28.9	32.3	34.8	37.2	40	42.3
19	23.9	27.2	30.1	33.7	36.2	38.6	41.5	43.8
20	25.0	28.4	31.4	35.0	37.6	40.0	43	45.3
21	26.2	29.6	32.7	36.3	38.9	41.4	44.5	46.8
22	27.3	30.8	33.9	37.7	40.3	42.8	46	48.3
23	28.4	32.0	35.2	39.0	41.6	44.2	47.5	49.7
24	29.6	33.2	36.4	40.3	43.0	45.6	48.5	51.2
25	30.7	34.4	37.7	41.6	44.3	46.9	50	52.6
26	31.8	35.6	38.9	42.9	45.6	48.3	61.5	54.1
27	32.9	36.7	40.1	44.1	47.0	49.6	53	55.5
28	34.0	37.9	41.3	45.4	48.3	51.0	54.5	56.9
29	35.1	39.1	42.6	46.7	49.6	52.3	56	58.3
30	36.3	40.3	43.8	48.0	50.9	53.7	57.5	59.7

Приложение 5

Квантили распределения Фишера F_{1-p} для $p = 0.05$.

f_2	f_1								
	1	2	3	4	5	6	12	24	∞
1	164.4	199.5	215.7	224.6	230.2	234.0	244.9	249	254.3
2	18.5	19.2	19.2	19.3	19.3	19.3	19.4	19.5	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	5.6	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	1.8
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	2.4	2.1	1.9	1.7
28	4.2	3.3	2.9	2.7	2.6	2.4	2.1	1.9	1.6
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

Приложение 6

Квантили распределения Кохрена G_{1-p}^* для $p = 0.05$.

k	f										
	1	2	3	4	5	6	8	10	16	36	∞
2	9985	9750	9392	9057	8772	8534	8159	7880	7341	6602	5000
3	9669	8709	7977	7454	7071	6771	6333	6025	5466	4748	3333
4	9065	7679	6841	6287	5895	5598	5175	4884	4366	3720	2500
5	8412	6838	5981	5441	5065	4783	4387	4118	3645	3066	2000
6	7808	6161	5321	4803	4447	4184	3817	3568	3135	2612	1667
7	7271	5612	4800	4307	3974	3726	3384	3154	2756	2278	1429
8	6798	5157	4377	3910	3595	3362	3043	2829	2462	2022	1250
9	6385	4775	4027	3584	3286	3067	2768	2568	2226	1820	1111
10	6020	4450	3733	3311	3029	2823	2541	2353	2032	1655	1000
12	5410	3924	3264	2880	2624	2439	2187	2020	1737	1403	0833
15	4709	3346	2758	2419	2195	2034	1815	1671	1429	1144	0667
20	3894	2705	2205	1921	1735	1602	1422	1303	1108	0879	0500
24	3434	2354	1907	1656	1493	1374	1216	1113	0942	0743	0417
30	2929	1980	1593	1377	1237	1137	1001	0921	0771	0604	0333
40	2370	1576	1259	1082	0968	0887	0780	0713	0595	0462	0250
60	1737	1131	0895	0765	0682	0623	0552	0497	0411	0316	0167
120	0998	0632	0495	0419	0371	0337	0292	0266	0218	0165	0083
∞	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000

* Все квантили G_{1-p} меньше единицы, поэтому в таблице приведены лишь десятичные знаки, следующие после запятой, перед которой при пользовании таблицей нужно ставить ноль целых. Например, при $n = 6, f = 3$ имеем $G_{0.95} = 0.5321$.

Приложение 7

Значения ν (ν') для различных уровней значимости.

Число степеней свободы f	Уровни значимости p			Число степеней свободы f	Уровни значимости p		
	0.10	0.05	0.01		0.10	0.05	0.01
1	1.406	1.412	1.414	12	2.297	2.461	2.759
2	1.645	1.689	1.723	13	2.326	2.493	2.800
3	1.791	1.869	1.955	14	2.354	2.523	2.837
4	1.894	1.996	2.130	15	2.380	2.551	2.871
5	1.974	2.093	2.265	16	2.404	2.577	2.903
6	2.041	2.172	2.374	17	2.426	2.600	2.932
7	2.097	2.237	2.464	18	2.447	2.623	2.959
8	2.146	2.294	2.540	19	2.467	2.644	2.984
9	2.190	2.343	2.606	20	2.486	2.664	3.008
10	2.229	2.387	2.663	21	2.504	2.683	3.030
11	2.264	2.426	2.714	22	2.520	2.701	3.051

Приложение 8

Квантили распределения Колмогорова.

p	λ_{1-p}	p	λ_{1-p}	p	λ_{1-p}
0.99	0.44	0.50	0.83	0.15	1.14
0.90	0.57	0.40	0.89	0.10	1.22
0.80	0.64	0.30	0.97	0.05	1.36
0.70	0.71	0.25	1.02	0.02	1.52
0.60	0.77	0.20	1.07	0.01	1.63

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	5
ЛЕКЦИЯ 1	7
1.1. Случайные величины. Классификация ошибок измерений. Абсолютная и относительная погрешность.	7
1.2. Оценка погрешностей функций приближенных аргументов.	10
1.3. Распределение случайных величин. Функция распределения и плотность распределения случайной величины. ..	13
ЛЕКЦИЯ 2	16
2.1. Числовые характеристики случайной величины. Свойства математического ожидания и дисперсии. Нормированная случайная величина.	16
2.2. Нормальное и стандартное распределения случайной величины. Функция Лапласа. Задача об абсолютном отклонении.	20
ЛЕКЦИЯ 3	24
3.1. Генеральная совокупность и случайная выборка. Выборочная функция распределения. Гистограммы. Понятие об оценках параметров генерального распределения.	24
3.2. Метод максимального правдоподобия.	27
3.3. Оценка математического ожидания и дисперсии нормально распределенной случайной величины. Дисперсия среднего серии измерений.	29
ЛЕКЦИЯ 4	33
4.1. Доверительные интервалы и доверительная вероятность, уровень значимости.	33
4.2. Проверка статистических гипотез, критерии значимости, ошибки первого и второго рода.	35
4.3. Построение доверительного интервала для математического ожидания непосредственно измеряемой величины. Распределение Стьюдента.	38
ЛЕКЦИЯ 5	41
5.1. Оценка случайной и суммарной ошибки косвенных измерений.	41

5.2. Оценка дисперсии нормально распределенной случайной величины.	44
5.3. Сравнение двух дисперсий. Распределение Фишера.	46
ЛЕКЦИЯ 6	49
6.1. Определение дисперсии по текущим измерениям. Сравнение нескольких дисперсий.	49
6.2. Сравнение двух средних. Расчет средневзвешенного значения.	52
6.3. Проверка однородности результатов измерений.	54
6.4. Сравнение выборочного распределения и распределения генеральной совокупности. Критерии согласия Пирсона и Колмогорова.	55
ПРИЛОЖЕНИЯ	58
Приложение 1	58
Приложение 2	60
Приложение 3	61
Приложение 4	62
Приложение 5	64
Приложение 6	65
Приложение 7	66
Приложение 8	66

БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ХИМИЧЕСКИЙ ФАКУЛЬТЕТ
Кафедра физической химии

А. В. Блохин

ТЕОРИЯ ЭКСПЕРИМЕНТА

Курс лекций

В двух частях

Часть 2

МИНСК
2002

Автор: Блохин А.В., кандидат химических наук.

Рецензенты:

кандидат химических наук **Н.Н. Горошко**;
старший преподаватель кафедры физической химии
Л.М. Володкович.

Печатается по решению
Редакционно-издательского совета
Белорусского государственного университета

Блохин А.В.

Б70 Теория эксперимента: Курс лекций. В 2 ч. Ч. 2.

А.В. Блохин. – Мн.: БГУ, 2002. – ... с.

ISBN

Аннотация:

Учебное пособие посвящено статистическим методам оптимизации экспериментальных исследований в физической химии и содержит основы методов регрессионного, корреляционного и дисперсионного анализов и планирования экстремального эксперимента.

ЛЕКЦИЯ 7

Системы случайных величин. Функция и плотность распределения системы двух случайных величин. Условные законы распределения. Стохастическая связь. Ковариация. Коэффициент корреляции, его свойства. Линии регрессии. Выборочный коэффициент корреляции; проверка гипотезы об отсутствии корреляции. Приближенная регрессия; метод наименьших квадратов.

7.1. Системы случайных величин. Функция и плотность распределения системы двух случайных величин. Условные законы распределения

На практике чаще всего приходится иметь дело с экспериментами, результатом которых является не одна случайная величина, а две и более, образующие систему. Свойства системы случайных величин не ограничиваются свойствами величин, в нее входящих; они определяются также взаимосвязью (зависимостями) этих случайных величин. Информация о каждой случайной величине, входящей в систему, содержится в ее законе распределения.

Рассмотрим систему из двух случайных величин X и Y . Функцией распределения такой системы называется вероятность совместного выполнения двух неравенств

$$F(x, y) = P(X < x, Y < y). \quad (7.1)$$

Плотность распределения системы $f(x, y)$ определяется как вторая смешанная производная $F(x, y)$

$$f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y}. \quad (7.2)$$

Вероятность попадания точки (X, Y) в произвольную область D равна

$$P[(X, Y) \in D] = \iint_{(D)} f(x, y) dx dy. \quad (7.3)$$

Свойства плотности распределения:

1) она является неубывающей функцией:

$$f(x, y) \geq 0; \quad (7.4)$$

- 2) вероятность попадания случайной точки на всю координатную плоскость равна вероятности достоверного события:

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy = 1; \quad (7.5)$$

- 3) функция распределения выражается через плотность распределения как

$$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(x, y) dx dy; \quad (7.6)$$

- 4) плотность распределения каждой из случайных величин можно получить следующим образом:

$$F_1(x) = F(x, \infty) = \int_{-\infty}^x \int_{-\infty}^{+\infty} f(x, y) dx dy, \quad (7.7)$$

$$f_1(x) = \frac{dF_1(x)}{dx} = \int_{-\infty}^{+\infty} f(x, y) dy, \quad (7.8)$$

$$f_2(y) = \frac{dF_2(y)}{dy} = \int_{-\infty}^{+\infty} f(x, y) dx. \quad (7.9)$$

Чтобы полностью охарактеризовать систему (т. е. получить ее закон распределения), кроме распределения каждой величины, входящей в систему, необходимо знать и связь между этими величинами. Эта зависимость характеризуется с помощью *условных законов распределения*.

Условным законом распределения величины Y , входящей в систему (X, Y) , называется ее закон распределения при условии, что другая случайная величина X приняла определенное значение x . Условная функция распределения обозначается $F(y/x)$, плотность распределения — $f(y/x)$. Для условных плотностей распределений справедлива *теорема умножения законов распределения*:

$$f(x, y) = f_1(x) f(y/x), \quad (7.10)$$

$$f(x, y) = f_2(y) f(x/y). \quad (7.11)$$

Тогда

$$f(y/x) = \frac{f(x,y)}{f_1(x)} = \frac{f(x,y)}{\int_{-\infty}^{\infty} f(x,y) dy}, \quad (7.12)$$

$$f(x/y) = \frac{f(x,y)}{f_2(y)} = \frac{f(x,y)}{\int_{-\infty}^{\infty} f(x,y) dx}. \quad (7.13)$$

7.2. Стохастическая связь. Ковариация. Коэффициент корреляции. Регрессия

Стохастической связью между случайными величинами называется такая связь, при которой с изменением одной величины меняется распределение другой. *Функциональной зависимостью* называется такая связь между случайными величинами, при которой при известном значении одной из величин можно точно указать значение другой.

В отличие от функциональной связи при стохастической связи с изменением величины X величина Y имеет лишь тенденцию изменяться. По мере увеличения тесноты стохастической зависимости она все более приближается к функциональной, а в пределе ей соответствует. Крайняя противоположность функциональной связи — полная независимость случайных величин.

Если случайные величины независимы, то согласно теореме умножения (7.10–7.11) получаем

$$f(y/x) = f_2(y) \text{ и } f(x/y) = f_1(x), \quad (7.14)$$

$$f(x,y) = f_1(x)f_2(y). \quad (7.15)$$

Условие (7.15) можно использовать в качестве необходимого и достаточного критерия независимости двух случайных величин, если известны плотности распределения системы и случайных величин, в нее входящих.

При неизвестном законе распределения системы для оценки тесноты стохастической связи чаще всего используется *коэффициент корреляции*. Дисперсия суммы двух случайных величин X и Y равна

$$D\{X + Y\} = M\{[X + Y - M(X + Y)]^2\} = M\{[X - M(X) + Y - M(Y)]^2\} =$$

$$\begin{aligned}
&= M[X - M(X)]^2 + 2M\{[X - M(X)][Y - M(Y)]\} + M[Y - M(Y)]^2 = \\
&= D(X) + 2M\{[X - M(X)][Y - M(Y)]\} + D(Y). \quad (7.16)
\end{aligned}$$

Если X и Y независимы, то

$$D(X + Y) = D(X) + D(Y).$$

Тогда зависимость между X и Y существует, если

$$M([X - m_x][Y - m_y]) \neq 0. \quad (7.17)$$

Величина (7.17) называется *корреляционным моментом*, или *ковариацией* $\text{cov}\{XY\}$, (cov_{xy}) случайных величин. Она характеризует не только зависимость величин, но и их рассеяние.

Из (7.17) следует, что если одна из величин мало отклоняется от своего математического ожидания, то ковариация будет мала даже при тесной стохастической связи. Чтобы избежать этого, для характеристики связи используют безразмерную величину, называемую *коэффициентом корреляции*:

$$r_{xy} = \frac{\text{cov}_{xy}}{\sigma_x \sigma_y} = \frac{M([X - m_x][Y - m_y])}{\sigma_x \sigma_y}, \quad (7.18)$$

где σ_x и σ_y — стандартные отклонения X и Y .

Случайные величины, для которых ковариация (значит, и коэффициент корреляции) равна нулю, называются *некоррелированными*. Равенство нулю коэффициента корреляции не всегда означает, что случайные величины X и Y независимы: связь может проявляться в моментах более высокого порядка (по сравнению с математическим ожиданием). Только в случае нормального распределения при $r_{xy} = 0$ связь между случайными величинами однозначно отсутствует.

Плотность нормального распределения системы двух случайных величин выражается следующей формулой:

$$\begin{aligned}
f(x, y) &= \frac{1}{2\pi \sigma_x \sigma_y \sqrt{1 - r^2}} \times, \\
&\times \exp \left\{ -\frac{1}{2(1 - r^2)} \left[\frac{(x - m_x)^2}{\sigma_x^2} - \frac{2r(x - m_x)(y - m_y)}{\sigma_x \sigma_y} + \frac{(y - m_y)^2}{\sigma_y^2} \right] \right\}, \quad (7.19)
\end{aligned}$$

где r — коэффициент корреляции. Если X и Y некоррелированы (т. е. $r = 0$), то из (7.19) следует, что

$$\begin{aligned}
f(x, y) &= \frac{1}{2\pi\sigma_x\sigma_y} \exp\left\{-\frac{1}{2}\left[\frac{(x-m_x)^2}{\sigma_x^2} + \frac{(y-m_y)^2}{\sigma_y^2}\right]\right\} = \\
&= \frac{1}{\sqrt{2\pi}\sigma_x} \exp\left[-\frac{(x-m_x)^2}{2\sigma_x^2}\right] \frac{1}{\sqrt{2\pi}\sigma_y} \exp\left[-\frac{(y-m_y)^2}{2\sigma_y^2}\right] = \\
&= f_1(x)f_2(y), \tag{7.20}
\end{aligned}$$

т. е. нормально распределенные случайные величины X и Y не только некоррелированы, но и независимы.

Отметим следующие свойства коэффициента корреляции:

- 1) величина r_{xy} не меняется от прибавления к X и Y неслучайных слагаемых;
- 2) величина r_{xy} не меняется от умножения X и Y на положительные числа;
- 3) если одну из величин, не меняя другой, умножить на -1 , то на -1 умножится и коэффициент корреляции.

Тогда, если от исходных величин перейти к нормированным

$$X_0 = \frac{X - m_x}{\sigma_x}, \quad Y_0 = \frac{Y - m_y}{\sigma_y},$$

величина r_{xy} не изменится: $r_{x_0, y_0} = r_{xy}$. Из (7.16) и (7.18) следует, что

$$\sigma^2(X + Y) = \sigma^2(X) + \sigma^2(Y) + 2r_{xy} \sqrt{\sigma^2(X)\sigma^2(Y)}. \tag{7.21}$$

Для нормированных величин $\sigma^2(X_0) = \sigma^2(Y_0) = 1$, тогда

$$\sigma^2(X_0 + Y_0) = 2 + 2r_{xy}. \tag{7.22}$$

Аналогично в случае разности $(X - Y)$ можно получить, что

$$\sigma^2(X_0 - Y_0) = 2 - 2r_{xy}. \tag{7.23}$$

По определению дисперсии

$$\sigma^2(X_0 + Y_0) \geq 0 \text{ и } \sigma^2(X_0 - Y_0) \geq 0,$$

следовательно

$$\begin{aligned}
2 + 2r_{xy} &\geq 0, \quad 2 - 2r_{xy} \geq 0, \\
r_{xy} &\geq -1, \quad r_{xy} \leq 1, \\
-1 &\leq r_{xy} \leq 1. \tag{7.24}
\end{aligned}$$

При $r_{xy} = \pm 1$ имеем линейные функциональные зависимости вида

$$y = b_0 + b_1 x,$$

при этом если $r_{xy} = 1$, то $b_1 > 0$; если $r_{xy} = -1$, то $b_1 < 0$.

Если между величинами X и Y имеется произвольная стохастическая связь, то $-1 < r_{xy} < 1$. При $r_{xy} > 0$ говорят о *положительной корреляционной связи* между X и Y , при $r_{xy} < 0$ — об *отрицательной*. Следует учитывать, что коэффициент корреляции характеризует не любую зависимость, а только линейную.

Для нормально распределенной системы двух случайных величин можно доказать, что

$$\begin{aligned} f(y/x) &= \frac{f(x, y)}{f_1(x)} = \\ &= \frac{1}{\sigma_y \sqrt{1-r^2} \sqrt{2\pi}} \exp \left[-\frac{1}{2(1-r^2)} \left(\frac{y-m_y}{\sigma_y} - r \frac{x-m_x}{\sigma_x} \right)^2 \right] = \\ &= \frac{1}{\sigma_y \sqrt{1-r^2} \sqrt{2\pi}} \exp \left[-\frac{1}{2(1-r^2) \sigma_y^2} \left(y - m_y - r \frac{\sigma_y}{\sigma_x} (x - m_x) \right)^2 \right]. \end{aligned} \quad (7.25)$$

Условная плотность распределения величины Y соответствует плотности нормального распределения с математическим ожиданием

$$m_{y/x} = m_y + r \frac{\sigma_y}{\sigma_x} (x - m_x) \quad (7.26)$$

и среднеквадратичным отклонением

$$\sigma_{y/x} = \sigma_y \sqrt{1-r^2}. \quad (7.27)$$

Величина $m_{y/x}$ называется *условным математическим ожиданием* величины Y при данном X . Линейная зависимость (7.26) — *регрессией* Y на X . По аналогии прямая

$$m_{x/y} = m_x + r \frac{\sigma_x}{\sigma_y} (y - m_y) \quad (7.28)$$

есть регрессия X на Y .

Линии регрессии совпадают только при наличии линейной функциональной зависимости. Из (7.26) и (7.28) видно, что для независимых X и Y линии регрессии параллельны координатным осям.

7.3. Выборочный коэффициент корреляции. Проверка гипотезы об отсутствии корреляции

При обработке результатов большинства физико-химических измерений возникает задача описания зависимости между исследуемыми случайными величинами. Для экспериментального изучения зависимости между двумя случайными величинами X и Y проводят n независимых опытов, при этом в каждом из них получают пару значений (x_i, y_i) , $i = 1, 2, \dots, n$. О наличии или отсутствии корреляции между X и Y можно качественно судить по виду поля корреляции, нанеся точки (x_i, y_i) на координатную плоскость.

Для количественной оценки тесноты связи служит *выборочный коэффициент корреляции*. Как было установлено ранее, состоятельными и несмещенными оценками для математических ожиданий m_x и m_y служат выборочные средние \bar{x} и \bar{y} , а генеральных дисперсий σ_x^2 и σ_y^2 — выборочные дисперсии s_x^2 и s_y^2 . Можно доказать, что состоятельной и несмещенной оценкой генеральной ковариации cov_{xy} служит *выборочная ковариация*

$$\text{cov}_{xy}^* = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}). \quad (7.29)$$

Пользуясь этой оценкой, рассчитывают выборочный коэффициент корреляции

$$r_{xy}^* = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{(n-1) s_x s_y}, \quad (7.30)$$

который является состоятельной оценкой коэффициента корреляции генеральной совокупности со смещением, равным $r(1-r^2)/2n$. Величина смещения убывает с увеличением числа опытов и при $n > 50$ составляет менее 1 %. Выборочный коэффициент корреляции обладает теми же свойствами, что и r_{xy} , и по абсолютной величине также не больше единицы:

$$-1 \leq r_{xy}^* \leq 1. \quad (7.31)$$

Величина выборочного коэффициента корреляции определяет меру криволинейности связи между X и Y . Поэтому возможны случаи,

когда при коэффициенте корреляции, значительно меньшем единицы, связь между X и Y оказывается близкой к функциональной, хотя и существенно нелинейной.

В случае, если полученное значение r^* близко к нулю, необходимо провести проверку гипотезы об отсутствии корреляции между случайными величинами. Требуется определить, значимо ли отличается r^* от нуля. Если число опытов n достаточно велико (более 20), то в условиях нулевой гипотезы ($H_0: r = 0$) можно использовать нормальное распределение со стандартом

$$\sigma_{r^*} \approx (1 - r^{*2}) / \sqrt{n}. \quad (7.32)$$

Тогда при $\beta = 0,95$ генеральный коэффициент корреляции находится в следующих доверительных границах:

$$r^* - \frac{1.96 \cdot (1 - r^{*2})}{\sqrt{n}} \leq r \leq r^* + \frac{1.96 \cdot (1 - r^{*2})}{\sqrt{n}}. \quad (7.33)$$

С вероятностью 0,95 можно ожидать, что существует корреляция между случайными величинами, если 0 не содержится внутри доверительного интервала.

На практике, особенно при числе опытов $n < 20$, часто приходится решать вопрос о том, насколько хорошо полученные экспериментальные точки подтверждают линейную связь между величинами X и Y . Ответить на этот вопрос можно следующим образом. Предположим, что две переменные X и Y действительно некоррелированы, т. е. при проведении бесконечно большого числа измерений выборочный коэффициент корреляции для них был бы равен нулю. При конечном числе измерений, однако, маловероятно, чтобы величина r^* была точно равна нулю из-за воздействия случайных факторов.

Обозначим через

$$P_n (|r^*| \geq r_1^*)$$

вероятность того, что n измерений двух некоррелированных переменных X и Y приведут к значению r^* (по модулю), не меньшему некоторого частного значения r_1^* . Результаты расчетов вероятностей P_n для выборок различного объема n и чисел r_1^* представлены в табл. 1. Для ответа на вопрос о том, насколько хорошо n пар полученных значений (x_i, y_i) подтверждают линейную связь между исследуемыми величинами, вначале по измеренным точкам вычисляют выборочный коэффициент корреляции r_1^* . Далее по табл. 1 находят вероятность P_n того, что n некоррелированных точек приведут к значению коэффициента

Таблица 1

Вероятность P_n того, что n измерений двух некоррелированных переменных дадут коэффициент корреляции $|r^*| \geq r_1^*$ (прочерками отмечены значения, меньшие 0,01)

n	r_1^*								
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
3	0.94	0.87	0.81	0.74	0.67	0.59	0.51	0.41	0.29
4	0.90	0.80	0.70	0.60	0.50	0.40	0.30	0.20	0.10
5	0.87	0.75	0.62	0.50	0.39	0.28	0.19	0.10	0.04
6	0.85	0.70	0.56	0.43	0.31	0.21	0.12	0.06	0.01
7	0.83	0.67	0.51	0.37	0.25	0.15	0.08	0.03	—
8	0.81	0.63	0.47	0.33	0.21	0.12	0.05	0.02	—
9	0.80	0.61	0.43	0.29	0.17	0.09	0.04	0.01	—
10	0.78	0.58	0.40	0.25	0.14	0.07	0.02	0.01	—
11	0.77	0.56	0.37	0.22	0.12	0.05	0.02	—	—
12	0.76	0.53	0.34	0.20	0.10	0.04	0.01	—	—
13	0.75	0.51	0.32	0.18	0.08	0.03	0.01	—	—
14	0.73	0.49	0.30	0.16	0.07	0.02	0.01	—	—
15	0.72	0.47	0.28	0.14	0.06	0.02	—	—	—
16	0.71	0.46	0.26	0.12	0.05	0.01	—	—	—
17	0.70	0.44	0.21	0.11	0.04	0.01	—	—	—
18	0.69	0.43	0.23	0.10	0.04	0.01	—	—	—
19	0.68	0.41	0.21	0.09	0.03	0.01	—	—	—
20	0.67	0.40	0.20	0.08	0.03	0.01	—	—	—
25	0.63	0.34	0.15	0.05	0.01	—	—	—	—
30	0.60	0.29	0.11	0.03	0.01	—	—	—	—
35	0.57	0.25	0.08	0.02	—	—	—	—	—
40	0.54	0.22	0.06	0.01	—	—	—	—	—
50	0.49	0.16	0.03	—	—	—	—	—	—
60	0.45	0.13	0.02	—	—	—	—	—	—
80	0.38	0.08	0.01	—	—	—	—	—	—
100	0.32	0.05	—	—	—	—	—	—	—

корреляции, не меньшего r_1^* . Если $P_n \leq 0,05$ (для «высокосignимых» корреляций $P_n \leq 0,01$), то гипотеза о линейной зависимости между величинами X и Y принимается (при выбранном уровне значимости 0,05 или 0,01 соответственно).

Например, по выборке из 5 пар значений (x_i, y_i) получено $r_1^* = 0,9$. Вероятность получения коэффициента r^* такого, что $|r^*| \geq 0,9$, для 5 некоррелированных точек равна $P_n = 0,04$ (табл. 1). Следовательно, гипотеза о линейной связи двух исследуемых величин может быть принята с уровнем значимости 0,05.

7.4. Приближенная регрессия. Метод наименьших квадратов

При исследовании корреляционной зависимости между двумя случайными величинами необходимо по данной выборке объемом n найти уравнение приближенной регрессии, чаще всего в виде следующего полинома:

$$\hat{y}(x) = b_0 + b_1x + b_2x^2 + b_3x^3 + \dots = b_0 + \sum_{j=1}^k b_jx^j, \quad (7.34)$$

где коэффициенты b_0 и b_j являются оценками соответствующих теоретических коэффициентов истинного уравнения регрессии

$$m_{y/x} = \varphi(x) = \beta_0 + \beta_1x + \beta_2x^2 + \beta_3x^3 + \dots = \beta_0 + \sum_{j=1}^k \beta_jx^j, \quad (7.35)$$

и оценить допускаемую при этом ошибку. Для этого обычно используют метод наименьших квадратов.

Рассмотрим некоторый класс функций, аналитическое выражение которых содержит некоторое число неопределенных коэффициентов, равное k . Наилучшее уравнение приближенной регрессии дает та функция из рассматриваемого класса, для которой сумма квадратов S имеет наименьшее значение:

$$S = \sum_{i=1}^n \left[y_i - \hat{y}(x_i) \right]^2 = \min. \quad (7.36)$$

Предположим, что экспериментальные точки отклоняются от уравнения истинной регрессии $\varphi(x)$ только в результате воздействия случайных факторов, а ошибки измерения нормально распределены. Полученные в опытах значения y_i будут распределены по нормальному закону с математическим ожиданием $m(y_i) = \varphi(x_i)$ и дисперсией σ_i^2 . При равноточных экспериментах $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 = \sigma^2$. Тогда плотность распределения величины Y_i принимает вид

$$f_i(y_i) = \frac{1}{\sqrt{2\pi} \sigma} \exp \left\{ -\frac{1}{2\sigma^2} [y_i - \varphi(x_i)]^2 \right\}. \quad (7.37)$$

В результате опытов случайные величины Y_i приняли совокупность значений y_i . Используем принцип максимального правдоподобия

бия: определим так математические ожидания $\varphi(x_i)$, чтобы вероятность этого события была максимальной. Обозначим через $p_i = f_i(y_i) \delta$ вероятность того, что случайная величина Y_i примет значение из интервала $y_i - \delta/2, y_i + \delta/2$. Вероятность совместного осуществления подобных событий для $i = 1, 2, \dots, n$ равна

$$\begin{aligned} P &= \delta^n \prod_{i=1}^n f_i(y_i) = \delta^n \sigma^{-n} (2\pi)^{-n/2} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n [y_i - \varphi(x_i)]^2 \right\} = \\ &= K \exp \left\{ -\frac{1}{\sigma^2} \sum_{i=1}^n [y_i - \varphi(x_i)]^2 \right\}, \end{aligned} \quad (7.38)$$

где K — коэффициент, не зависящий от $\varphi(x_i)$.

Очевидно, что при заданном σ^2 вероятность P максимальна при условии, что

$$\sum_{i=1}^n [y_i - \varphi(x_i)]^2 = \min.$$

Таким образом, при нормальном распределении случайных величин оптимальность метода наименьших квадратов легко обосновывается.

Нахождение коэффициентов уравнения приближенной регрессии по этому методу связано с задачей определения минимума функции многих переменных. Пусть

$$\hat{y}(x) = f(x, b_0, b_1, b_2, \dots, b_k). \quad (7.40)$$

Требуется найти значения коэффициентов $b_0, b_1, b_2, \dots, b_k$ так, чтобы

$$S = \sum_{i=1}^n \left[y_i - \hat{y}(x_i) \right]^2 = \min.$$

Если S принимает минимальное значение, то

$$\frac{\partial S}{\partial b_0} = 0, \quad \frac{\partial S}{\partial b_1} = 0, \quad \frac{\partial S}{\partial b_2} = 0, \dots, \quad \frac{\partial S}{\partial b_k} = 0, \quad (7.41)$$

что соответствует следующей системе уравнений:

$$\sum_{i=1}^n 2 \left[y_i - \hat{y}(x_i) \right] \frac{\partial \hat{y}(x_i)}{\partial b_0} = 0,$$

$$\sum_{i=1}^n 2 \left[y_i - \hat{y}(x_i) \right] \frac{\partial \hat{y}(x_i)}{\partial b_1} = 0, \quad (7.42)$$

.....,

$$\sum_{i=1}^n 2 \left[y_i - \hat{y}(x_i) \right] \frac{\partial \hat{y}(x_i)}{\partial b_k} = 0.$$

Преобразуем (7.42)

$$\sum_{i=1}^n y_i \frac{\partial \hat{y}(x_i)}{\partial b_0} - \sum_{i=1}^n \hat{y}(x_i) \frac{\partial \hat{y}(x_i)}{\partial b_0} = 0,$$

$$\sum_{i=1}^n y_i \frac{\partial \hat{y}(x_i)}{\partial b_1} - \sum_{i=1}^n \hat{y}(x_i) \frac{\partial \hat{y}(x_i)}{\partial b_1} = 0, \quad (7.43)$$

.....,

$$\sum_{i=1}^n y_i \frac{\partial \hat{y}(x_i)}{\partial b_k} - \sum_{i=1}^n \hat{y}(x_i) \frac{\partial \hat{y}(x_i)}{\partial b_k} = 0.$$

В последней системе содержится столько же $(k + 1)$ уравнений, сколько и неизвестных коэффициентов в уравнении (7.40), т. е. она является *системой нормальных уравнений*. Поскольку $S \geq 0$ при любых значениях коэффициентов, то у нее должен существовать по меньшей мере один минимум. Поэтому если система (7.43) имеет единственное решение, то оно и является минимумом для S .

ЛЕКЦИЯ 8

Линейная регрессия от одного параметра. Регрессионный анализ. Аппроксимация, параболическая регрессия. Оценка тесноты нелинейной связи, корреляционный анализ. Метод множественной корреляции.

8.1. Линейная регрессия от одного параметра

Пусть из опытов получена выборка точек (x_i, y_i) объемом n . Найдем методом наименьших квадратов коэффициенты линейного уравнения регрессии

$$\hat{y} = b_0 + b_1 x. \quad (8.1)$$

Система нормальных уравнений (7.43) с учетом того, что

$$\hat{y}(x_i) = b_0 + b_1 x_i,$$

принимает вид

$$\begin{aligned} \sum_{i=1}^n y_i - \sum_{i=1}^n (b_0 + b_1 x_i) &= 0, \\ \sum_{i=1}^n y_i x_i - \sum_{i=1}^n (b_0 + b_1 x_i) x_i &= 0, \end{aligned} \quad (8.2)$$

или после преобразования

$$\begin{aligned} nb_0 + b_1 \sum_{i=1}^n x_i &= \sum_{i=1}^n y_i, \\ b_0 \sum_{i=1}^n x_i + b_1 \sum_{i=1}^n x_i^2 &= \sum_{i=1}^n y_i x_i. \end{aligned} \quad (8.3)$$

Решив систему уравнений, получим

$$b_0 = \frac{\sum_{i=1}^n y_i \sum_{i=1}^n x_i^2 - \sum_{i=1}^n x_i \sum_{i=1}^n x_i y_i}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2}, \quad (8.4)$$

$$\begin{aligned}
b_1 &= \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \\
&= \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{(n-1) s_x^2}. \tag{8.5}
\end{aligned}$$

Из системы уравнений (8.3) видно, что между коэффициентами b_0 и b_1 существует корреляционная зависимость, выражение для которой можно получить, например, из первого уравнения системы:

$$b_0 = \bar{y} - b_1 \bar{x}. \tag{8.6}$$

Выборочный коэффициент корреляции с учетом (8.5) равен

$$r_{xy}^* = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{(n-1) s_x s_y} = \frac{b_1 (n-1) s_x^2}{(n-1) s_x s_y} = \frac{b_1 s_x}{s_y} \tag{8.7}$$

и оценивает силу линейной связи между Y и X .

8.2. Регрессионный анализ

Итак, уравнение линейной регрессии определено. Проведем статистический анализ полученных результатов, заключающийся в оценке значимости коэффициентов регрессии и проверки адекватности полученного уравнения экспериментальным данным. Подобный анализ и называется *регрессионным*.

Примем, что

- 1) входной параметр x измеряется с гораздо большей точностью по сравнению с выходной величиной y ;
- 2) значения y_i получены независимым образом и нормально распределены;
- 3) если при каждом заданном значении x_i проводится серия параллельных опытов, то выборочные дисперсии s_i^2 однородны.

8.2.1. Проверка адекватности приближенного уравнения регрессии эксперименту

Рассмотрим три наиболее часто встречающихся варианта проверки адекватности полученного уравнения регрессии.

1. Пусть при каждом значении x_i проведена серия из m параллельных опытов. Тогда дисперсия воспроизводимости с числом степеней свободы $f_{\text{воспр.}} = n(m - 1)$ равна

$$s_{\text{воспр.}}^2 = \frac{\sum_{i=1}^n s_i^2}{n}. \quad (8.8)$$

Дисперсия адекватности определяется формулой

$$s_{\text{ад.}}^2 = \frac{m \sum_{i=1}^n \left(\bar{y}_i - \hat{y}(x_i) \right)^2}{n - l}, \quad (8.9)$$

где l — число коэффициентов в уравнении регрессии (при линейной регрессии $l = 2$),

$$\bar{y}_i = \frac{1}{m} \sum_{u=1}^m y_{iu}. \quad (8.10)$$

Число степеней свободы дисперсии адекватности равно $f_{\text{ад.}} = n - l$.

Адекватность уравнения проверяется по критерию Фишера

$$F = s_{\text{ад.}}^2 / s_{\text{воспр.}}^2. \quad (8.11)$$

Если вычисленное значение F окажется меньше табличной величины $F_{1-p}(f_1, f_2)$ для уровня значимости p и числа степеней свободы $f_1 = f_{\text{ад.}}$ и $f_2 = f_{\text{воспр.}}$, то уравнение адекватно эксперименту.

2. Основная серия опытов проведена без параллельных, а дисперсия воспроизводимости определена в отдельной серии из m опытов, тогда

$$s_{\text{ад.}}^2 = \frac{\sum_{i=1}^n \left(y_i - \hat{y}(x_i) \right)^2}{n - l}, \quad (8.12)$$

$$s_{\text{воспр.}}^2 = \frac{\sum_{u=1}^m (y_u^0 - \bar{y}^0)^2}{m-1}, \quad \text{где } \bar{y}^0 = \frac{1}{m} \sum_{u=1}^m y_u^0. \quad (8.13)$$

Адекватность уравнения проверяется по критерию Фишера (8.11), при этом $f_2 = f_{\text{воспр.}} = m - 1$.

3. Основная серия опытов выполнена без параллельных, и нет данных для расчета дисперсии воспроизводимости. Тогда по критерию Фишера сравнивается дисперсия адекватности и дисперсия относительно среднего

$$F = \frac{s_y^2(f_1)}{s_{\text{ад.}}^2(f_2)}, \quad (8.14)$$

где

$$s_y^2 = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n-1}. \quad (8.15)$$

Чем больше полученное F превышает табличное $F_{1-p}(f_1, f_2)$ для уровня значимости p и чисел степеней свободы $f_1 = n - 1$ и $f_2 = n - l$, тем эффективнее уравнение регрессии.

8.2.2. Оценка значимости коэффициентов уравнения регрессии

Значимость коэффициентов уравнения регрессии оценивается по критерию Стьюдента

$$t_j = \frac{|b_j|}{s(b_j)}, \quad (8.16)$$

где b_j — j -й коэффициент уравнения регрессии; $s(b_j)$ — среднее квадратичное отклонение j -го коэффициента. Если t_j больше табличной величины $t_{1-p/2}$ для выбранного уровня значимости p и числа степеней свободы f дисперсии j -го коэффициента, то коэффициент b_j значимо отличается от нуля.

В случае линейной регрессии средние квадратичные отклонения коэффициентов рассчитываются следующим образом:

$$s(b_0) = \sqrt{\left(s^2 \sum_{i=1}^n x_i^2 \right) / \left(n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2 \right)}, \quad (8.17)$$

$$s(b_1) = \sqrt{\left(s^2 n \right) / \left(n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2 \right)}, \quad (8.18)$$

где дисперсия s^2 в общем случае определяется как

$$s^2 = \frac{f_{\text{воспр.}} s_{\text{воспр.}}^2 + f_{\text{ад.}} s_{\text{ад.}}^2}{f_{\text{воспр.}} + f_{\text{ад.}}} = \frac{n(m-1) s_{\text{воспр.}}^2 + (n-l) s_{\text{ад.}}^2}{n(m-1) + (n-l)}. \quad (8.19)$$

Число степеней свободы средневзвешенной дисперсии s^2 равно

$$f = n(m-1) + (n-l) = nm - n + n - l = nm - l.$$

Дисперсии воспроизводимости и адекватности рассчитываются по формулам (8.8) и (8.9) или (8.12) и (8.13). Если у экспериментатора нет оснований сомневаться в линейном характере изучаемой зависимости и опыты проведены без параллельных (т. е. $m = 1$), то $s^2 = s_{\text{ад.}}^2$ и $f = f_{\text{ад.}} = n - l$. Дисперсия адекватности в этом случае определяется по формуле (8.12).

Для оценки случайных ошибок в определении коэффициентов приближенного уравнения регрессии можно также воспользоваться критерием Стьюдента. Рассмотрим величину

$$t = \frac{b_0 - \beta_0}{s(b_0)}, \quad (8.20)$$

где β_0 — истинное значение коэффициента b_0 . Произведя выкладки, аналогичные представленным в лекции 4, получим

$$b_0 - s(b_0) t_{1-p/2} \leq \beta_0 \leq b_0 + s(b_0) t_{1-p/2}, \quad (8.21)$$

или

$$\beta_0 = b_0 \pm s(b_0) t_{1-p/2}, \quad (8.22)$$

где $t_{1-p/2}$ — квантиль t -распределения для числа степеней свободы f и выбранного уровня значимости p .

Аналогично можно построить доверительный интервал для коэффициента b_1 :

$$b_1 - s(b_1)t_{1-p/2} \leq \beta_1 \leq b_1 + s(b_1)t_{1-p/2}, \quad (8.23)$$

$$\beta_1 = b_1 \pm s(b_1)t_{1-p/2}. \quad (8.24)$$

С учетом (8.22) и (8.24), уравнение регрессии принимает следующий вид:

$$y = \beta_0 + \beta_1 x = (b_0 \pm s(b_0)t_{1-p/2}) + (b_1 \pm s(b_1)t_{1-p/2})x.$$

8.2.3. Оценка доверительного интервала для искомой функции

На практике нередко возникает необходимость в оценке точек, резко выделяющихся из общей линейной закономерности. Подобную оценку легко произвести, построив доверительный интервал («коридор ошибок») искомой функции. Под «коридором ошибок» понимают границы, отсчитываемые по обе стороны от полученной прямой и показывающие пределы, в которых должны лежать экспериментальные точки. Точки, лежащие за пределами этого коридора, следует признать ошибочными и исключить из общей выборки.

Воспользуемся критерием Стьюдента и рассмотрим величину

$$t = \frac{\hat{y} - m_{y/x}}{\hat{s}(y)}, \quad (8.25)$$

где $m_{y/x}$ — условное математическое ожидание Y при заданном X ; $\hat{s}(y)$ — выборочное среднеквадратичное отклонение, соответствующее выборочной дисперсии

$$s^2(\hat{y}) = s^2(b_0) + (x^2 - 2x\bar{x})s^2(b_1) \quad (8.26)$$

с числом степеней свободы $f = nm - 2$, если среднеквадратичные отклонения коэффициентов рассчитываются на основе средневзвешенной дисперсии s^2 , определяемой по формуле (8.19), и $f = n - 2$, если $s^2 = s_{ад}^2$. Тогда границы коридора ошибок для произвольного значения аргумента x определяются следующим выражением:

$$m_{y/x} = \hat{y}(x) \pm t_{1-p/2} \cdot \hat{s}(y), \quad (8.27)$$

где $t_{1-p/2}$ — квантиль t -распределения для числа степеней свободы f и выбранного уровня значимости p (обычно 0,05).

Процедура выделения из общей совокупности точек, содержащих грубые ошибки, заключается в следующем. Вначале методом наименьших квадратов обрабатываются все полученные экспериментальные данные, не выбрасывая ни одной точки. Далее по формуле (8.27) для каждой ординаты (для каждого заданного значения x) определяется доверительный интервал при выбранной доверительной вероятности. Если оказывается, что одна или несколько точек при этом выпадают из рассчитанных для них интервалов и величина отклонения превышает систематическую погрешность измерения, то их следует признать ошибочными и исключить из рассмотрения. Затем весь расчет коэффициентов, их случайных ошибок и коридора ошибок повторяется заново.

8.3. Оценка тесноты нелинейной связи

Если уравнение регрессии получено с достаточной точностью, то силу стохастической связи между величинами Y и X можно охарактеризовать величиной

$$\gamma = \frac{(n-1)s_{ад.}^2}{(n-1)s_y^2}. \quad (8.28)$$

Дисперсия адекватности (остаточная дисперсия) и дисперсия относительно среднего рассчитываются по формулам (8.12) и (8.15) соответственно. Связь тем сильнее, чем меньше γ . Величина

$$\Theta = \sqrt{1-\gamma} \quad (8.29)$$

называется *корреляционным отношением*, для которого справедливо

$$0 \leq \Theta \leq 1. \quad (8.30)$$

Чем больше Θ , тем сильнее связь.

В общем случае анализ силы связи по корреляционному отношению называют *корреляционным анализом*. Функциональная зависимость между случайными величинами существует, если $\Theta = 1$. Однако при $\Theta = 0$ однозначно говорить об отсутствии связи можно только в случае нормального распределения случайных величин.

При линейной регрессии корреляционное отношение равно коэффициенту корреляции:

$$\Theta = \sqrt{1 - \frac{(n-2)s_{ад.}^2}{(n-1)s_y^2}} = |r^*|. \quad (8.31)$$

8.4. Аппроксимация. Параболическая регрессия

В общем случае при описании функциональной зависимости между двумя случайными величинами используют полиномы некоторой степени, коэффициенты которых могут и не иметь определенного физического смысла. Такая операция называется *аппроксимацией* экспериментальных данных. Полученная эмпирическая формула обычно справедлива только для сравнительно узкого интервала измерений и неприменима вне этого интервала. При использовании метода наименьших квадратов коэффициенты приближенного уравнения регрессии определяются решением системы линейных уравнений.

Допустим, что зависимость между величинами X и Y описывается параболой второго порядка

$$\hat{y}(x) = b_0 + b_1x + b_2x^2. \quad (8.32)$$

Тогда

$$\frac{\partial \hat{y}(x)}{\partial b_0} = 1, \quad \frac{\partial \hat{y}(x)}{\partial b_1} = x, \quad \frac{\partial \hat{y}(x)}{\partial b_2} = x^2, \quad (8.33)$$

и система нормальных уравнений (7.43) принимает вид

$$\begin{aligned} b_0n + b_1 \sum_{i=1}^n x_i + b_2 \sum_{i=1}^n x_i^2 &= \sum_{i=1}^n y_i, \\ b_0 \sum_{i=1}^n x_i + b_1 \sum_{i=1}^n x_i^2 + b_2 \sum_{i=1}^n x_i^3 &= \sum_{i=1}^n x_i y_i, \\ b_0 \sum_{i=1}^n x_i^2 + b_1 \sum_{i=1}^n x_i^3 + b_2 \sum_{i=1}^n x_i^4 &= \sum_{i=1}^n x_i^2 y_i. \end{aligned} \quad (8.34)$$

Решая систему (8.34), находят коэффициенты искомой квадратичной функции. При описании функциональных зависимостей полиномами большей степени коэффициенты определяются из аналогичных по структуре систем уравнений.

На практике адекватности уравнения регрессии эксперименту добиваются повышением степени аппроксимирующего полинома. При использовании полинома k -степени требуется определять $k + 1$ коэффициент. Увеличение степени полинома прекращают, если дисперсия адекватности (остаточная дисперсия) уравнения регрессии $k + 1$ сте-

пени (s_{k+1}^2) перестает быть значимо меньше дисперсии адекватности, вычисленной для полинома k -степени (s_k^2). Значимость различия исследуется по критерию Фишера

$$F = s_k^2 / s_{k+1}^2,$$

где

$$s_k^2 = \frac{\sum_{i=1}^n (y_i - \hat{y}(x_i))^2}{n - (k + 1)}, \quad s_{k+1}^2 = \frac{\sum_{i=1}^n (y_i - \hat{y}(x_i))^2}{n - (k + 2)}. \quad (8.35)$$

Если полученное F меньше табличного $F_{1-p}(f_1, f_2)$ для уровня значимости p и чисел степеней свободы $f_1 = f_k = n - k - 1$ и $f_2 = f_{k+1} = n - k - 2$, то увеличение степени полинома нужно прекратить и в качестве приближенного уравнения регрессии использовать полином k -степени.

8.5. Приведение некоторых функциональных зависимостей к линейному виду

При малых объемах выборки увеличение порядка полинома может иногда приводить к росту остаточной дисперсии. Чтобы избежать этого, при решении многих задач производят замену переменных. Например, зависимости типа

$$\hat{z} = a_0 a_1^x \quad \text{или} \quad \hat{z} = a_0 t^{a_1} \quad (8.36)$$

сводятся к линейным $\hat{y} = b_0 + b_1 x$ следующим образом:

$$\ln \hat{z} = \ln a_0 + x \ln a_1, \quad \hat{y} = \ln \hat{z}, \quad b_0 = \ln a_0, \quad b_1 = \ln a_1, \quad (8.37)$$

$$\ln \hat{z} = \ln a_0 + a_1 \ln t, \quad \hat{y} = \ln \hat{z}, \quad b_0 = \ln a_0, \quad b_1 = a_1, \quad x = \ln t. \quad (8.38)$$

Коэффициенты уравнений (8.37) и (8.38) находятся методом наименьших квадратов.

Рассмотрим некоторые наиболее часто встречающиеся случаи линеаризации зависимостей при обработке результатов физико-химических экспериментов.

1. Температурная зависимость константы равновесия реакции для небольшого интервала температур имеет вид

$$\ln K = \frac{\Delta S}{R} - \frac{\Delta H}{R} \cdot \frac{1}{T}, \quad (8.39)$$

где ΔS и ΔH — энтропия и энтальпия реакции. Непосредственно измеряемыми величинами являются константа равновесия K и температура T . Произведем замену переменных:

$$\hat{y} = b_0 + b_1 x, \text{ где } \hat{y} = \ln K, \quad b_0 = \frac{\Delta S}{R}, \quad b_1 = -\frac{\Delta H}{R}, \quad x = \frac{1}{T}.$$

Коэффициенты b_0 и b_1 определяются методом наименьших квадратов.

Энтальпия и энтропия реакции с учетом случайных ошибок равны

$$\Delta S = R \cdot (b_0 \pm s(b_0) t_{1-p/2}) = R b_0 \pm R s(b_0) t_{1-p/2},$$

$$\Delta H = -R \cdot (b_1 \pm s(b_1) t_{1-p/2}) = -(R b_1 \pm R s(b_1) t_{1-p/2}).$$

2. Температурная зависимость давления насыщенного пара вещества в узком интервале температур имеет вид

$$\ln P = a - \frac{\Delta H}{R} \cdot \frac{1}{T}, \quad (8.40)$$

где a — константа, ΔH — энтальпия парообразования (испарения или сублимации). Непосредственно определяемыми величинами являются давление насыщенного пара P и температура T . Произведем замену переменных:

$$\hat{y} = b_0 + b_1 x, \text{ где } \hat{y} = \ln P, \quad a = b_0, \quad b_1 = -\frac{\Delta H}{R}, \quad x = \frac{1}{T}.$$

Энтальпия парообразования с учетом случайной ошибки равна

$$\Delta H = -R \cdot (b_1 \pm s(b_1) t_{1-p/2}) = -(R b_1 \pm R s(b_1) t_{1-p/2}).$$

3. Константа скорости реакции первого порядка описывается следующим уравнением:

$$k = \frac{1}{t} \ln \frac{C_0}{C}, \quad (8.41)$$

или

$$\ln C = \ln C_0 - k t,$$

где k — константа скорости реакции, C_0 и C — исходная и текущая концентрация реагирующего вещества к моменту времени t соответственно. Произведем замену переменных:

$$\hat{y} = b_0 + b_1 x, \text{ где } \hat{y} = \ln C, b_0 = \ln C_0, b_1 = k, x = t.$$

Определив коэффициент b_1 методом наименьших квадратов, получим значение константы скорости реакции с учетом случайной ошибки:

$$k = -(b_1 \pm s(b_1) t_{1-p/2}).$$

8.6. Метод множественной корреляции

На практике часто бывает необходимым исследовать корреляционную связь между многими (а не только двумя) величинами. В случае, когда необходимо установить зависимость величины Y от более чем одного параметра, обычно используют уравнения множественной регрессии следующего вида

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2 + \dots + b_k x_k. \quad (8.42)$$

Коэффициенты уравнения находят методом наименьших квадратов, т. е. определяют из условия

$$S = \sum_{i=1}^n \left(y_i - \hat{y}_i \right)^2 = \min, \quad (8.43)$$

где $\hat{y}_i = \hat{y}(x_{1i}, x_{2i}, \dots, x_{ki})$. Условия минимума функции S следующие:

$$\frac{\partial S}{\partial b_0} = 0, \quad \frac{\partial S}{\partial b_1} = 0, \quad \dots, \quad \frac{\partial S}{\partial b_k} = 0. \quad (8.44)$$

Коэффициенты уравнения приближенной регрессии находят из решения системы $(k+1)$ нормальных уравнений, полученных из условий (8.44).

Рассмотрим случай, когда величина Y линейно зависит от двух переменных X_1 и X_2 . Пусть из опытов получена выборка точек (x_{1i}, x_{2i}, y_i) объемом n . Найдем методом наименьших квадратов коэффициенты линейного уравнения регрессии

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2. \quad (8.45)$$

Тогда

$$\frac{\partial \hat{y}}{\partial b_0} = 1, \quad \frac{\partial \hat{y}}{\partial b_1} = x_1, \quad \frac{\partial \hat{y}}{\partial b_2} = x_2. \quad (8.46)$$

Система нормальных уравнений, соответствующих условиям (8.44), принимает следующий вид:

$$\begin{aligned} \sum_{i=1}^n 2 \left[y_i - \hat{y}_i \right] \frac{\partial \hat{y}}{\partial b_0} &= 0, \\ \sum_{i=1}^n 2 \left[y_i - \hat{y}_i \right] \frac{\partial \hat{y}}{\partial b_1} &= 0, \\ \sum_{i=1}^n 2 \left[y_i - \hat{y}_i \right] \frac{\partial \hat{y}}{\partial b_2} &= 0. \end{aligned} \tag{8.47}$$

С учетом того, что $\hat{y}_i = b_0 + b_1 x_{1i} + b_2 x_{2i}$ и значений частных производных (8.46), после арифметических преобразований получаем

$$\begin{aligned} b_0 n + b_1 \sum_{i=1}^n x_{1i} + b_2 \sum_{i=1}^n x_{2i} &= \sum_{i=1}^n y_i, \\ b_0 \sum_{i=1}^n x_{1i} + b_1 \sum_{i=1}^n x_{1i}^2 + b_2 \sum_{i=1}^n x_{1i} x_{2i} &= \sum_{i=1}^n x_{1i} y_i, \\ b_0 \sum_{i=1}^n x_{2i} + b_1 \sum_{i=1}^n x_{1i} x_{2i} + b_2 \sum_{i=1}^n x_{2i}^2 &= \sum_{i=1}^n x_{2i} y_i. \end{aligned} \tag{8.48}$$

Решая полученную систему уравнений относительно b_0 , b_1 и b_2 , находим наилучшую аппроксимацию для соотношения (8.45). Силу линейной связи между переменными X_1 и X_2 можно оценить на основании выборочного коэффициента корреляции

$$r^*(x_1, x_2) = \frac{\sum_{i=1}^n (x_{1i} - \bar{x}_1)(x_{2i} - \bar{x}_2)}{(n-1) \cdot s(x_1) \cdot s(x_2)}. \tag{8.49}$$

ЛЕКЦИЯ 9

Дисперсионный анализ, его задачи. Проведение однофакторного и двухфакторного дисперсионного анализа.

9.1. Задачи дисперсионного анализа. Однофакторный дисперсионный анализ

Средние значения измеряемых величин зависят от комплекса основных факторов (качественных и количественных), определяющих условия проведения опыта, и случайных факторов. *Задачей дисперсионного анализа и является изучение влияния тех или иных факторов на изменчивость средних.* В зависимости от числа источников дисперсии (числа рассматриваемых факторов) различают однофакторный и многофакторный дисперсионный анализ. Многофакторный дисперсионный анализ более эффективен по сравнению с классическим методом исследования, при котором изменяется только один фактор при постоянстве всех остальных, что не позволяет определить влияние взаимодействия различных факторов на результаты эксперимента.

При дисперсионном анализе каждое наблюдение используется для одновременной оценки всех факторов и их взаимодействий. Суть дисперсионного анализа заключается в выделении и оценке отдельных факторов, влияющих на значения среднего. При этом суммарная выборочная дисперсия разлагается на составляющие, обусловленные действием независимых факторов. Влияние данного фактора признается значимым, если соответствующая ему выборочная дисперсия значимо отличается от дисперсии воспроизводимости, обусловленной случайными ошибками. *Проверка значимости оценок дисперсий проводится по критерию Фишера.*

В дальнейшем примем, что:

- 1) случайные ошибки нормально распределены;
- 2) эксперименты равноточны;
- 3) изучаемые факторы влияют только на изменчивость средних, но не на дисперсию наблюдений (она постоянна).

При дисперсионном анализе рассматриваются факторы двух видов: *со случайными уровнями* и *с фиксированными*. В первом случае выбор уровней фактора производится из бесконечной совокупности возможных значений. Если все уровни выбираются случайным образом, то математическая модель объекта называется *моделью со случайными уровнями факторов*. Если же каждый фактор может принимать только некоторые из фиксированных значений, то говорят о *мо-*

дели с фиксированными уровнями факторов. В случае модели смешанного типа одна группа факторов рассматривается на случайных уровнях, а другая — на фиксированных.

Рассмотрим влияние на результаты опытов единичного фактора A , принимающего k различных значений (фактор A имеет k фиксированных уровней $a_i, i = 1, 2, \dots, k$). Обозначим через y_{ij} результат j -опыта в серии из n_i числа измерений ($j = 1, 2, \dots, n_i$), выполненных на i -уровне фактора A (табл. 2).

Таблица 2

Исходные данные для однофакторного дисперсионного анализа

Номер наблюдения	Уровни фактора A			
	a_1	a_2	...	a_k
1	y_{11}	y_{21}	...	y_{k1}
2	y_{12}	y_{22}	...	y_{k2}
...
n	y_{1n}	y_{2n}	...	y_{kn}
Итоги:	B_1, C_1	B_2, C_2	...	B_k, C_k

Предположим, что результат каждого опыта можно представить в виде следующей модели:

$$y_{ij} = \mu + \alpha_i + \varepsilon_{ij}, \quad (9.1)$$

где μ — суммарный эффект во всех опытах; α_i — эффект, обусловленный влиянием фактора A на i -уровне; ε_{ij} — случайная ошибка опыта на i -уровне. Примем также, что наблюдения на фиксированном уровне фактора A нормально распределены относительно среднего значения ($\mu + \alpha_i$) с общей дисперсией $\sigma^2_{\text{ош.}}$. Для того чтобы решить вопрос о значимости влияния фактора A , следует проверить нулевую гипотезу равенства математических ожиданий сумм ($\mu + \alpha_i$) на различных уровнях этого фактора:

$$H_0 : m_1 = m_2 = \dots = m_k = m, \quad (9.2)$$

где $m_i = M\{\mu + \alpha_i\}$.

Рассмотрим случай, когда на каждом уровне выполнено равное число опытов ($n_1 = n_2 = \dots = n_k = n$). Общее число опытов равно

$$N = n_1 + n_2 + \dots + n_k = kn. \quad (9.3)$$

Обозначим сумму результатов всех опытов (итогов) на i -уровне через

$$B_i = \sum_{j=1}^n y_{ij}, \quad (9.4)$$

а сумму квадратов итогов на i -уровне через

$$C_i = \sum_{j=1}^n y_{ij}^2. \quad (9.5)$$

Тогда среднее значение наблюдений на i -уровне равно

$$\bar{y}_i = \frac{\sum_{j=1}^n y_{ij}}{n} = \frac{B_i}{n}, \quad (9.6)$$

а общее среднее для всей выборки из N наблюдений —

$$\bar{y} = \frac{1}{N} \sum_{i=1}^k \sum_{j=1}^n y_{ij} = \frac{1}{k} \sum_{i=1}^k \bar{y}_i = \frac{1}{kn} \sum_{i=1}^k B_i. \quad (9.7)$$

Общая выборочная дисперсия опытов определяется выражением

$$\begin{aligned} s^2 &= \frac{\sum_{i=1}^k \sum_{j=1}^n (y_{ij} - \bar{y})^2}{N-1} = \frac{1}{N-1} \left[\sum_{i=1}^k \sum_{j=1}^n y_{ij}^2 - \frac{1}{N} \left(\sum_{i=1}^k \sum_{j=1}^n y_{ij} \right)^2 \right] = \\ &= \frac{1}{N-1} \left[\sum_{i=1}^k C_i - \frac{1}{N} \left(\sum_{i=1}^k B_i \right)^2 \right], \end{aligned} \quad (9.8)$$

а выборочная дисперсия на i -уровне —

$$\begin{aligned} s_i^2 &= \frac{1}{n-1} \sum_{j=1}^n (y_{ij} - \bar{y}_i)^2 = \frac{1}{n-1} \left[\sum_{j=1}^n y_{ij}^2 - \frac{1}{n} \left(\sum_{j=1}^n y_{ij} \right)^2 \right] = \\ &= \frac{1}{n-1} \left[C_i - \frac{B_i^2}{n} \right]. \end{aligned} \quad (9.9)$$

Если выборочные дисперсии s_i^2 однородны (проверка по критерию Кохрена), то лучшей оценкой дисперсии $\sigma_{\text{ош}}^2$, характеризующей влияние случайных факторов, будет выборочная дисперсия

$$s_{\text{ош}}^2 = \frac{1}{k} \sum_{i=1}^k s_i^2 \quad (9.10)$$

с числом степеней свободы $f_{\text{ош}} = k(n-1) = N-k$. Приближенно оценить дисперсию фактора A можно следующим образом:

$$\sigma_A^2 \approx s^2 - s_{\text{ош}}^2. \quad (9.11)$$

Для получения более точной оценки рассмотрим отклонение средних на фиксированных уровнях от общего среднего:

$$\frac{1}{k-1} \sum_{i=1}^k (\bar{y}_i - \bar{y})^2 \approx \sigma_A^2 + \frac{\sigma_{\text{ош}}^2}{n} \approx \sigma_A^2 + \frac{s_{\text{ош}}^2}{n}. \quad (9.12)$$

В данном случае под дисперсией фактора A понимают математическое ожидание среднего квадрата отклонений, обусловленного влиянием этого фактора. Выборочная дисперсия

$$s_A^2 = \frac{n}{k-1} \sum_{i=1}^k (\bar{y}_i - \bar{y})^2 \approx n\sigma_A^2 + s_{\text{ош}}^2 \quad (9.13)$$

с числом степеней свободы $f_A = k-1$ используется для проверки нулевой гипотезы (9.2) по критерию Фишера.

При этом, если нулевая гипотеза ($H_0 : \sigma_A^2 = \sigma_{\text{ош}}^2$) верна, выполняется следующее условие:

$$\left(s_A^2 / s_{\text{ош}}^2 \right) \leq F_{1-p}, \quad (9.14)$$

т. е. различие между дисперсиями s_A^2 и $s_{\text{ош}}^2$ является незначимым, и следовательно влияние фактора A на результаты опытов тоже незначимо (сопоставимо с эффектом случайности). При проверке гипотезы используется односторонний критерий, так как альтернативной гипотезой является $H_1 : \sigma_A^2 > \sigma_{\text{ош}}^2$. Если же

$$\left(s_A^2 / s_{\text{ош}}^2 \right) > F_{1-p}, \quad (9.15)$$

то нулевая гипотеза о равенстве математических ожиданий сумм ($\mu + \alpha_i$) отвергается (влияние фактора A значимо). Чтобы выяснить,

какие средние различны, можно использовать критерий Стьюдента, сравнивая средние попарно. Оценить влияние фактора A можно на основании (9.13):

$$\sigma_A^2 = \frac{s_A^2 - s_{\text{ош}}^2}{n}. \quad (9.16)$$

Если на каждом уровне выполнено разное число опытов, выборочная дисперсия фактора A рассчитывается по формуле

$$s_A^2 = \frac{1}{k-1} \left[\sum_{i=1}^k \frac{B_i^2}{n_i} - \frac{1}{N} \left(\sum_{i=1}^k B_i \right)^2 \right], \quad (9.17)$$

а выборочная дисперсия, характеризующая влияние случайных факторов, по формуле

$$s_{\text{ош}}^2 = \sum_{i=1}^k f_i s_i^2 / \sum_{i=1}^k f_i, \quad (9.18)$$

где $f_i = n_i - 1$. Число степеней свободы $s_{\text{ош}}^2$ равно $f_{\text{ош}} = N - k$.

Если дисперсия s_A^2 значительно отличается от дисперсии $s_{\text{ош}}^2$, т. е. выполняется неравенство (9.15), то дисперсия фактора A оценивается по формуле

$$\sigma_A^2 \approx \frac{k-1}{N-1} (s_A^2 - s_{\text{ош}}^2). \quad (9.19)$$

9.2. Двухфакторный дисперсионный анализ

Рассмотрим влияние на результаты опытов двух факторов A и B . Фактор A исследуется на k уровнях ($i = 1, 2, \dots, k$), фактор B — на m уровнях ($j = 1, 2, \dots, m$). Пусть при каждом сочетании уровней факторов выполнено n параллельных опытов ($q = 1, 2, \dots, n$). Тогда общее число опытов равно $N = nkm$. Обозначим через y_{ijq} результат q -го опыта, выполненного на i -уровне фактора A и j -уровне фактора B .

Предположим, что результат каждого опыта можно представить следующим образом:

$$y_{ijq} = \mu + \alpha_i + \beta_j + \alpha_i \beta_j + \varepsilon_{ijq}, \quad (9.20)$$

где μ — общее среднее (суммарный эффект во всех опытах); α_i и β_j — эффекты, обусловленные влиянием фактора A на i -уровне и фактором

B на j -уровне соответственно; ε_{ijq} — случайная ошибка опыта, распределенная нормально с нулевым математическим ожиданием и дисперсией $\sigma_{\text{ош}}^2$; $\alpha_i\beta_j$ — эффект взаимодействия факторов. Величина $\alpha_i\beta_j$ характеризует отклонение среднего в (ij) -серии опытов от суммы первых трех членов в ур-и (9.20), а соответствующую ей дисперсию σ_{AB}^2 можно оценить только при наличии параллельных опытов.

При отсутствии параллельных опытов (табл. 3) или в случае, если эффектом взаимодействия факторов пренебрегают, для описания результатов экспериментов используется линейная модель

$$y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}. \quad (9.21)$$

Таблица 3

Исходные данные для двухфакторного дисперсионного анализа без параллельных опытов. Факторы A и B исследуются на 3 уровнях

Уровни фактора B	Уровни фактора A			
	a_1	a_2	$a_3 (a_k)$	Средние:
b_1	y_{11}	y_{21}	$y_{31} (y_{k1})$	\bar{y}'_1
b_2	y_{12}	y_{22}	$y_{32} (y_{k2})$	\bar{y}'_2
$b_3 (b_m)$	y_{13}	y_{23}	$y_{33} (y_{km})$	$\bar{y}'_3 (\bar{y}'_m)$
Средние:	\bar{y}_1	\bar{y}_2	$\bar{y}_3 (\bar{y}_k)$	—

Обозначим через \bar{y}_i и \bar{y}'_j средние по столбцам и по строкам:

$$\bar{y}_i = \frac{\sum_{j=1}^m y_{ij}}{m}, \quad \bar{y}'_j = \frac{\sum_{i=1}^k y_{ij}}{k}, \quad (9.22)$$

а через \bar{y} — среднее всех опытов:

$$\bar{y} = \frac{1}{km} \sum_{i=1}^k \sum_{j=1}^m y_{ij} = \frac{1}{k} \sum_{i=1}^k \bar{y}_i. \quad (9.23)$$

Рассмотрим влияние факторов A и B на рассеяние средних по столбцам и по строкам соответственно относительно общего среднего. Рассеяние в средних по строкам не зависит от фактора A , так как все его уровни усреднены, и определяется влиянием фактора B и случайных факторов.

Тогда с учетом того, что дисперсия среднего в k раз меньше дисперсии случайной ошибки единичного измерения, имеем

$$\sigma_B^2 + \frac{\sigma_{\text{ош}}^2}{k} \approx \frac{1}{m-1} \sum_{j=1}^m (\bar{y}'_j - \bar{y})^2. \quad (9.24)$$

Аналогичным образом можно показать, что

$$\sigma_A^2 + \frac{\sigma_{\text{ош}}^2}{m} \approx \frac{1}{k-1} \sum_{i=1}^k (\bar{y}_i - \bar{y})^2. \quad (9.25)$$

Таким образом, чтобы оценить дисперсии факторов A и B , необходимо знать дисперсию случайной ошибки.

Оценить влияние случайных факторов при отсутствии параллельных опытов можно следующим образом. Рассеяние результатов опытов в i -столбце относительно его среднего обусловлено влиянием фактора B и фактора случайности:

$$s_i^2 = \frac{1}{m-1} \sum_{j=1}^m (y_{ij} - \bar{y}_i)^2 \approx \sigma_B^2 + \sigma_{\text{ош}}^2. \quad (9.26)$$

Равенство (9.26) станет более точным, если использовать средневзвешенное значение дисперсии по всем столбцам:

$$\sigma_B^2 + \sigma_{\text{ош}}^2 \approx \frac{1}{k} \sum_{i=1}^k s_i^2. \quad (9.27)$$

Вычитая (9.24) из (9.27), получим

$$\sigma_{\text{ош}}^2 - \frac{\sigma_{\text{ош}}^2}{k} \approx \frac{1}{k} \sum_{i=1}^k s_i^2 - \frac{1}{m-1} \sum_{j=1}^m (\bar{y}'_j - \bar{y})^2, \quad (9.28)$$

или после арифметических преобразований

$$\sigma_{\text{ош}}^2 \approx \frac{1}{(k-1)(m-1)} \left[(m-1) \sum_{i=1}^k s_i^2 - k \sum_{j=1}^m (\bar{y}'_j - \bar{y})^2 \right] \cong s_{\text{ош}}^2. \quad (9.29)$$

Полученную оценку для дисперсии случайной ошибки с числом степеней свободы $f_{\text{ош}} = (k-1)(m-1)$ обозначим через $s_{\text{ош}}^2$. Определим также следующие выборочные дисперсии:

$$s_A^2 = \frac{m}{k-1} \sum_{i=1}^k (\bar{y}_i - \bar{y})^2 \approx m\sigma_A^2 + s_{\text{ош}}^2, \quad (9.30)$$

$$s_B^2 = \frac{k}{m-1} \sum_{j=1}^m (\bar{y}'_j - \bar{y})^2 \approx k\sigma_B^2 + s_{\text{ош}}^2 \quad (9.31)$$

с числом степеней свободы $f_A = (k - 1)$ и $f_B = (m - 1)$.

Проверка нулевой гипотезы о незначимости влияния факторов A и B проводится по критерию Фишера: если

$$\frac{s_A^2}{s_{\text{ош}}^2} \leq F_{1-p}(f_A, f_{\text{ош}}) \text{ и (или) } \frac{s_B^2}{s_{\text{ош}}^2} \leq F_{1-p}(f_B, f_{\text{ош}}), \quad (9.32)$$

то влияние фактора признается незначимым ($\alpha_i = 0$ и (или) $\beta_j = 0$).

Если одно (или оба) из неравенств (9.32) не выполняется, то влияние соответствующего фактора (факторов) значимо. Определить, какие именно средние различны, можно по критерию Стьюдента.

Рассмотрим теперь случай, когда при каждом сочетании уровней факторов A и B выполнено n параллельных опытов ($u = 1, 2, \dots, n$), что дает возможность оценить влияние взаимодействия этих факторов на результаты опытов.

Так, например, в табл. 3 вместо одного значения y_{11} появится серия значений $y_{111}, y_{112}, \dots, y_{11n}$. Обозначим через \bar{y}_{ij} среднее в ячейке (среднее серии параллельных опытов):

$$\bar{y}_{ij} = \frac{1}{n} \sum_{u=1}^n y_{iju} \quad (9.33)$$

Тогда

$$\bar{y}_i = \frac{1}{m} \sum_{j=1}^m \bar{y}_{ij}, \quad \bar{y}'_j = \frac{1}{k} \sum_{i=1}^k \bar{y}_{ij}, \quad (9.34)$$

$$\bar{y} = \frac{1}{km} \sum_{i=1}^k \sum_{j=1}^m \bar{y}_{ij} = \frac{1}{k} \sum_{i=1}^k \bar{y}_i \quad (9.35)$$

и дисперсии s_A^2 и s_B^2 рассчитываются по формулам (9.30) и (9.31).

В качестве оценки дисперсии воспроизводимости используем средневзвешенное значение дисперсий результатов в каждой ячейке

$$s_{\text{ош}}^2 = \frac{1}{mk} \sum_{i=1}^k \sum_{j=1}^m s_{ij}^2, \quad (9.36)$$

где

$$s_{ij}^2 = \frac{1}{n-1} \sum_{u=1}^n (y_{iju} - \bar{y}_{ij})^2. \quad (9.37)$$

Число степеней свободы дисперсии $s_{\text{ош}}^2$ равно $f_{\text{ош}} = mk(n-1)$.

Введем также выборочную дисперсию, характеризующую влияние взаимодействия факторов

$$s_{AB}^2 \approx \frac{n}{(k-1)(m-1)} \left[\sum_{i=1}^k \sum_{j=1}^m (\bar{y}_{ij} - \bar{y}_i)^2 - k \sum_{j=1}^m (\bar{y}_j - \bar{y})^2 \right], \quad (9.38)$$

с числом степеней свободы $f_{AB} = (k-1)(m-1)$.

Проверка значимости влияния факторов и их взаимодействия проводится по критерию Фишера, но неодинаково для моделей с фиксированными и случайными уровнями:

1. Для модели с фиксированными уровнями выборочные дисперсии s_A^2 , s_B^2 и s_{AB}^2 сравниваются с оценкой дисперсии воспроизводимости $s_{\text{ош}}^2$. Если выполняются неравенства

$$\begin{aligned} (s_A^2 / s_{\text{ош}}^2) > F_{1-p}(f_A, f_{\text{ош}}), \quad (s_B^2 / s_{\text{ош}}^2) > F_{1-p}(f_B, f_{\text{ош}}), \\ (s_{AB}^2 / s_{\text{ош}}^2) > F_{1-p}(f_{AB}, f_{\text{ош}}), \end{aligned} \quad (9.39)$$

то влияние факторов и их взаимодействия значимо.

2. Для модели со случайными уровнями проверка значимости взаимодействия факторов проводится так же, как и для модели с фиксированными уровнями. Влияние факторов значимо, если выполняются следующие неравенства:

$$\begin{aligned} (s_A^2 / s_{AB}^2) > F_{1-p}(f_A, f_{AB}), \\ (s_B^2 / s_{AB}^2) > F_{1-p}(f_B, f_{AB}). \end{aligned} \quad (9.40)$$

ЛЕКЦИЯ 10

Планирование эксперимента при дисперсионном анализе. Постановка задачи при планировании экстремальных экспериментов. Полный факторный эксперимент типа 2^2 : матрица планирования, вычисление коэффициентов уравнения регрессии.

10.1. Планирование эксперимента при дисперсионном анализе

При двухфакторном дисперсионном анализе минимальное число опытов (в условиях линейной модели), обеспечивающее перебор всех возможных сочетаний уровней факторов, определяется произведением числа их уровней: $N = km$. Подобный эксперимент называется *полным факторным экспериментом* (ПФЭ). Если изучается влияние на процесс k факторов при одинаковом числе уровней n , то необходимое число опытов при ПФЭ равно

$$N = n^k. \quad (10.1)$$

Так, если $k = 2$ и $n = 3$ (табл. 3, лекция 9), то $N = 3^2 = 9$.

Эксперимент, в котором пропущены некоторые сочетания уровней, называется *дробным факторным экспериментом* (ДФЭ). Сокращение числа опытов неизбежно приводит к потере части информации, при этом обычно пренебрегают эффектами взаимодействия факторов.

Рассмотрим *трехфакторный дисперсионный анализ при одинаковом числе уровней n для каждого фактора*. Пусть $n = 2$. Тогда при ПФЭ потребуется провести $N = 2^3 = 8$ опытов (табл. 4).

Таблица 4

Полный факторный эксперимент 2^3

Уровни факторов	a_1		a_2	
	b_1	b_2	b_1	b_2
c_1	* y_{111}	y_{121}	y_{211}	* y_{221}
c_2	y_{112}	* y_{122}	* y_{212}	y_{222}

При отсутствии параллельных опытов результаты наблюдений можно представить в виде линейной модели

$$y_{ijq} = \mu + \alpha_i + \beta_j + \gamma_q + \varepsilon_{ijq}, \quad (10.2)$$

при этом линейные эффекты оказываются смешанными с эффектами взаимодействия: эффект A с BC взаимодействием, эффект B с AC взаимодействием, эффект C с AB взаимодействием. Однако число опытов

в условиях линейной модели можно существенно сократить при использовании ДФЭ, спланированного по схеме латинского квадрата.

Латинским квадратом $n \times n$ называют квадратную таблицу, составленную из n элементов (чисел или букв) таким образом, чтобы каждый элемент повторялся в каждой строке и каждом столбце только один раз. Из двух элементов образуется латинский квадрат 2×2 :

$$\begin{array}{cc} A B & \text{или} \\ B A & \end{array} \begin{array}{c} c_1 c_2; \\ c_2 c_1 \end{array}; \quad (10.3)$$

из трех — латинский квадрат 3×3 :

$$\begin{array}{ccc} A B C & & c_1 c_2 c_3 \\ B C A & \text{или} & c_2 c_3 c_1 \\ C A B & & c_3 c_1 c_2 \end{array}. \quad (10.4)$$

Стандартными латинскими квадратами называются квадраты, у которых первая строка и столбец построены или в алфавитном порядке, или в порядке натурального ряда (квадраты (10.3) и (10.4)). Получены эти квадраты путем одношаговой циклической перестановки.

При ДФЭ по схеме латинского квадрата вводится в планирование третий фактор, при этом основой служит ПФЭ типа n^2 . Так, при $n = 2$ на ПФЭ типа 2^2 (для факторов A и B) накладывается латинский квадрат 2×2 (табл. 5). План эксперимента, соответствующий табл. 5, называется *матрицей планирования* и представлен в табл. 6. Число опытов при этом сокращается до четырех вместо восьми при ПФЭ.

Таблица 5
2 x 2 латинский квадрат

A	B	
	b_1	b_2
a_1	c_1	c_2
a_2	c_2	c_1

Хотя латинский квадрат 2×2 является частью плана, всю табл. 5 также называют латинским квадратом. В нем каждый элемент повторяется только один раз в каждой строке и каждом столбце, что в равной степени сказывается при подсчете средних по строкам и столбцам. Приведенный в табл. 6 план представляет собой половину — *полуреплику* от ПФЭ типа 2^3 (вошедшие в полуреплику опыты отмечены в табл. 4 звездочками).

Таблица 6

План ДФЭ по схеме латинского квадрата 2 x 2
($k = 3, n = 2, N = 4$)

Номер опыта	<i>A</i>	<i>B</i>	<i>C</i>	Итоги
1	a_1	b_1	c_1	y_{111}
2	a_1	b_2	c_2	y_{122}
3	a_2	b_1	c_2	y_{212}
4	a_2	b_2	c_1	y_{221}

Аналогично планируется ДФЭ по схеме латинского квадрата 3 x 3 (табл. 7). За основу взят ПФЭ типа 3^2 , третий фактор (*C*) введен в рассмотрение по схеме латинского квадрата (10.4). ДФЭ 3^2 можно рассматривать как 1/3 реплику от ПФЭ типа 3^3 .

Таблица 7

Латинский квадрат 3 x 3

<i>A</i>	<i>B</i>		
	b_1	b_2	b_3
a_1	c_1 y_1	c_2 y_2	c_3 y_3
a_2	c_2 y_4	c_3 y_5	c_1 y_6
a_3	c_3 y_7	c_1 y_8	c_2 y_9

В общем случае при планировании дробного факторного эксперимента по схеме латинского квадрата число опытов по сравнению с ПФЭ уменьшается в n раз (так, если $n = 4$, то при ПФЭ $N = 4^3 = 64$, а при ДФЭ по схеме латинского квадрата 4×4 — $N = 4^2 = 16$).

Дисперсионный анализ латинского квадрата, выполненного без параллельных опытов, проводится аналогично двухфакторному дисперсионному анализу. При этом для факторов *A* и *B* рассматривается их влияние на рассеяние средних по столбцам и по строкам относительно общего среднего соответственно, а для фактора *C* — на рассеяние средних по латинским буквам C_q . Так, например, для ДФЭ, представленного в табл. 7, средние по латинским буквам равны

$$C_1 = \frac{y_1 + y_6 + y_8}{3}, C_2 = \frac{y_2 + y_4 + y_9}{3}, C_3 = \frac{y_3 + y_5 + y_7}{3}. \quad (10.5)$$

Значимость линейных эффектов проверяют по критерию Фишера. Адекватность принятой линейной модели можно проверить, выполнив для каждого сочетания уровней факторов (для каждой ячейки латинского квадрата) одинаковое число параллельных опытов. При этом наличие параллельных наблюдений используется только для оценки случайной ошибки опыта. Если эффекты взаимодействия незначимы, то остаточная дисперсия будет незначимо отличаться от дисперсии воспроизводимости, обусловленной ошибкой опыта.

10.2. Постановка задачи при планировании экстремальных экспериментов

Решение экстремальных задач физической химии и химической технологии (например, определение оптимальных условий проведения опыта и протекания процесса, оптимального состава материалов) возможно на основе математической модели объекта — *функции отклика*, связывающей выходной параметр, характеризующий результаты эксперимента, с переменными, определяющими условия проведения опыта (*факторами*):

$$y = \varphi(x_1, x_2, \dots, x_k). \quad (10.6)$$

На основе теоретического анализа физико-химических процессов при наличии достаточной информации об их механизмах можно составить детерминированную математическую модель объекта. Однако при проведении большинства исследований механизмы процессов, протекающих в изучаемых объектах, остаются неизвестными, поэтому для решения задач оптимизации необходимо использовать методы математической статистики.

При статистическом подходе математическая модель объекта или процесса представляется в виде полинома, т.е. отрезка ряда Тейлора, в который разлагается неизвестная функция (10.6):

$$y = \beta_0 + \sum_{j=1}^k \beta_j x_j + \sum_{\substack{u,j=1 \\ u \neq j}}^k \beta_{uj} x_u x_j + \sum_{j=1}^k \beta_{jj} x_j^2 + \\ + \sum_{\substack{i,u,j=1 \\ i \neq u \neq j}}^k \beta_{iuj} x_i x_u x_j + \dots, \quad (10.7)$$

где

$$\beta_0 = \varphi(0), \quad \beta_j = \frac{\partial \varphi(0)}{\partial x_j}, \quad \beta_{uj} = \frac{\partial \varphi(0)}{\partial x_u \partial x_j},$$

$$\beta_{jj} = \frac{\partial \varphi(0)}{2\partial x_j^2}, \quad \beta_{uij} = \frac{\partial \varphi(0)}{\partial x_i \partial x_u \partial x_j}. \quad (10.8)$$

Из-за воздействия случайных факторов на результаты опыта при обработке и анализе экспериментальных данных для полиномиальной модели (10.7) находят выборочные коэффициенты регрессии $b_0, b_j, b_{uj}, b_{jj}, b_{uij}$, которые являются оценками соответствующих теоретических коэффициентов. Уравнение регрессии записывается в виде

$$\hat{y} = b_0 + \sum_{j=1}^k b_j x_j + \sum_{\substack{u,j=1 \\ u \neq j}}^k b_{uj} x_u x_j + \sum_{j=1}^k b_{jj} x_j^2 + \sum_{\substack{i,u,j=1 \\ i \neq u \neq j}}^k b_{uij} x_i x_u x_j + \dots$$

$$+ \sum_{\substack{i,u,j=1 \\ i \neq u \neq j}}^k b_{uij} x_i x_u x_j + \dots, \quad (10.9)$$

где b_0 — свободный член; b_j — линейные эффекты; b_{uj} — эффекты парного взаимодействия; b_{jj} — квадратичные эффекты; b_{uij} — эффекты тройного взаимодействия.

В зависимости от целей исследования и имеющейся информации можно ограничиться расчетом только части коэффициентов, пренебрегая влиянием остальных эффектов (например, в условиях линейной модели значимыми считаются только линейные эффекты, квадратичной модели — линейные и квадратичные эффекты, при этом в обоих случаях принимается, что эффекты взаимодействия факторов пренебрежимо малы).

Следует отметить, что на основании оценок теоретических коэффициентов нельзя определить аналитическое выражение функции отклика и, следовательно, получить информацию о механизме процесса. Полиномиальные модели используются только для решения задач оптимизации и управления процессами.

Под *планированием эксперимента* понимают оптимальное (наиболее эффективное) управление ходом эксперимента с целью получения максимально возможной информации на основе минимально допустимого количества опытных данных. Весь эксперимент обычно разби-

вается на несколько этапов. Информация, полученная после каждого этапа, используется для планирования исследований на следующем этапе. Планирование эксперимента позволяет варьировать все факторы и получать одновременно количественные оценки всех эффектов, и при этом, в отличие от классического регрессионного анализа, избежать корреляции между коэффициентами уравнения регрессии.

10.3. Полный факторный эксперимент типа 2^2 : матрица планирования, вычисление коэффициентов уравнения регрессии

При полном факторном эксперименте (ПФЭ) число опытов равно числу всех возможных комбинаций уровней факторов и при одинаковом числе уровней для каждого фактора определяется формулой

$$N = n^k, \quad (10.10)$$

где n — число уровней, k — число факторов ($j = 1, 2, \dots, k$). ПФЭ 2^k называется такое проведение опытов, при котором каждый из k факторов рассматривается только на двух уровнях. При этом уровни факторов представляют собой границы варьирования данного параметра.

Допустим, что изучается влияние на выход продукта (y) двух параметров (факторов): температуры (z_1) в интервале 50–100 °С и давления (z_2) в диапазоне 1–2 атм. При реализации ПФЭ требуется выполнить $N = 2^2 = 4$ опыта. Произведем кодирование факторов (замену переменных):

$$x_j = \frac{z_j - z_j^0}{\Delta z_j}, \quad (10.11)$$

где

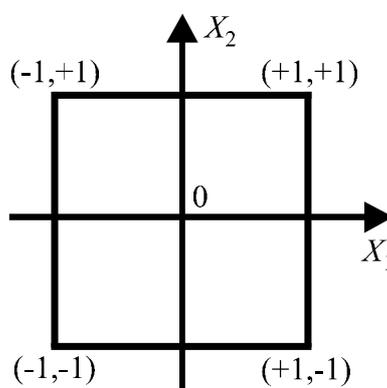
$$z_j^0 = \frac{z_j^{\max} + z_j^{\min}}{2}, \quad \Delta z_j = \frac{z_j^{\max} - z_j^{\min}}{2}, \quad (10.12)$$

z_j^{\max} и z_j^{\min} — верхняя и нижняя границы варьирования j -фактора. Точка (z_1^0, z_2^0) называется *центром плана*, или *основным уровнем*; величины Δz_1 и Δz_2 — интервалами варьирования по осям z_1 и z_2 .

Как следует из уравнений (10.11) и (10.12), для переменных x_1 и x_2 нижний уровень равен -1 , верхний — $+1$, координаты центра плана равны нулю. В табл. 8 представлен план ПФЭ 2^2 , который в безразмерном масштабе может быть интерпретирован в виде четырех вершин квадрата (рис. 1).

Полный факторный эксперимент 2^2

№ опыта	Факторы в натуральном масштабе		Факторы в безразмерном масштабе		Выход продукта, y
	z_1 (°C)	z_2 (атм)	x_1	x_2	
1	50	1	-1	-1	y_1
2	50	2	-1	+1	y_2
3	100	1	+1	-1	y_3
4	100	2	+1	+1	y_4

Рис. 1. Полный факторный эксперимент 2^2

Вычислим коэффициенты линейного уравнения регрессии

$$\hat{y} = b_0 + b_1x_1 + b_2x_2. \quad (10.13)$$

Для нахождения b_0 в план ПФЭ надо ввести столбец фиктивной переменной $x_0 = 1$; соответствующая матрица планирования представлена в табл. 9. В математической статистике доказывается, что при планировании эксперимента по предложенной схеме и нахождении коэффициентов уравнения регрессии по методу наименьших квадратов любой коэффициент определяется скалярным произведением столбца y на соответствующий столбец факторов x_j в безразмерном масштабе (табл. 9), деленным на число опытов в матрице планирования:

$$b_j = \frac{\sum_{i=1}^N x_{ji}y_i}{N}. \quad (10.14)$$

Таблица 9

**Матрица планирования ПФЭ типа 2^2
с фиктивной переменной**

№ опыта	x_0	x_1	x_2	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

Так, значение коэффициента b_1 определяется выражением

$$b_1 = \frac{1}{4} \sum_{i=1}^4 x_{1i} y_i = \frac{[-y_1 + y_2 - y_3 + y_4]}{4}. \quad (10.15)$$

Если ввести в рассмотрение эффект парного взаимодействия, то уравнение регрессии примет вид

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2 + b_{12} x_1 x_2. \quad (10.16)$$

Для нахождения коэффициента b_{12} необходимо расширить матрицу планирования, представленную в табл. 9, добавив в нее столбец $x_1 x_2$, характеризующий эффект взаимодействия (табл. 10).

Таблица 10

Расширенная матрица планирования ПФЭ типа 2^2

№ 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

Значения фактора взаимодействия в безразмерном масштабе определяются произведением соответствующих значений факторов x_1 и x_2 :

$$(x_1 x_2)_i = x_{1i} \cdot x_{2i}. \quad (10.17)$$

Коэффициент b_{12} определяется так же, как и линейные эффекты:

$$b_{12} = \frac{1}{N} \sum_{i=1}^N (x_1 x_2)_i y_i = \frac{1}{4} [y_1 - y_2 - y_3 + y_4]. \quad (10.18)$$

ЛЕКЦИЯ 11

Матрица планирования ПФЭ 2^3 . Проверка значимости коэффициентов и адекватности уравнения регрессии, полученных при обработке результатов ПФЭ 2^2 и 2^3 . Дробный факторный эксперимент. Планы типа 2^{k-1} .

11.1. Матрица планирования полного факторного эксперимента типа 2^3

Рассмотрим планирование ПФЭ типа 2^3 , при котором исследуется влияние на результат опыта уже трех факторов. При реализации такого ПФЭ требуется выполнить $N = 8$ опытов. Проведем кодирование факторов по уравнениям (10.11) – (10.12). План проведения опытов представлен в табл. 11, геометрически в безразмерном масштабе он может быть интерпретирован в виде восьми вершин куба (рис. 2).

Таблица 11

Полный факторный эксперимент 2^3

№ опыта	Факторы в безразмерном масштабе			Выход продукта, y
	x_1	x_2	x_3	
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

Уравнение регрессии с учетом эффектов взаимодействия факторов запишется в следующем виде:

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_{12}x_1x_2 + b_{13}x_1x_3 + b_{23}x_2x_3 + b_{123}x_1x_2x_3, \quad (11.1)$$

где коэффициенты b_{12} , b_{13} и b_{23} характеризуют эффекты парного взаимодействия, b_{123} — эффект тройного взаимодействия.

Для нахождения коэффициентов уравнения (11.1) необходимо составить расширенную матрицу планирования ПФЭ с фиктивной переменной, представленную в табл. 12.

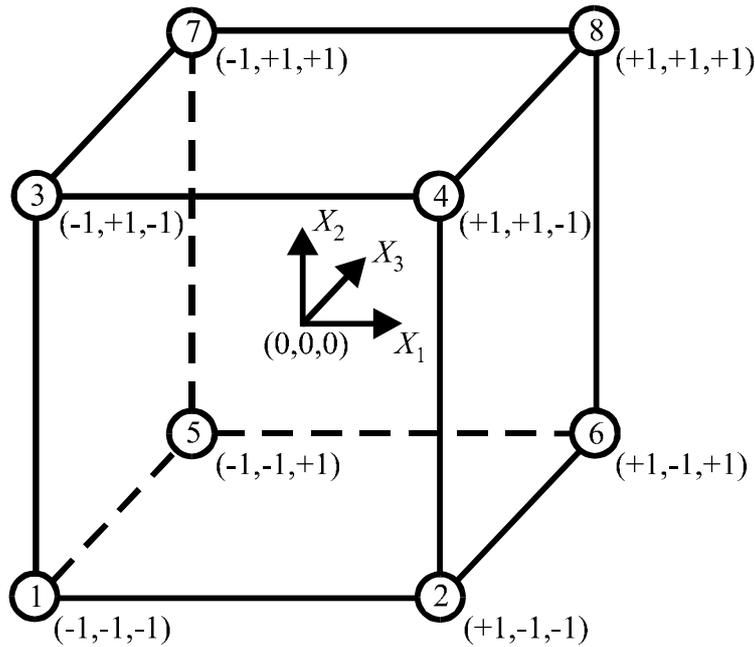


Рис. 2. Полный факторный эксперимент 2^3

Таблица 12

Расширенная матрица планирования ПФЭ типа 2^3

№	x_0	x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3	$x_1x_2x_3$	y
1	+1	-1	-1	-1	+1	+1	+1	-1	y_1
2	+1	+1	-1	-1	-1	-1	+1	+1	y_2
3	+1	-1	+1	-1	-1	+1	-1	+1	y_3
4	+1	+1	+1	-1	+1	-1	-1	-1	y_4
5	+1	-1	-1	+1	+1	-1	-1	+1	y_5
6	+1	+1	-1	+1	-1	+1	-1	-1	y_6
7	+1	-1	+1	+1	-1	-1	+1	-1	y_7
8	+1	+1	+1	+1	+1	+1	+1	+1	y_8

Как и при ПФЭ 2^2 , коэффициенты уравнения регрессии (11.1) определяются скалярным произведением столбца y на соответствующий столбец факторов или их взаимодействий в безразмерном масштабе, деленным на число опытов в матрице планирования (см. уравнения (10.14) и (10.18)).

Так, например, коэффициент b_{123} рассчитывается по следующему выражению:

$$b_{123} = \frac{1}{8} [-y_1 + y_2 + y_3 - y_4 + y_5 - y_6 - y_7 + y_8]. \quad (11.2)$$

11.2. Проверка значимости коэффициентов и адекватности уравнения регрессии, полученных при обработке результатов ПФЭ 2^2 и 2^3

Для оценки значимости коэффициентов уравнения регрессии и проверки адекватности уравнения эксперименту достаточно провести серию параллельных опытов, выполненных при каком-то одном сочетании факторов.

Пусть в центре плана (в точках (z_1^0, z_2^0) и (z_1^0, z_2^0, z_3^0) для ПФЭ 2^2 и 2^3 соответственно) проведена серия из m опытов. Тогда выборочная дисперсия воспроизводимости, характеризующая влияние случайных факторов, равна

$$s_{\text{воспр}}^2 = \frac{\sum_{u=1}^m (y_u^0 - \bar{y}^0)^2}{m-1}, \quad (11.3)$$

где y_u^0 — результат u -го опыта ($u = 1, 2, \dots, m$), \bar{y}^0 — среднее значение серии опытов. В математической статистике доказывается, что для спланированных экспериментов все коэффициенты уравнений регрессии определяются с одинаковой точностью, равной

$$s(b_j) = \frac{s_{\text{воспр}}}{\sqrt{N}}. \quad (11.4)$$

Значимость коэффициентов проверяется по критерию Стьюдента. В условиях нулевой гипотезы $H_0: \beta_j = 0$; отношение абсолютной величины коэффициента к его ошибке имеет распределение Стьюдента. Для каждого коэффициента определяется t -отношение:

$$t_j = \frac{|b_j|}{s(b_j)} = \frac{|b_j|}{s_{\text{воспр}}} \sqrt{N}, \quad (11.5)$$

которое сравнивается с табличным значением критерия Стьюдента $t_p(f)$ для выбранного уровня значимости p (обычно 0,05) и числа степеней свободы $f = m - 1$. Если для рассматриваемого коэффициента $t_j > t_p(f)$, то он значимо отличается от нуля. Выборочные коэффициенты, для которых $t_j \leq t_p(f)$, незначимы, и их следует исключить из уравнения регрессии.

Допустим, при проверке значимости коэффициентов уравнения (11.1) оказалось, что все коэффициенты, характеризующие эффекты

взаимодействия факторов, незначимы. После их исключения получаем линейное уравнение регрессии

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + b_3x_3, \quad (11.6)$$

при этом значения b_0 , b_1 , b_2 и b_3 не требуется вычислять заново из-за того, что коэффициенты уравнения некоррелированы между собой. В отличие от классического регрессионного анализа, исключение незначимого коэффициента не сказывается на величинах остальных коэффициентов уравнения регрессии, а сами выборочные коэффициенты, полученные при реализации ПФЭ, являются *несмешанными* оценками теоретических коэффициентов.

Адекватность уравнения проверяется по критерию Фишера

$$F = \left(s_{\text{ад}}^2 / s_{\text{воспр}}^2 \right), \quad (11.7)$$

Дисперсия адекватности (остаточная дисперсия) равна

$$s_{\text{ад}}^2 = s_{\text{ост}}^2 = \frac{1}{N-l} \sum_{i=1}^N (y_i - \hat{y}_i)^2, \quad (11.8)$$

где l — число значимых коэффициентов (для рассматриваемого случая $l = 4$). Уравнение адекватно описывает эксперимент, если

$$F \leq F_{1-p}(f_1, f_2), \quad (11.9)$$

где $F_{1-p}(f_1, f_2)$ — табличное значение критерия Фишера для $p = 0,05$ и чисел степеней свободы $f_1 = f_{\text{ад}} = N - l$ и $f_2 = f_{\text{воспр}} = m - 1$.

Рассмотрим также схему проведения регрессионного анализа для спланированного эксперимента в случае, когда каждый опыт в матрице планирования повторялся m раз. В качестве примера используем ПФЭ 2^3 ; при получении уравнения регрессии ограничимся линейным приближением (уравнение (11.6)). Матрица планирования такого эксперимента представлена в табл. 13.

Для каждого сочетания уровней факторов определяется среднее значение измеряемой величины и выборочная дисперсия:

$$\bar{y}_i = \frac{1}{m} \sum_{u=1}^m y_{iu}, \quad (11.10)$$

$$s_i^2 = \frac{1}{m-1} \sum_{u=1}^m (y_{iu} - \bar{y}_i)^2. \quad (11.11)$$

Таблица 13

**Матрица планирования ПФЭ 2^3 в условиях линейной модели
с одинаковым числом параллельных опытов
при каждом сочетании уровней факторов**

№	x_0	x_1	x_2	x_3	y	\bar{y}_i	s_i^2
1	+1	-1	-1	-1	$y_{11}, y_{12}, \dots, y_{1m}$	\bar{y}_1	s_1^2
2	+1	+1	-1	-1	$y_{21}, y_{22}, \dots, y_{2m}$	\bar{y}_2	s_2^2
3	+1	-1	+1	-1	$y_{31}, y_{32}, \dots, y_{3m}$	\bar{y}_3	s_3^2
4	+1	+1	+1	-1	$y_{41}, y_{42}, \dots, y_{4m}$	\bar{y}_4	s_4^2
5	+1	-1	-1	+1	$y_{51}, y_{52}, \dots, y_{5m}$	\bar{y}_5	s_5^2
6	+1	+1	-1	+1	$y_{61}, y_{62}, \dots, y_{6m}$	\bar{y}_6	s_6^2
7	+1	-1	+1	+1	$y_{71}, y_{72}, \dots, y_{7m}$	\bar{y}_7	s_7^2
8	+1	+1	+1	+1	$y_{81}, y_{82}, \dots, y_{8m}$	\bar{y}_8	s_8^2

Однородность дисперсий проверяется по критерию Кохрена. Отношение максимальной дисперсии к сумме всех дисперсий

$$G = \frac{s_{\max}^2}{\sum_{i=1}^N s_i^2} \quad (11.12)$$

сравнивается с табличным значением $G_{1-p}(f_1, f_2)$ для $p = 0,05$ и чисел степеней свободы $f_1 = m - 1$ и $f_2 = N$. Если $G \leq G_{1-p}(f_1, f_2)$, то выборочные дисперсии однородны. Тогда наилучшей оценкой дисперсии воспроизводимости будет средневзвешенная дисперсия

$$s_{\text{воспр}}^2 = \frac{1}{N} \sum_{i=1}^N s_i^2 \quad (11.13)$$

с числом степеней свободы $f_{\text{воспр}} = N(m - 1)$.

Коэффициенты уравнения регрессии определяются по формуле

$$b_j = \frac{1}{N} \sum_{i=1}^N x_{ji} \bar{y}_i \quad (11.14)$$

Поскольку дисперсия среднего в m раз меньше дисперсии единичного измерения, т. е.

$$s^2(\bar{y}) = s_{\text{воспр}}^2 / m, \quad (11.15)$$

то выборочные среднеквадратичные отклонения коэффициентов рассчитываются следующим образом:

$$s(b_j) = \frac{s_{\text{воспр}}}{\sqrt{Nm}} = \frac{1}{N\sqrt{m}} \sqrt{\sum_{i=1}^N s_i^2}. \quad (11.16)$$

Значимость коэффициентов проверяется по критерию Стьюдента: если

$$t_j = \frac{|b_j|}{s(b_j)} > t_p(f), \quad (11.17)$$

где $t_p(f)$ — табличное значение критерия Стьюдента для $p = 0,05$ и числа степеней свободы $f = N(m - 1)$, то коэффициент значимо отличается от нуля.

Адекватность уравнения регрессии эксперименту проверяется по критерию Фишера. Дисперсия адекватности равна

$$s_{\text{ад}}^2 = \frac{m \sum_{i=1}^N (\bar{y}_i - \hat{y}_i)^2}{N - l}, \quad (11.18)$$

где l — число значимых коэффициентов в уравнении регрессии.

Уравнение адекватно эксперименту, если

$$F = \frac{s_{\text{ад}}^2}{s_{\text{воспр}}^2} \leq F_{1-p}(f_{\text{ад}}, f_{\text{воспр}}), \quad (11.19)$$

где $F_{1-p}(f_{\text{ад}}, f_{\text{воспр}})$ — табличное значение критерия Фишера для $p = 0,05$ и чисел степеней свободы $f_{\text{ад}} = N - l$ и $f_{\text{воспр}} = N(m - 1)$. В противном случае для описания результатов эксперимента необходимо увеличить порядок аппроксимирующего полинома.

11.3. Дробный факторный эксперимент. Планы типа 2^{k-1}

Число необходимых опытов в условиях линейной модели существенно сокращается при проведении дробных факторных экспериментов (дробных реплик от ПФЭ). В качестве реплики обычно используется полный факторный эксперимент для меньшего числа факторов. При этом вычисление коэффициентов уравнения и оценка их значимости проводится так же, как и в рассмотренных выше примерах ПФЭ

2^2 и 2^3 . Число опытов в дробной реплике должно быть больше или равно числу неизвестных коэффициентов в уравнении регрессии.

Спланируем дробный факторный эксперимент для получения линейного уравнения регрессии небольшого участка поверхности отклика при трех независимых факторах:

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + b_3x_3. \quad (11.20)$$

Постановка ПФЭ 2^3 требует проведения 8 опытов. Для решения же поставленной задачи можно ограничиться 4 опытами, если в матрице планирования ПФЭ 2^2 (табл. 10, лекция 10) использовать столбец x_1x_2 в качестве плана для x_3 . Матрица планирования такого сокращенного эксперимента — ДФЭ типа 2^{3-1} , или *полуреплики* от ПФЭ 2^3 , — представлена в табл. 14.

Таблица 14

Матрица планирования ДФЭ типа 2^{3-1}

№ опыта	x_0	x_1	x_2	x_3	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

Проведение ДФЭ по предложенной схеме позволяет оценить свободный член и три коэффициента при линейных членах уравнения (11.20), однако при этом они будут являться несмешанными оценками теоретических коэффициентов только в том случае, если генеральные коэффициенты регрессии при парных взаимодействиях равны нулю. В противном случае найденные выборочные коэффициенты будут смешанными оценками теоретических:

$$b_1 \rightarrow \beta_1 + \beta_{23}, \quad b_2 \rightarrow \beta_2 + \beta_{13}, \quad b_3 \rightarrow \beta_3 + \beta_{12}. \quad (11.21)$$

Генеральные коэффициенты не могут быть оценены по отдельности на основании только 4 опытов, поскольку при этом столбцы для линейных членов и парных произведений одинаковы (например, элементы вычисленного столбца для произведения x_2x_3 в точности совпадают с элементами столбца x_1). Чтобы определить, оценкой суммы каких именно генеральных коэффициентов являются выборочные коэффициенты, удобно пользоваться *генерирующим соотношением*

$$x_3 = x_1x_2, \quad (11.22)$$

в общем случае означаящим, какой именно столбец ПФЭ 2^k был использован в качестве плана для введения $(k + 1)$ -го фактора в ДФЭ. При умножении обеих частей (11.22) на x_3 , получаем

$$x_3^2 = x_1x_2x_3. \quad (11.23)$$

Единичный столбец

$$I = x_1x_2x_3 \quad (11.24)$$

называется *определяющим контрастом* и позволяет определить, элементы каких столбцов в расширенной матрице планирования одинаковы. Умножая I по очереди на x_1 , x_2 и x_3 , получаем

$$\begin{aligned} x_1 &= x_1^2x_2x_3 = x_2x_3, & x_2 &= x_1x_2^2x_3 = x_1x_3, \\ x_3 &= x_1x_2x_3^2 = x_1x_2, \end{aligned} \quad (11.25)$$

в точности соответствующих системе смешанных оценок (11.21).

При постановке ДФЭ с числом факторов $k \geq 4$ в зависимости от генерирующего соотношения выборочные коэффициенты регрессии оказываются смешанными оценками того или иного сочетания генеральных коэффициентов. Поэтому важно заранее определиться с тем, какая информация является наиболее важной в данном исследовании, и в зависимости от поставленной задачи подобрать нужную дробную реплику.

Рассмотрим, например, планирование ДФЭ типа 2^{4-1} , представляющего собой полуреплику от ПФЭ 2^4 . В качестве реплики используем ПФЭ 2^3 (табл. 12). Используем два генерирующих соотношения:

$$x_4 = x_1x_2x_3, \quad (11.26)$$

$$x_4 = x_1x_3. \quad (11.27)$$

Для соотношения (11.26) определяющим контрастом будет

$$I = x_1x_2x_3x_4. \quad (11.28)$$

Тогда

$$\begin{aligned} x_1 &= x_2x_3x_4, & b_1 &\rightarrow \beta_1 + \beta_{234}; & x_2 &= x_1x_3x_4, & b_2 &\rightarrow \beta_2 + \beta_{134}; \\ x_3 &= x_1x_2x_4, & b_3 &\rightarrow \beta_3 + \beta_{124}; & x_4 &= x_1x_2x_3, & b_4 &\rightarrow \beta_4 + \beta_{123}; \\ x_1x_2 &= x_3x_4, & b_{12} &\rightarrow \beta_{12} + \beta_{34}; & x_1x_3 &= x_2x_4, & b_{13} &\rightarrow \beta_{13} + \beta_{24}; \\ & & & & x_1x_4 &= x_2x_3, & b_{14} &\rightarrow \beta_{14} + \beta_{23}. \end{aligned} \quad (11.29)$$

В реальных задачах влияние тройных взаимодействий обычно равно нулю. Следовательно, генерирующее соотношение (11.26) следует ис-

пользовать, если наибольший интерес представляют оценки для линейных эффектов.

Для соотношения (11.27) определяющим контрастом будет

$$I = x_1 x_3 x_4. \quad (11.30)$$

Тогда

$$\begin{aligned} x_1 = x_3 x_4, b_1 &\rightarrow \beta_1 + \beta_{34}; & x_2 = x_1 x_2 x_3 x_4, b_2 &\rightarrow \beta_2 + \beta_{1234}; \\ x_3 = x_1 x_4, b_3 &\rightarrow \beta_3 + \beta_{14}; & x_4 = x_1 x_3, b_4 &\rightarrow \beta_4 + \beta_{13}; \\ x_1 x_2 = x_2 x_3 x_4, b_{12} &\rightarrow \beta_{12} + \beta_{234}; & x_2 x_3 = x_1 x_2 x_4, b_{23} &\rightarrow \beta_{23} + \beta_{124}; \\ x_2 x_4 = x_1 x_2 x_3, b_{24} &\rightarrow \beta_{24} + \beta_{123}. \end{aligned} \quad (11.31)$$

Следовательно, дробную реплику с генерирующим соотношением $x_4 = x_1 x_3$ следует использовать, если наибольший интерес представляют эффекты парных взаимодействий.

В общем случае число опытов в дробной реплике должно удовлетворять следующему соотношению:

$$k + 1 \leq N < 2^k, \quad (11.32)$$

где k — число факторов. Если число опытов равно числу определяемых коэффициентов в линейном уравнении регрессии ($N = k + 1$), дробная реплика представляет собой *линейный насыщенный план*, для которого все линейные эффекты смешаны с эффектами взаимодействия. Число степеней свободы остаточной дисперсии в таких планах равно нулю, поэтому для проверки адекватности линейного уравнения необходимо проведение дополнительных опытов.

Итак, рассмотренные двухуровневые планы ПФЭ 2^k и ДФЭ 2^{k-1} обладают следующими свойствами: вычисления просты; все коэффициенты регрессии определяются независимо друг от друга и с одинаковой и минимальной дисперсией; каждый коэффициент рассчитывается по результатам всех опытов.

ЛЕКЦИЯ 12

Оптимизация методом крутого восхождения по поверхности отклика. Описание функции отклика в области, близкой к экстремуму. Композиционные планы Бокса-Уилсона. Ортогональные планы второго порядка, расчет коэффициентов уравнения регрессии. Метод последовательного симплекс-планирования.

12.1. Оптимизация методом крутого восхождения по поверхности отклика

Задача оптимизации сводится к опытному определению такого сочетания уровней k факторов (координаты точки в $(k+1)$ -мерном факторном пространстве), при котором достигается максимальное (минимальное) значение выходного параметра y (или нескольких параметров), т. е. функция отклика системы

$$y = \varphi(x_1, x_2, \dots, x_k)$$

принимает экстремальное значение.

Рассмотрим случай, когда на систему оказывают влияние только два фактора (x_1 и x_2 в безразмерном масштабе). Построим контурные сечения $y = const$ поверхности отклика при $k = 2$ (рис. 3 а).

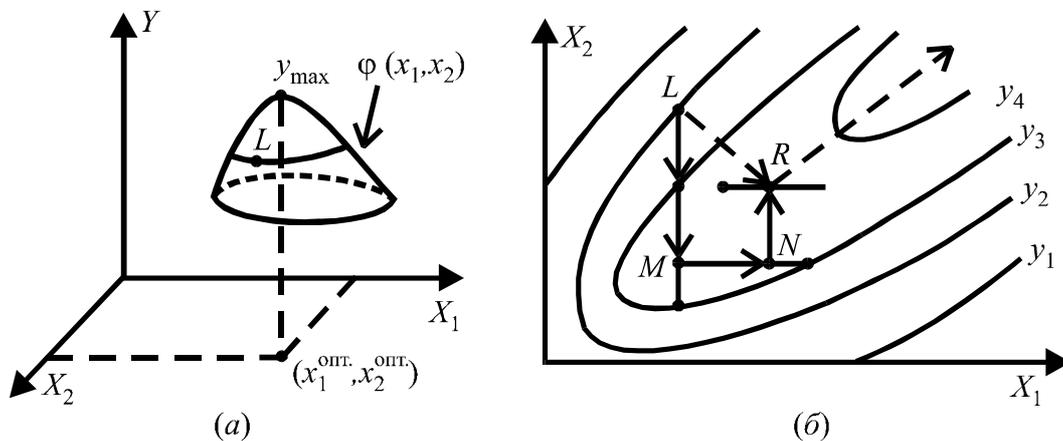


Рис. 3. Движение по поверхности отклика (а) к экстремуму в традиционном эксперименте и в методе крутого восхождения (б)

Поиск экстремальной точки поверхности отклика в традиционном эксперименте проводится следующим образом. В точке L с известным значением y фиксируется один из факторов, например x_1 , и начинается движение из этой точки вдоль оси x_2 . Движение по x_2 продолжается до

тех пор, пока не прекращается прирост y (рис. 3 б). В точке M с наилучшим значением выходного параметра фиксируется фактор x_2 и начинается движение в направлении оси x_1 . В точке N со следующим наилучшим значением y снова фиксируется x_1 и начинается движение по x_2 и т. д. Очевидно, что путь к экстремуму по ломаной кривой $LMNR$ (рис. 3 б) не является оптимальным.

Кратчайшим, наиболее крутым путем достижения экстремума будет движение из точки L по градиенту перпендикулярно изолиниям $y = const$ (на рис. 3 б этот путь показан пунктирной линией). Для рассматриваемого случая градиент функции отклика равен

$$\text{grad } \varphi = \left(\frac{\partial \varphi}{\partial x_1} \right) \vec{i} + \left(\frac{\partial \varphi}{\partial x_2} \right) \vec{j}, \quad (12.1)$$

где \vec{i} и \vec{j} — орты координатных осей. Предполагается, что функция φ непрерывна, дифференцируема и не имеет особых точек.

Для реализации метода крутого восхождения Бокс и Уилсон предложили шаговый метод движения по поверхности отклика. В окрестности точки L ставится эксперимент для локального описания поверхности отклика линейным уравнением регрессии:

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2. \quad (12.2)$$

Движение из точки L начинается в направлении градиента линейного приближения

$$\frac{\partial \hat{y}}{\partial x_1} = b_1; \quad \frac{\partial \hat{y}}{\partial x_2} = b_2. \quad (12.3)$$

Для случая, представленного на рис. 3 б, выборочные коэффициенты при линейных членах в окрестности точки L имеют разные знаки: $b_1 > 0$, $b_2 < 0$, поэтому при движении к максимуму функции отклика значение x_1 увеличивается, а x_2 уменьшается. Движение по градиенту линейного приближения продолжается до тех пор, пока не прекращается прирост y . В точке с наибольшим значением y (центр плана) ставится новая серия опытов и определяется новое направление движения по поверхности отклика. Такой шаговый процесс продолжается до достижения области, близкой к экстремуму.

При постановке опытов величина шага должна быть пропорциональна произведению коэффициента на интервал варьирования: $b_j \Delta z_j$. Например, при движении из точки L следующий эксперимент ставит-

ся в точке со значениями x_1 и x_2 , отличающимися от начальных на величины $2b_1\Delta z_1$ и $2b_2\Delta z_2$ соответственно. В общем случае направление градиента будет зависеть от выбранного интервала варьирования независимых факторов. При изменении в n раз интервала варьирования некоторого j -фактора величина шага для него меняется в n^2 раз, так как при этом в n раз изменяется и коэффициент регрессии b_j . Инвариантными к изменению интервала остаются только знаки составляющих градиента. При увеличении числа рассматриваемых факторов более двух оптимизация методом крутого восхождения по поверхности отклика проводится аналогичным способом.

12.2. Описание функции отклика в области, близкой к экстремуму. Композиционные планы Бокса-Уилсона

В области, близкой к экстремуму, (или «почти стационарной области») функция отклика существенно нелинейна, поэтому для ее адекватного описания необходимо использовать нелинейные полиномы. В настоящее время для этой цели наиболее широко применяют полиномы второго порядка, для получения которых имеются хорошо разработанные планы эксперимента.

Для описания полиномом второго порядка эксперимента, реализованного для нахождения оптимальных условий процесса, число опытов N в плане должно быть не меньше числа определяемых коэффициентов в уравнении регрессии второго порядка для k факторов

$$\hat{y} = b_0 + \sum_{j=1}^k b_j x_j + \sum_{\substack{u,j=1 \\ u \neq j}}^k b_{uj} x_u x_j + \sum_{j=1}^k b_{jj} x_j^2. \quad (12.4)$$

Выборочные коэффициенты (12.4) являются оценками соответствующих коэффициентов уравнения теоретической регрессии:

$$m_y = \beta_0 + \sum_{j=1}^k \beta_j x_j + \sum_{\substack{u,j=1 \\ u \neq j}}^k \beta_{uj} x_u x_j + \sum_{j=1}^k \beta_{jj} x_j^2. \quad (12.5)$$

В зависимости от числа рассматриваемых факторов число коэффициентов l уравнения регрессии (12.4) определяется по формуле

$$l = (k+1) + k + C_k^2 = 2k + 1 + \frac{k!}{2!(k-2)!} = \frac{(k+1)(k+2)}{2}, \quad (12.6)$$

где C_k^2 — количество сочетаний из k факторов по два, равное числу эффектов парного взаимодействия.

В области, близкой к экстремуму, становятся значимыми эффекты парного взаимодействия и квадратичные эффекты. Поэтому то, что адекватное описание результатов эксперимента требует использования полиномов второго порядка, может служить признаком нахождения в почти стационарной области. Близость к этой области можно также установить, поставив дополнительно к ПФЭ 2^k или ДФЭ 2^{k-1} серию опытов в центре плана. Среднее значение результатов этих опытов является оценкой для свободного члена уравнения (12.5):

$$\overline{y^0} \rightarrow \beta_0. \quad (12.7)$$

Выборочный коэффициент b_0 , вычисляемый по формуле

$$b_0 = \frac{1}{N} \sum_{i=1}^N x_{0i} y_i, \quad (12.8)$$

оценивает сумму свободного и квадратичных членов:

$$b_0 \rightarrow \beta_0 + \sum_{j=1}^k \beta_{jj}. \quad (12.9)$$

Поэтому, чем больше разность

$$(b_0 - \overline{y^0}) \rightarrow \sum_{j=1}^k \beta_{jj}, \quad (12.10)$$

тем значимее квадратичные эффекты.

Для описания поверхности отклика полиномами второго порядка независимые факторы в планах должны принимать не менее трех разных значений. Эксперимент, в котором каждый из k факторов рассматривается на трех уровнях и реализуются все возможные сочетания уровней факторов, является ПФЭ типа 3^k . В качестве примера в табл. 15 представлена матрица планирования ПФЭ 3^2 .

Проведение ПФЭ 3^k требует большого числа опытов, намного превышающего число определяемых коэффициентов l в уравнении (12.4) уже при $k > 2$:

k	2	3	4
3^k	9	27	81
l	6	10	15

Таблица 15

Матрица планирования ПФЭ 3^2

№ опыта	x_1	x_2	y
1	-1	-1	y_1
2	0	-1	y_2
3	+1	-1	y_3
4	-1	0	y_4
5	0	0	y_5
6	+1	0	y_6
7	-1	+1	y_7
8	0	+1	y_8
9	+1	+1	y_9

Сократить общее число опытов при условии получения несмешанных оценок для линейных эффектов и эффектов взаимодействия можно с помощью *композиционных планов Бокса-Уилсона*. Ядро таких планов при $k < 5$ составляет ПФЭ 2^k , и полуреплика от него при $k \geq 5$.

Если линейное уравнение регрессии оказалось неадекватным эксперименту, необходимо:

- 1) добавить $2k$ звездных точек, расположенных на координатных осях факторного пространства. Координаты звездных точек в общем случае равны

$$(\pm\alpha, 0, \dots, 0), (0, \pm\alpha, \dots, 0), \dots, (0, \dots, 0, \pm\alpha),$$

где α — расстояние от центра плана до звездной точки, или звездное плечо;

- 2) увеличить число экспериментов в центре плана n_0 .

Общее число опытов в матрице композиционного плана при $k \leq 4$ составляет

$$N = 2^k + 2k + n_0. \quad (12.11)$$

Рассмотрим построение композиционных планов на примере $k = 2$ (рис. 4). Точки 1, 2, 3, 4 образуют ПФЭ 2^2 , точки 5, 6, 7, 8 являются звездными точками с координатами $(\pm\alpha, 0)$ и $(0, \pm\alpha)$, координаты n_0 опытов в центре плана нулевые — $(0, 0)$.

Композиционный план второго порядка для двух факторов представлен в табл. 16, при этом в центре плана выполнена серия из трех опытов (№ 9–11).

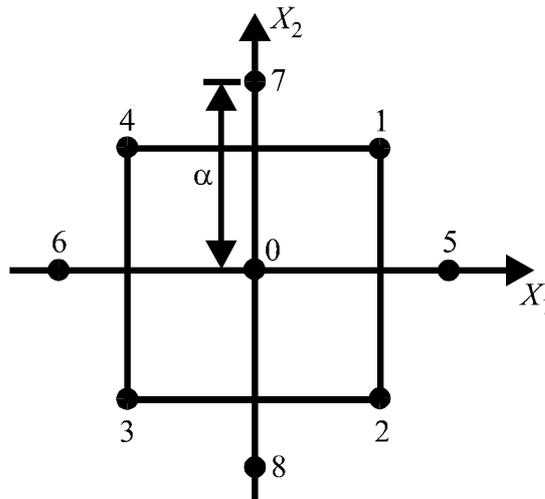


Рис. 4. Композиционный план второго порядка для двух факторов

Таблица 16

Композиционный план второго порядка для двух факторов

№ опыта	x_0	x_1	x_2	x_1x_2	x_1^2	x_2^2	y
1	+1	+1	+1	+1	+1	+1	y_1
2	+1	+1	-1	-1	+1	+1	y_2
3	+1	-1	-1	+1	+1	+1	y_3
4	+1	-1	+1	-1	+1	+1	y_4
5	+1	$+\alpha$	0	0	α^2	0	y_5
6	+1	$-\alpha$	0	0	α^2	0	y_6
7	+1	0	$+\alpha$	0	0	α^2	y_7
8	+1	0	$-\alpha$	0	0	α^2	y_8
9	+1	0	0	0	0	0	y_9
10	+1	0	0	0	0	0	y_{10}
11	+1	0	0	0	0	0	y_{11}

12.3. Ортогональные планы второго порядка, расчет коэффициентов уравнения регрессии

Выбор звездного плеча в композиционных планах Бокса–Уилсона может быть произвольным, однако расчеты коэффициентов уравнения регрессии при $k < 5$ существенно упрощаются, если величина плеча определяется исходя из следующего уравнения:

$$\alpha^4 + 2k\alpha^2 - 2^{k-1}(k + 0.5n_0) = 0. \quad (12.12)$$

Значения α^2 , определенные по (12.12), приведены в табл. 17.

Таблица 17

Значения α^2 для k факторов и n_0 опытов в центре плана

n_0	k			n_0	k		
	2	3	4		2	3	4
1	1.00	1.476	2.00	6	1.742	2.325	2.950
2	1.160	1.650	2.164	7	1.873	2.481	3.140
3	1.317	1.831	2.390	8	2.00	2.633	3.310
4	1.475	2.00	2.580	9	2.113	2.782	3.490
5	1.606	2.164	2.770	10	2.243	2.928	3.66

Выбрав α , проведем следующее линейное преобразование квадратичных столбцов:

$$x'_j = x_j^2 - \overline{x_j^2} = x_j^2 - \frac{1}{N} \sum_{i=1}^N x_{ji}^2. \quad (12.13)$$

Композиционные планы, полученные таким образом, называются *ортогональными планами второго порядка*.

Ортогональный план второго порядка при $k=2$ и $n_0=1$ представлен в табл. 18. За его основу взят композиционный план для двух факторов (табл. 16) с общим числом опытов $N=9$. Величину звездного плеча определим по табл. 17: $\alpha^2=1$, $\alpha=1$. Средние значения элементов квадратичных столбцов в табл. 16 равны

$$\overline{x_1^2} = \frac{1}{9} \sum_{i=1}^9 x_{1i}^2 = \overline{x_2^2} = \frac{1}{9} \sum_{i=1}^9 x_{2i}^2 = \frac{1}{9} (4 + 2\alpha^2) = \frac{2}{3}. \quad (12.14)$$

В математической статистике доказывается, что для ортогональных планов второго порядка все коэффициенты уравнения регрессии определяются независимо друг от друга по формуле

$$b_j = \frac{\sum_{i=1}^N x_{ji} y_i}{\sum_{i=1}^N x_{ji}^2}, \quad (12.15)$$

а дисперсии коэффициентов равны

$$s^2(b_j) = s_{\text{воспр}}^2 / \sum_{i=1}^N x_{ji}^2. \quad (12.16)$$

Таблица 18

Ортогональный план второго порядка для двух факторов

№ опыта	x_0	x_1	x_2	x_1x_2	x_1'	x_2'	y
1	+1	+1	+1	+1	+1/3	+1/3	y_1
2	+1	+1	-1	-1	+1/3	+1/3	y_2
3	+1	-1	-1	+1	+1/3	+1/3	y_3
4	+1	-1	+1	-1	+1/3	+1/3	y_4
5	+1	$+\alpha$	0	0	+1/3	-2/3	y_5
6	+1	$-\alpha$	0	0	+1/3	-2/3	y_6
7	+1	0	$+\alpha$	0	-2/3	+1/3	y_7
8	+1	0	$-\alpha$	0	-2/3	+1/3	y_8
9	+1	0	0	0	-2/3	-2/3	y_9

Для определения дисперсии воспроизводимости необходимо выполнить серию опытов в центре плана. В результате расчетов по матрице с преобразованными столбцами для квадратичных эффектов (табл. 18) получаем следующее уравнение:

$$\hat{y} = b_0' + b_1x_1 + b_2x_2 + b_{12}x_1x_2 + b_{11}(x_1^2 - \overline{x_1^2}) + b_{22}(x_2^2 - \overline{x_2^2}). \quad (12.17)$$

Чтобы перейти к обычной записи уравнения регрессии в виде

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + b_{12}x_1x_2 + b_{11}x_1^2 + b_{22}x_2^2, \quad (12.18)$$

определим b_0 по формуле

$$b_0 = b_0' - b_{11}\overline{x_1^2} - b_{22}\overline{x_2^2} \quad (12.19)$$

с дисперсией, равной

$$s^2(b_0) = s^2(b_0') + \left(\overline{x_1^2}\right)^2 s^2(b_{11}) + \left(\overline{x_2^2}\right)^2 s^2(b_{22}). \quad (12.20)$$

Значимость коэффициентов проверяется по критерию Стьюдента. Если

$$t_j = \frac{|b_j|}{s(b_j)} > t_p(f), \quad (12.21)$$

где $t_p(f)$ — табличное значение критерия Стьюдента для $p = 0,05$ и числа степеней свободы дисперсии воспроизводимости, то коэффициент значимо отличается от нуля.

Коэффициенты уравнения регрессии, получаемые при помощи ортогональных планов второго порядка, определяются с разной точностью. В случае, когда $k \leq 4$, согласно (12.16) имеем

$$s^2(b'_0) = \frac{s_{\text{воспр}}^2}{N};$$

$$s^2(b_j) = \frac{s_{\text{воспр}}^2}{2^k + 2\alpha^2}, j = 1, 2, \dots, k;$$

$$s^2(b_{uj}) = \frac{s_{\text{воспр}}^2}{2^k}, u, j = 1, 2, \dots, k, u \neq j; \quad (12.22)$$

$$s^2(b_{jj}) = \frac{s_{\text{воспр}}^2}{2^k (1 - x_j^2)^2 + 2(\alpha^2 - x_j^2)^2 + (n_0 + 2k - 2)(x_j^2)^2}, j = 1, 2, \dots, k.$$

После исключения незначимых коэффициентов проводится проверка адекватности уравнения по критерию Фишера. Уравнение адекватно эксперименту, если

$$F = \frac{s_{\text{ад}}^2}{s_{\text{воспр}}^2} \leq F_{1-p}(f_{\text{ад}}, f_{\text{воспр}}), \quad (12.23)$$

где $F_{1-p}(f_{\text{ад}}, f_{\text{воспр}})$ — критерий Фишера для $p = 0,05$; $f_{\text{ад}} = N - l$ — число степеней свободы дисперсии адекватности (l — число значимых коэффициентов в уравнении регрессии); $f_{\text{воспр}}$ — число степеней свободы дисперсии воспроизводимости.

12.4. Метод последовательного симплекс-планирования

В рассмотренных выше планах ПФЭ 2^2 и 2^3 экспериментальные точки располагались в вершинах квадрата и куба соответственно. В качестве экспериментального плана можно также использовать *регулярный симплекс*. Симплексом в k -мерном пространстве называют выпуклый многогранник, имеющий ровно $(k + 1)$ вершину, каждая из которых определяется пересечением k гиперплоскостей данного пространства. Симплекс называется регулярным, если расстояния между всеми его вершинами равны. Примерами регулярных симплексов являются правильный треугольник в двумерном пространстве и тетраэдр в трехмерном.

На практике планирование эксперимента с использованием регулярных симплексов применяется для решения задач оптимизации при

движении к почти стационарной области. Для получения регулярного симплекса проводится линейное преобразование уровней факторов

$$x_j = \frac{z_j - z_j^0}{\Delta z_j}, \quad (12.24)$$

где z_j^0 — j -я координата центра плана; Δz_j — интервал варьирования по j -фактору.

Оптимизация методом последовательного симплекс-планирования проводится следующим образом: исходная серия опытов планируется так, чтобы экспериментальные точки образовывали регулярный симплекс в факторном пространстве. После проведения опытов определяется вершина симплекса, соответствующая наихудшим результатам. Далее строится новый симплекс, для чего наихудшая точка исходного симплекса заменяется новой, расположенной симметрично относительно центра грани симплекса, находящейся против наихудшей точки. Новая точка вместе с оставшимися точками образует новый симплекс, центр тяжести которого смещен в сторону повышения качества процесса. После реализации опыта в дополнительной точке опять проводится выявление наихудшей вершины симплекса и т. д. При достижении области оптимума, симплекс начинает вращение вокруг вершины с максимальным значением отклика.

На рис. 5 показаны схемы достижения экстремума поверхности отклика методами крутого восхождения и симплекс-планирования на примере зависимости целевой функции y от двух факторов. При оптимизации методом крутого восхождения (рис. 5 а) в окрестности точки M поставлен ПФЭ 2^2 , движение по градиенту линейного приближения осуществлялось в опытах 5–9. Далее был поставлен новый ПФЭ 2^2 (точки 10–13) с центром в точке 7, в которой было получено наилучшее значение y . Движение по новому градиенту (точки 14–15) приводит к экстремуму.

При оптимизации методом симплекс-планирования (рис. 5 б) в исходном симплексе (точки 1–3) худшей точкой оказалась точка 2. Ее зеркальным отражением относительно c_1 — центра грани 1–3 — является точка 4. В новом симплексе 1, 3, 4 худшей оказалась точка 1, в результате ее зеркального отражения получен симплекс 3, 4, 5 и т. д. Область оптимума достигается при реализации симплекса 9, 10, 11.

Хотя оба рассмотренных метода требуют проведения примерно одинакового числа опытов, симплекс-планирование имеет ряд важных

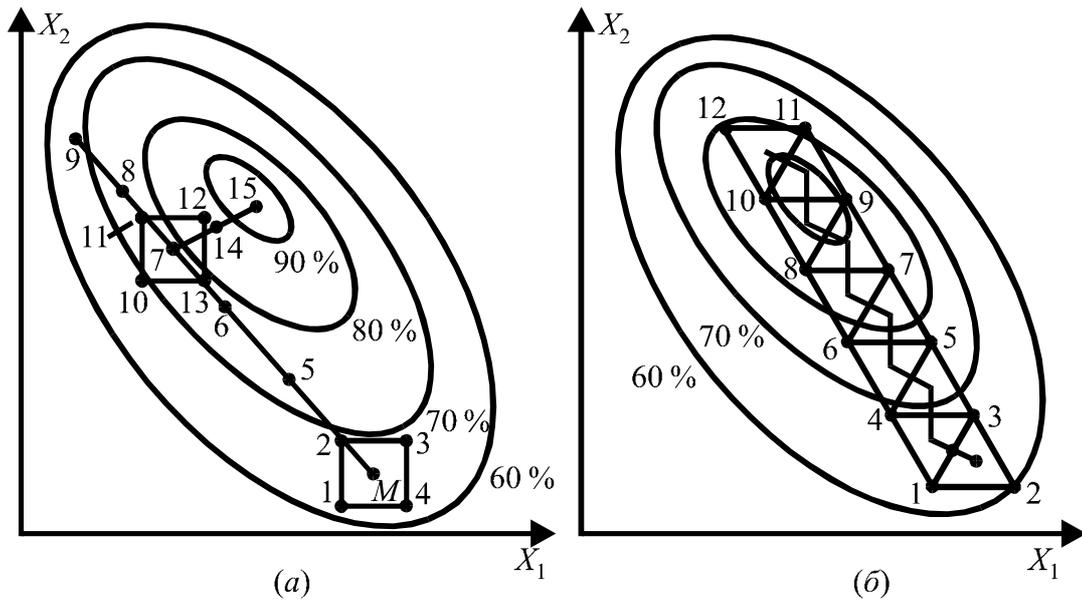


Рис. 5. Достижение экстремума поверхности отклика методами кругового восхождения (а) и симплекс-планирования (б)

преимуществ: при использовании этого метода параметр оптимизации y может измеряться приближенно, достаточно иметь возможность проранжировать его величину; можно одновременно учитывать несколько параметров оптимизации; метод не предъявляет жестких требований к локальной аппроксимации поверхности отклика уравнением регрессии.

На практике рекомендуется ориентировать исходный симплекс в факторном пространстве следующим образом: центр симплекса совпадает с началом координат, одна из вершин лежит на координатной оси, а остальные располагаются симметрично относительно координатных осей, плоскостей и гиперплоскостей (в многомерном случае). Тогда координаты вершин симплекса при $k = 5$ задаются матрицей

$$X = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 & x_5 \\ -x_1 & x_2 & x_3 & x_4 & x_5 \\ 0 & -2x_2 & x_3 & x_4 & x_5 \\ 0 & 0 & -3x_3 & x_4 & x_5 \\ 0 & 0 & 0 & -4x_4 & x_5 \\ 0 & 0 & 0 & 0 & -5x_5 \end{bmatrix}. \quad (12.25)$$

При $k < 5$ координаты вершин симплекса определяются частью матрицы (12.25), при этом число столбцов равно числу факторов, а число

строк равно $k + 1$; при $k > 5$ для каждого добавленного фактора в матрицу (12.25) добавляются соответствующие столбец и строка. В общем случае число опытов в симплексной матрице для k независимых факторов $N = (k + 1)$ равно числу коэффициентов линейного уравнения регрессии, т. е. симплексные планы являются насыщенными.

Если длину стороны симплекса принять равной 1, то

$$x_j = \sqrt{\frac{1}{2j(j+1)}}. \quad (12.26)$$

Для практического использования матрицы (12.25) ее числовые элементы заранее подсчитаны по формуле (12.26)

$$X = \begin{bmatrix} 0.5 & 0.289 & 0.204 & 0.158 & 0.129 \\ -0.5 & 0.289 & 0.204 & 0.158 & 0.129 \\ 0 & -0.578 & 0.204 & 0.158 & 0.129 \\ 0 & 0 & -0.612 & 0.158 & 0.129 \\ 0 & 0 & 0 & -0.632 & 0.129 \\ 0 & 0 & 0 & 0 & -0.645 \end{bmatrix}. \quad (12.27)$$

План эксперимента в безразмерном масштабе для k факторов состоит из k столбцов и $k + 1$ строки матрицы (12.27). Коэффициенты уравнения линейной регрессии

$$\hat{y} = b_0 + \sum_{j=1}^k b_j x_j \quad (12.28)$$

вычисляются следующим образом:

$$b_0 = \frac{1}{N} \sum_{i=1}^N y_i, \quad b_j = 2 \sum_{i=1}^N x_{ji} y_i. \quad (12.29)$$

Если в одной из вершин симплекса поставить серию параллельных опытов и рассчитать дисперсию воспроизводимости, то выборочные дисперсии коэффициентов определяются по формуле

$$s^2(b_j) = \frac{s_{\text{воспр}}^2}{\sum_{i=1}^N x_{ji}^2} = 2s_{\text{воспр}}^2. \quad (12.30)$$

Следует отметить, что коэффициенты уравнения регрессии, полученные по симплексному плану, определяются с меньшей точностью по сравнению с коэффициентами, полученными при реализации ПФЭ 2^k и ДФЭ 2^{k-1} , для которых

$$s^2(b_j) = s_{\text{воспр}}^2 / N. \quad (12.31)$$

После реализации исходного симплекса требуется провести отражение наихудшей точки относительно центра противоположной грани. Координаты отраженной точки равны:

$$x_j^{(k+2)} = 2x_j^{(c)} - x_j^{(l)}, \quad j = 1, 2, \dots, k, \quad (12.32)$$

где $x_j^{(l)}$ — j -я координата наихудшей точки; $x_j^{(k+2)}$ — j -я координата новой точки, получаемой в результате отражения; $x_j^{(c)}$ — j -я координата центра противоположной грани, определяемая по формуле

$$x_j^{(c)} = \frac{1}{k} \sum_{i=1}^{k+1} x_j^{(i)}, \quad i \neq l, \quad (12.33)$$

где $x_j^{(i)}$ — j -я координата i -й вершины симплекса ($i = 1, 2, \dots, k+1$).

Координаты центра оптимального симплекса (точка S) в почти стационарной области находятся следующим образом:

$$x_j^{(S)} = \frac{1}{(k+1)} \sum_{i=1}^{k+1} x_j^{(i)}. \quad (12.34)$$

СОДЕРЖАНИЕ

ЛЕКЦИЯ 7	3
7.1. Системы случайных величин. Функция и плотность распределения системы двух случайных величин. Условные законы распределения.	3
7.2. Стохастическая связь. Ковариация. Коэффициент корреляции. Регрессия.	5
7.3. Выборочный коэффициент корреляции. Проверка гипотезы об отсутствии корреляции.	9
7.4. Приближенная регрессия. Метод наименьших квадратов.	12
ЛЕКЦИЯ 8	15
8.1. Линейная регрессия от одного параметра.	15
8.2. Регрессионный анализ.	16
8.2.1. Проверка адекватности приближенного уравнения регрессии эксперименту.	17
8.2.2. Оценка значимости коэффициентов уравнения регрессии.	18
8.2.3. Оценка доверительного интервала для искомой функции.	20
8.3. Оценка тесноты нелинейной связи.	21
8.4. Аппроксимация. Параболическая регрессия.	22
8.5. Приведение некоторых функциональных зависимостей к линейному виду.	23
8.6. Метод множественной корреляции.	25
ЛЕКЦИЯ 9	27
9.1. Задачи дисперсионного анализа. Однофакторный дисперсионный анализ.	27
9.2. Двухфакторный дисперсионный анализ.	31
ЛЕКЦИЯ 10	36
10.1. Планирование эксперимента при дисперсионном анализе.	36
10.2. Постановка задачи при планировании экстремальных экспериментов.	39

10.3. Полный факторный эксперимент типа 2^2 : матрица планирования, вычисление коэффициентов уравнения регрессии.	41
ЛЕКЦИЯ 11	44
11.1. Матрица планирования полного факторного эксперимента типа 2^3	44
11.2. Проверка значимости коэффициентов и адекватности уравнения регрессии, полученных при обработке результатов ПФЭ 2^2 и 2^3	46
11.3. Дробный факторный эксперимент. Планы типа 2^{k-1}	49
ЛЕКЦИЯ 12	53
12.1. Оптимизация методом крутого восхождения по поверхности отклика.	53
12.2. Описание функции отклика в области, близкой к экстремуму. Композиционные планы Бокса-Уилсона.	55
12.3. Ортогональные планы второго порядка, расчет коэффициентов уравнения регрессии.	58
12.4. Метод последовательного симплекс-планирования.	61