

МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ ФЕДЕРАЦИИ
ДАЛЬНЕВОСТОЧНЫЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ УНИВЕРСИТЕТ
ТЕХНОЛОГИЧЕСКИЙ ФАКУЛЬТЕТ

А.М. Емельянов, Н.П. Кидяева,
Е.А. Подолько, Е.М. Шпилёв

СТАТИСТИЧЕСКИЕ МЕТОДЫ ОБРАБОТКИ,
ПЛАНИРОВАНИЯ ИНЖЕНЕРНОГО ЭКСПЕРИМЕНТА

*Учебное пособие
для магистрантов и аспирантов*

Благовещенск
Издательство ДальГАУ
2015

УДК 519.24

Статистические методы обработки, планирования инженерного эксперимента: учебное пособие / сост. А.М. Емельянов, Н.П. Кидяева, Е.А. Подолько, Е.М. Шпилев. – Благовещенск: ДальГАУ, 2015. – 93 с.

В работе последовательно излагаются статистические методы обработки экспериментальных данных: построение гистограмм; точечные и интервальные оценки; критерии Барлетта и Кохрена; метод регрессионного анализа; изложены элементы планирования, постановка факторного эксперимента.

В приложении приведены необходимые статистические таблицы для вычисления квантилей и других статистических констант.

Пособие рассчитано на магистрантов и аспирантов инженерных специальностей.

Рецензент – З.Ф. Кривуца, канд.физ.-мат.наук, доцент

Рекомендовано к печати методическим советом технологического факультета Дальневосточного государственного аграрного университета (Протокол №3 от 19 ноября 2014 года).

Издательство ДальГАУ

2015

ВВЕДЕНИЕ

Вопросам математической обработки результатов инженерного эксперимента посвящено большое количество технической литературы. К сожалению, в литературе рассматриваемые вопросы изложены разрозненно и не всегда в виде, удобном для использования. В настоящих методических указаниях предпринята попытка в систематизированном виде представить один из методов математической обработки данных.

Основной задачей обработки экспериментального материала является выделение из него полезной информации и представление ее в виде, удобном для анализа. При этом информация преобразовывается так, чтобы отдельные стороны явления или процесса выделялись четко и ярко, а полученные результаты можно было бы оценить количественными показателями. Такой подход обусловлен непрерывно повышающейся требовательностью к достоверности результатов исследований.

Пусть в результате однофакторного эксперимента получены измерения $x_1, x_2, \dots, x_n$ исследуемого параметра при нескольких постоянных значениях независимой переменной величины. Для увеличения точности опытных данных при каждом постоянном значении независимой переменной проводится ряд параллельных опытов. Полученные экспериментальные данные представляют собою дискретные случайные величины. Целью математической обработки данных является получение эмпирической зависимости $y = f(x)$. Оценка коэффициентов данного уравнения, проверка гипотезы о значимости коэффициентов, адекватности модели (уравнения) является предметом *регрессионного анализа*, а полученное уравнение $y = f(x)$ – уравнением регрессии.

Дальнейшее развитие теория приведенных экспериментальных исследований получила после того, как совместно с регрессионным анализом стали использовать различные методы планирования эксперимента. При постановке факторных экспериментов обработка их результатов основывается на законах математической статистики.

1 ОСНОВНЫЕ ПОНЯТИЯ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

1.1 Статистическая функция распределения. Гистограмма

Исходным материалом при решении задач математической статистики является последовательность независимых наблюдений случайной величины X распределения $F(x)$. Наблюдаемые значения случайной величины $x_1, x_2, \dots, x_n$ называются выборкой из генеральной совокупности X с функцией распределения $F(x)$, количество наблюдений – объемом выборки. С точки зрения теории вероятностей выборка $\{x_1, x_2, \dots, x_n\}$ представляет собой последовательность независимых одинаково распределенных случайных величин.

Предположим, что изучается некоторая случайная величина X , закон распределения которой точно неизвестен, необходимо определить этот закон или проверить экспериментально гипотезу о том, что величина X распределена по тому или иному закону. С этой целью над случайной величиной X производится ряд независимых опытов. Полученная совокупность данных называется простой статистической совокупностью. Данный материал может быть обработан различными способами. Одним из способов такой обработки является построение статистической функции распределения случайной величины.

Статистической функцией распределения случайной величины X называется частота события $X < x$ в данном статистическом материале

$$F^*(x) = P^*(X < x).$$

График статистической функции распределения представляет собой ступенчатую линию, скачки которой соответствуют наблюдаемым значениям случайной величины и по величине равны частотам этих значений.

Для придания компактности и наглядности статистическому материалу производится дополнительная обработка – строится статистический ряд. Весь диапазон наблюдаемых значений X делится на интервалы или «разряды», подсчитывается количество значений m_i , приходящееся на каждый i – й

разряд. Это число делится на общее число наблюдений n – находится относительная частота, соответствующая данному разряду

$$P_i^* = \frac{m_i}{n}$$

Сумма частот всех разрядов равна единицы.

Полученные данные сводятся в таблицу, в которой приведены разряды в порядке их расположения вдоль оси абсцисс и соответствующие частоты. Эта таблица называется статистическим рядом.

Статистический ряд часто оформляется в виде гистограммы. Гистограмма строится в следующем порядке. По оси абсцисс откладываются разряды, и на каждом из разрядов как на основании строится прямоугольник, площадь которого равна частоте данного разряда. Для построения гистограммы нужно частоту каждого разряда разделить на его длину и полученное число взять в качестве высоты прямоугольника.

Если середины вершин прямоугольника гистограммы соединить ломаной линией, то получим полигон частот, также являющийся одним из способов графического представления эмпирических представлений.

Рассмотрим выше приведенные понятия на примере.

Пример 1.1. Произведено 500 измерений диаметра гильзы плунжерной пары топливного насоса. Результаты отклонений диаметра от номинального сведены в статистический ряд.

Таблица 1.1

Статистический ряд результатов отклонения диаметра гильзы

Y_i	-4; -3	-3; -2	-2; -1	-1; 0	0; 1	1; 2	2; 3	3; 4
m_i	6	25	72	133	120	88	46	10
P_i^*	0,012	0,05	0,144	0,266	0,240	0,176	0,092	0,020

Y_i – обозначены интервалы (разряды) отклонения диаметра гильзы, m_i – число наблюдений в данном интервале, $P_i^* = m_i/n$ – соответствующие частоты.

Гистограмма для отклонения диаметра гильзы, построенная по данным статистического ряда, приведена на рисунке 1.1.

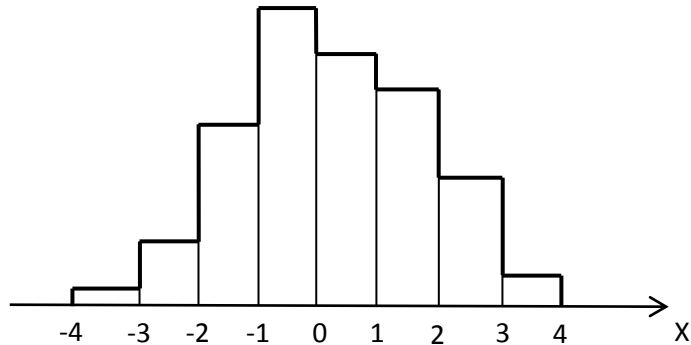


Рис.1.1.1. Гистограмма для отклонения диаметра гильзы

Пользуясь данными статистического ряда, можно приближенно построить и статистическую функцию распределения величины X . Для практики обычно достаточно построить статистическую функцию распределения по нескольким точкам. В качестве этих точек возьмем границы $x_1, x_2, \dots, x_n$, разрядов статистического ряда. Тогда:

$$F^*(x_1) = F^*(-4) = 0;$$

$$F^*(x_2) = P_1^* = F^*(-3) = 0,012;$$

$$F^*(x_3) = P_1^* + P_2^* = F^*(-2) = 0,012 + 0,050 = 0,062;$$

$$F^*(x_4) = \sum_{i=1}^3 P_i^* = F^*(-1) = 0,026;$$

$$F^*(x_5) = \sum_{i=1}^4 P_i^* = F^*(0) = 0,472;$$

$$F^*(x_6) = \sum_{i=1}^5 P_i^* = F^*(1) = 0,712;$$

$$F^*(x_7) = \sum_{i=1}^6 P_i^* = F^*(2) = 0,888;$$

$$F^*(x_8) = \sum_{i=1}^7 P_i^* = F^*(3) = 0,980;$$

$$F^*(x_9) = \sum_{i=1}^8 P_i^* = F^*(4) = 1,000$$

Статистическая функция распределения изображена на рисунке 1.2.

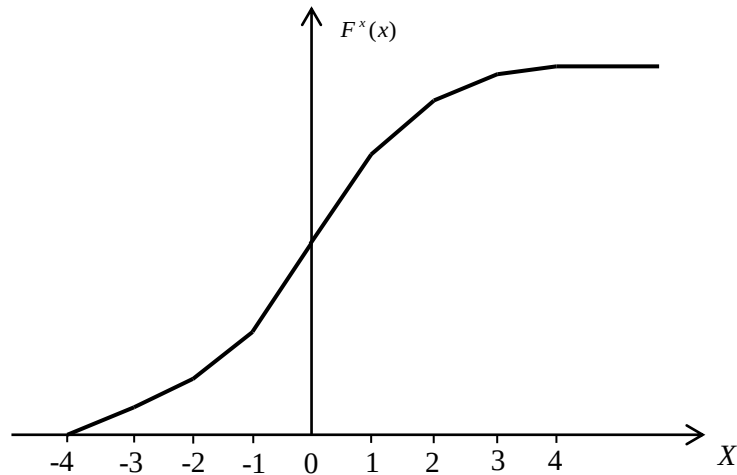


Рис. 1.2 Статистическая функция распределения

1.2 Числовые характеристики статистического распределения

В теории вероятности рассматриваются различные числовые характеристики случайных величин, в частности, математическое ожидание, дисперсия. Аналогичные числовые характеристики существуют и для статистических распределений. Для основной характеристики положения математического ожидания случайной величины такой аналогией является среднее арифметическое наблюдаемых значений случайной величины или выборочное среднее

$$M^*[X] = \frac{\sum_{i=1}^n x_i}{n} = m_x, \quad (1.1)$$

где x_i – значение случайной величины, наблюдаемое в i -ом опыте, n – число опытов.

Получим выборочную дисперсию – характеристику, аналогичную дисперсии случайной величины, которая представляет собой математическое ожидание случайной величины $\overset{\circ}{X}^2 = (x - m_x)^2$

$$\sigma^2 = M\left[\overset{\circ}{X}\right] = M\left[(x - m_x)^2\right], \quad (1.2)$$

где m_x – математическое ожидание случайной величины X .

Если в этом выражении заменить математическое ожидание средним арифметическим, то получим выборочную дисперсию случайной величины

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - m_x)^2}{n}. \quad (1.3)$$

Среднее квадратическое отклонение определяется как положительный корень из дисперсии:

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - m_x)^2}{n}}. \quad (1.4)$$

Среднее квадратическое отклонение характеризует рассеяние (разброс) данных относительно математического ожидания.

1.3 Оценки для математического ожидания и дисперсии (точечная оценка параметров распределения)

Любое значение некоторого параметра, вычисленное на основе ограниченного числа опытов, всегда будет содержать элемент случайности. Такое приближенное, случайное значение называется оценкой параметра.

Рассмотрим общую задачу. Имеется случайная величина X , закон распределения которой содержит неизвестный параметр a . Требуется найти подходящую оценку для параметра a по результатам n независимых опытов, в каждом из которых величина X приняла определенное значение. Обозначим через $\bar{a}$ оценку для параметра a . К оценке $\bar{a}$ предъявляется ряд требований:

1. Оценка должна быть **состоятельной**, т.е. при увеличении числа опытов n она приближается (сходится по вероятности) к параметру a .
2. Оценка должна быть **несмещенной**, т.е. пользуясь величиной $\bar{a}$ вместо a не должно быть систематической ошибки в стороны завышения или занижения

$$M[\bar{a}] = a.$$

3. Оценка должна быть *эффективной*, т.е. обладать по сравнению с другими наименьшей дисперсией

$$\sigma^2[\bar{a}] = \min$$

Пусть имеется случайная величина X с математическим ожиданием m и дисперсией σ^2 , оба параметра неизвестны. Над величиной X произведено n независимых опытов, давших результаты: $x_1, x_2, \dots, x_n$.

Оценкой для математического ожидания, удовлетворяющей выше сформулированным требованиям, является выборочное среднее

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} \quad (1.5)$$

Оценкой для дисперсии является исправленная выборочная дисперсия:

$$S^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} \quad (1.6)$$

Пример 1.2. Для определения точности измерительного прибора, систематическая ошибка которого практически равна нулю, было произведено пять независимых измерений, результаты которых представлены в таблице 1.2.

Таблица 1.2

n	1	2	3	4	5
$x_i, \text{ мм}$	2781	2836	2807	2763	2858

Определить несмещенную оценку математического ожидания и дисперсию ошибок измерения прибора.

Искомые оценки определяются по формулам (1.5) и (1.6)

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{1}{5} (2781 + 2836 + 2807 + 2763 + 2858) = 2809$$

$$S^2 = \frac{1}{5-1} [(2781-2809)^2 + (2836-2809)^2 + (2807-2809)^2 +$$

$$+ (2763 - 2809)^2 + (2858 - 2809)^2 \Big] = 1508,5 \text{ мм}^2$$

1.4 Основные законы распределения

1.4.1 Нормальный закон распределения и его параметры

Изучение различных явлений показывает, что многие случайные величины: ошибки измерений, величины износа деталей, ошибки стрельбы, и т.д. могут быть описаны нормальным законом распределения. Нормальный закон распределения занимает особое место среди других законов распределения. Главная особенность, выделяющая нормальный закон среди других заключается в том, что он является предельным законом, к которому, приближаются другие законы при неограниченном увеличении числа опытов.

Нормальным называют распределение вероятностей непрерывной случайной величины, которое описывается плотностью:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-m)^2}{2\sigma^2}}, \quad (1.7)$$

где m – математическое ожидание, σ – среднее квадратическое отклонение.

График плотности нормального распределения имеет вид:

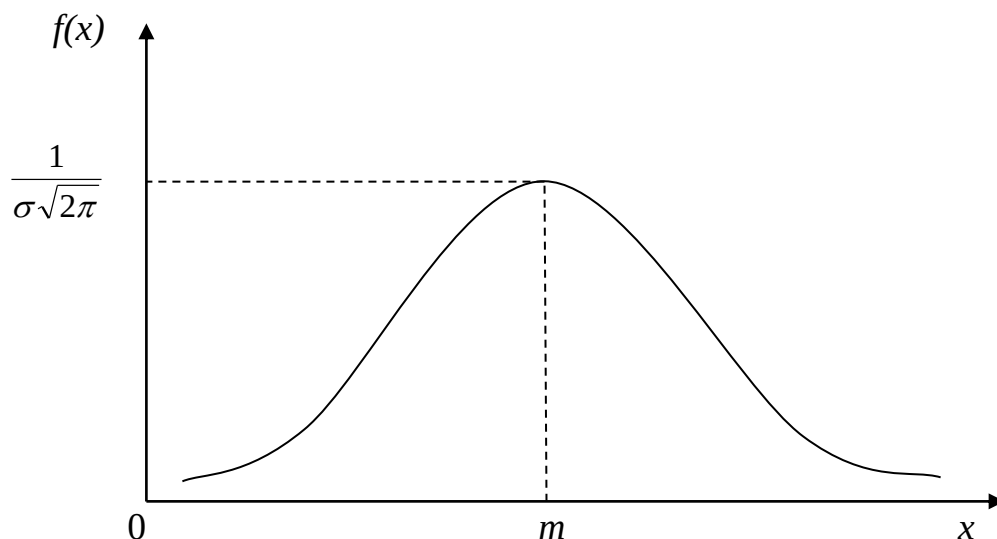


Рис.1.3 График плотности нормального распределения

График плотности нормального распределения называют нормальной кривой (кривой Гаусса). Кривая распределения по нормальному закону имеет

симметричный вид. Максимальная ордината кривой равна $\frac{1}{\sigma\sqrt{2\pi}}$, соответствует точке $x = m$, по мере удаления от точки m плотность распределения уменьшается, при $x \rightarrow \pm\infty$ правая асимптотически приближается к оси абсцисс.

Рассмотрим смысл параметров m и σ , характеризующих нормальное распределение.

Точка $x = m$ является центром рассеивания. Если изменять центр рассеивания m , кривая распределения смещается вдоль оси абсцисс, не изменяя своей формы (рисунок 1.4). Центр рассеивания характеризует положение распределения по оси абсцисс.

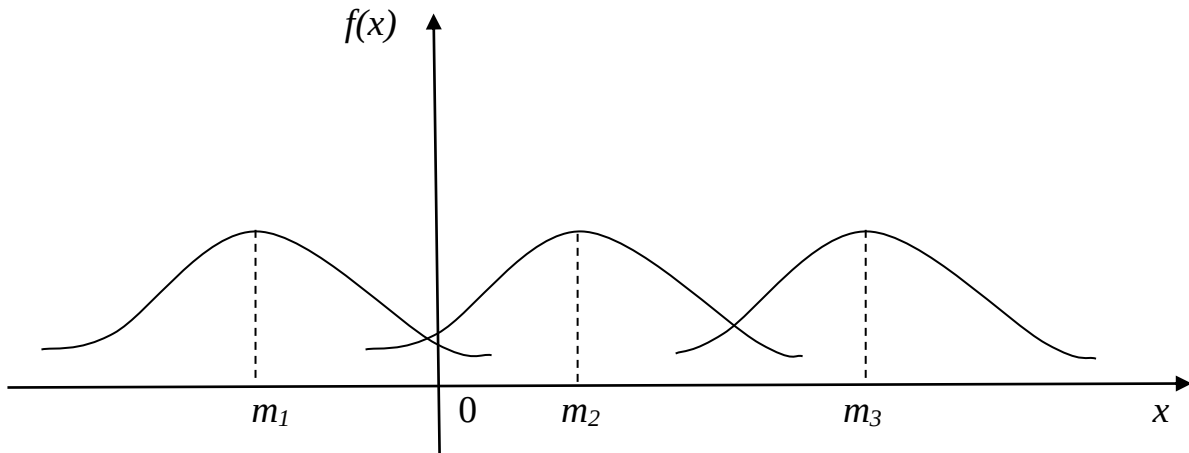


Рис. 1.4

Среднее квадратическое отклонение определяет форму кривой нормального распределения. Максимальная ордината кривой распределения обратно пропорциональна σ . При увеличении σ ордината уменьшается. Поскольку площадь кривой распределения равна единице, то при увеличении σ кривая распределения становится более плоской, при уменьшении σ кривая распределения вытягивается вверх. На рисунке 1.5 изображены три нормальные кривые при $m = 0$.

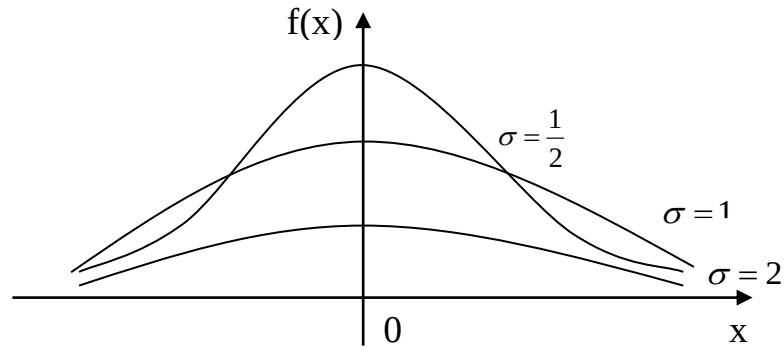


Рис. 1.5

1.4.2 Вероятность попадания случайной величины в заданный интервал. Функция Лапласа

Вероятность попадания случайной величины в заданный интервал от α до β вычисляется по формуле

$$P(\alpha < x < \beta) = F(\beta) - F(\alpha), \quad (1.8)$$

где $F(x)$ – функция распределения случайной величины.

Найдем функцию распределения $F(x)$ случайной величины X . Плотность распределения случайной величины X

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-m)^2}{2\sigma^2}}.$$

Отсюда находим функцию распределения

$$F(x) = \int_{-\infty}^x f(x) dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(x-m)^2}{2\sigma^2}} dx. \quad (1.9)$$

Сделав в данном интеграле замену переменной $\frac{x-m}{\sigma} = t$,

получим

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-m}{\sigma}} e^{-\frac{t^2}{2}} dt \quad (1.10)$$

Полученный интеграл не выражается через элементарные функции. Значения этого интеграла выражаются через функцию Лапласа

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt. \quad (1.11)$$

Данная функция представляет собой функцию нормального распределения с параметрами $m = 0$, $\sigma = 1$.

Функция распределения $F(x)$ случайной величины через функцию Лапласа выражается следующим образом

$$F(x) = \Phi\left(\frac{x - m}{\sigma}\right). \quad (1.12)$$

Искомая вероятность попадания случайной величины X в заданный интервал определяется по выражению:

$$P(\alpha < x < \beta) = \Phi\left(\frac{\beta - m}{\sigma}\right) - \Phi\left(\frac{\alpha - m}{\sigma}\right). \quad (1.13)$$

Функция Лапласа обладает следующими свойствами:

1. Область определения функции – вся числовая ось;
2. $\Phi(-\infty) = 0$;
3. $\Phi(+\infty) = 1$;
4. $\Phi(x)$ – функция неубывающая
5. Функция Лапласа нечетная $\Phi(-x) = -\Phi(x)$.

Значения функции Лапласа имеются в таблицах в учебниках по теории вероятностей (приложение 1).

Пример 1.3. Случайная величина X распределена по нормальному закону. Математическое ожидание равно 30, среднее квадратическое отклонение равно 10. Найти вероятность того, что X примет значение, принадлежащее интервалу (10;50).

Решение.

$$P(10 < x < 50) = \Phi\left(\frac{50 - 30}{10}\right) - \Phi\left(\frac{10 - 30}{10}\right) = \Phi(2) - \Phi(-2) = 2\Phi(2) = 2 \cdot 0,4772 = 0,9544.$$

Пример 1.4. Случайная величина X имеет нормальное распределение. Математическое ожидание $m = 0,5$, дисперсия $\sigma^2 = \frac{1}{8}$. Найти вероятность попадания случайной величины в интервал $(0,4; 0,6)$.

Решение.

$$P(0,4 < x < 0,6) = \Phi\left(\frac{0,6 - 0,5}{\sqrt{\frac{1}{8}}}\right) - \Phi\left(\frac{0,4 - 0,5}{\sqrt{\frac{1}{8}}}\right) = \Phi(0,2828) - \Phi(-0,2828) =$$

$$= 2\Phi(0,2828) = 2 \cdot 0,11031 = 0,22061.$$

Пример 1.5. На токарном станке изготавливают деталь. Длина детали представляет собой случайную величину, распределенную по нормальному закону. Математическое ожидание $m = 10$, дисперсия $\sigma^2 = \frac{1}{2000}$. Найти вероятность брака, если допустимые размеры детали должны быть $10 \pm 0,05$.

Решение.

По условию задачи $m = 10$, $\sigma^2 = \frac{1}{2000}$, $\alpha = 9,95$; $\beta = 10,05$.

Вероятность брака равна

$$P_{\text{бп}} = 1 - P(9,95 < x < 10,05) = 1 - \left[\Phi\left(\frac{10,05 - 10}{\frac{1}{\sqrt{2000}}}\right) - \Phi\left(\frac{9,95 - 10}{\frac{1}{\sqrt{2000}}}\right) \right] = 1 - [\Phi(2,2361) - \Phi(-2,2361)] =$$

$$= 1 - 2\Phi(2,2361) = 1 - 2 \cdot 0,4875 = 0,025.$$

При решении практических задач часто приходится решать вопрос вычисления вероятного попадания нормально распределённой случайной величины в интервал $(m-l; m+l)$ симметричный относительно центра рассеивания (рисунок 1.6. а).

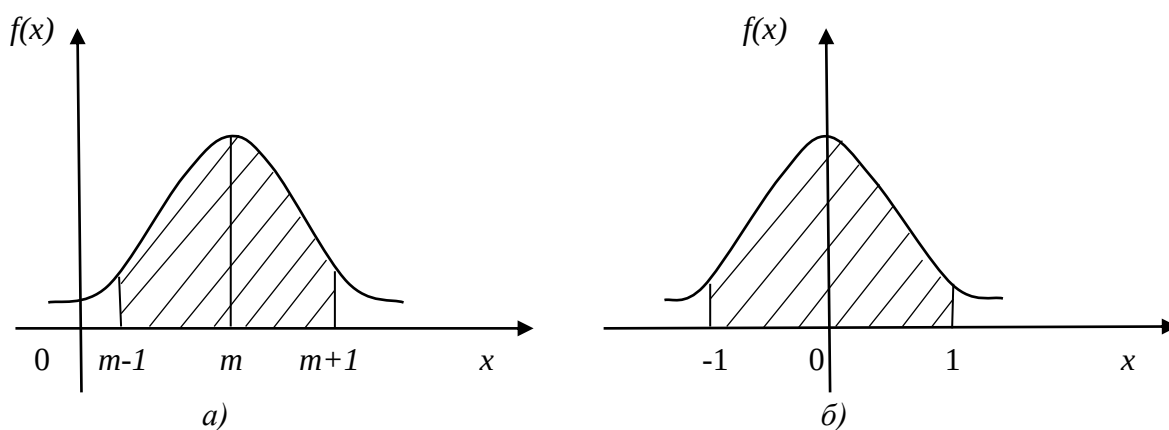


Рис.1.6

Вероятность попадания случайной величины в интервал определяется по формуле

$$P(m-l < X < m+l) = \Phi\left(\frac{l}{\sigma}\right) - \Phi\left(-\frac{l}{\sigma}\right) \quad (1.14)$$

Поскольку функция Лапласа нечетная, то $\Phi\left(-\frac{l}{\sigma}\right) = -\Phi\left(\frac{l}{\sigma}\right)$.

Искомая вероятность будет равна

$$P(m-l < X < m+l) = 2\Phi\left(\frac{l}{\sigma}\right) \quad (1.15)$$

Правая часть данного равенства не зависит от центра рассеивания, поэтому, при $m=0$ (рисунок 1.6. б), получаем

$$P(-l < x < l) = 2\Phi\left(\frac{l}{\sigma}\right) \quad (1.16)$$

Пример 1.6. Найти вероятность попадания в полосу шириной $2l = 3,5\text{ м}$, если ошибки стрельбы подчиняются нормальному закону распределения с параметрами $\alpha = 0$, $\sigma = 1,9$.

Решение. По условию задачи $\alpha = -1,75$; $\beta = 1,75$.

$$P(-1,75 < x < 1,75) = 2\Phi\left(\frac{1,75}{1,9}\right) = 2\Phi(0,9211) = 2 \cdot 0,3212 = 0,6424.$$

Примем за единицу измерения отклонения случайной величины от её центра рассеивания среднее квадратичное отклонение, тогда на основании формулы при различном значении σ имеем

$$P(-\sigma < x < \sigma) = 2\Phi(1) = 0,6826,$$

$$P(-2\sigma < x < 2\sigma) = 2\Phi(2) = 0,9544,$$

$$P(-3\sigma < x < 3\sigma) = 2\Phi(3) = 0,9972.$$

Из полученных равенств видно, что для нормально распределённой случайной величины X все рассеивания укладываются на участки $m \pm 3\sigma$ с вероятностью 0,9972. В этом заключается «правило трёх сигм»: если случайная величина распределена нормально, то абсолютная величина её отклонения от математического ожидания практически не превосходит утроенного среднего квадратического отклонения.

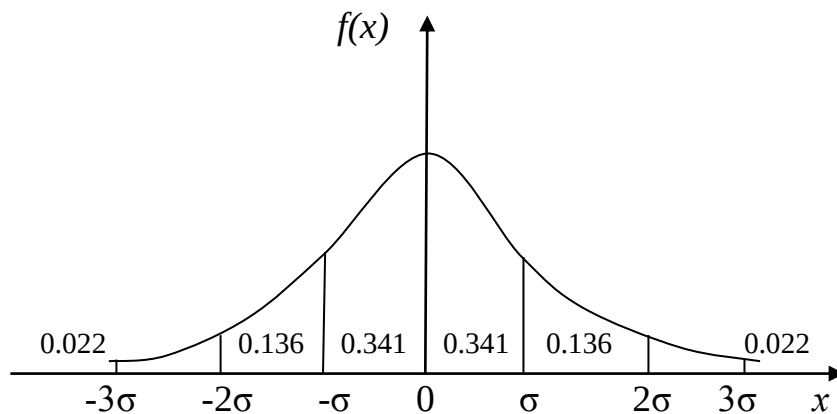


Рис. 1.7

В некоторых случаях наряду со средним квадратическим отклонением пользуются вероятным отклонением. Вероятным отклонением случайной величины X , распределенной по нормальному закону, называется половина длины участка, симметричного относительно центра рассеивания, вероятность попадания в который равна половине. Среднее квадратическое и вероятное отклонения связаны соотношением

$$\frac{E}{\sigma} = 0,674; \quad E = 0,674\sigma.$$

Нормальный закон распределения играет исключительно важную роль в теории вероятностей и занимает среди других законов распределения

особое положение. Главная особенность, выделяющая нормальный закон среди других, состоит в том, что он является предельным законом, к которому приближаются другие законы при неограниченном увеличении числа опытов.

Пример 1.7. Случайная величина X , распределенная по нормальному закону, представляет собой ошибку измерения некоторого расстояния. При измерении допускается систематическая ошибка в сторону завышения на 1,2 мм, среднее квадратическое отклонение ошибки измерения равно 0,8 мм. Найти вероятность того, что отклонение измеренного значения от истинного не превзойдет по абсолютной величине 1,6 мм.

Решение. Ошибка измерения есть случайная величина X , подчиненная нормальному закону с параметрами $m=1,2$ и $\sigma=0,8$. Нужно найти вероятность попадания случайной величины на интервал от $\alpha=-1,6$ до $\beta=1,6$. По формуле (1.13) имеем

$$P(-1,6 < x < 1,6) = \Phi\left(\frac{1,6-1,2}{0,8}\right) - \Phi\left(\frac{-1,6-1,2}{0,8}\right) = \Phi(0,5) - \Phi(-3,5).$$

Пользуясь таблицами функции $\Phi(x)$ (приложение 1), найдем

$$\Phi(0,5) = 0,6915, \quad \Phi(-3,5) = 0,0002.$$

Отсюда

$$P(-1,6 < x < 1,6) = 0,6915 - 0,0002 = 0,6913.$$

1.4.3 Распределение χ^2 «хи квадрат»

Пусть $x_1, x_2, \dots, x_n$ случайные величины, распределенные по нормальному закону с параметрами m_i и σ_i . Если вычислить квадраты нормированных величин U_i

$$U_i = \left(\frac{x_i - m_i}{\sigma_i} \right)^2, \quad (1.17)$$

а затем их сумму, то получим новую случайную величину χ^2 с $\kappa = n$ степенями свободы. Следует отметить, что каждое соотношение, связывающее n наблюдений, уменьшает число степеней свободы на единицу.

Плотность распределения вероятностей для χ^2 имеет

$$P(\chi^2) = \frac{1}{2^{\frac{\kappa}{2}} \Gamma\left(\frac{\kappa}{2}\right)} (x^2)^{\frac{\kappa}{2}-1} e^{-\frac{x^2}{2}} \quad (1.18)$$

Распределение χ^2 определяется одним параметром – числом степеней свободы κ . С увеличением числа степеней свободы распределение медленно приближается к нормальному.

Распределение χ^2 находит широкое теоретическое и практическое применение, например, проверка согласия экспериментального наблюдения с предполагаемым распределением вероятности; получение доверительных интервалов для дисперсии и среднего квадратического отклонения и т.д.

1.4.4 Распределение Стьюдента

Рассмотрим независимые случайные величины U и V , где U – нормированная случайная величина, распределенная по нормальному закону, а V распределена по закону χ^2 с κ степенями свободы. Тогда величина

$$t = \frac{U}{\sqrt{\frac{V}{\kappa}}} \quad (1.19)$$

имеет распределение, которое называют распределением Стьюдента.

С учетом параметров m и S^2 случайной величины X , распределенной по нормальному закону, выражение можно переписать в следующем виде

$$t = \sqrt{n} \cdot \frac{\bar{x} - m}{S}, \quad (1.20)$$

где

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

Плотность вероятности для закона Стьюдента имеет вид

$$S_{n-1}(t) = \frac{\Gamma\left(\frac{n}{2}\right)}{\sqrt{(n-1) \cdot \pi} \cdot \Gamma\left(\frac{n-1}{2}\right)} \cdot \left(1 + \frac{t^2}{n-1}\right)^{-\frac{n}{2}} \quad (1.21)$$

где $\Gamma(x)$ – известная гамма-функция.

С возрастанием числа степеней свободы распределение Стьюдента быстро приближается к нормальному.

Распределение Стьюдента широко используется при проверке статистических гипотез и при нахождении доверительных интервалов для средних значений случайных величин.

1.4.5 Распределение F Фишера-Снедекора

Если U и V независимые случайные величины, распределенные по закону χ^2 со степенями κ_1 и κ_2 , то величина

$$F = \frac{\frac{U}{\kappa_1}}{\frac{V}{\kappa_2}} \quad (1.22)$$

имеет распределение, называемое распределением F Фишера - Снедекора, $\kappa_1 = n_1 - 1$, $\kappa_2 = n_2 - 1$. Распределение F можно также представить в виде отклонения двух исправленных дисперсий двух случайных величин x_1 и x_2 с нормальной плотностью распределения

$$F = \frac{S_1^2}{S_2^2}. \quad (1.23)$$

Плотность F распределения имеет вид

$$P(F) = \frac{\Gamma\left(\frac{\kappa_1 + \kappa_2}{2}\right) \cdot \kappa_1^{\frac{\kappa_1}{2}} \cdot \kappa_2^{\frac{\kappa_2}{2}}}{\Gamma\left(\frac{\kappa_1}{2}\right) \cdot \Gamma\left(\frac{\kappa_2}{2}\right)} \cdot \frac{x^{\frac{\kappa_1 - \kappa_2}{2}}}{(\kappa_1 + \kappa_2 x)^{\frac{\kappa_1 + \kappa_2}{2}}} \quad (1.24)$$

Распределение F определяется двумя параметрами - числами степеней свободы (приложение 2). Данное распределение широко применяется в дисперсионном анализе и при построении математических моделей.

1.5 Методы оценивания параметров

1.5.1 Метод максимального правдоподобия

Наилучшее измеренное значение, $\tilde{a}$ параметра, a есть такое значение параметра, при котором функция распределения вероятностей имеет максимальное значение. Метод оценки параметра на основании данного положения называется методом максимального правдоподобия.

Данный метод приводит к оценкам, обладающим состоятельностью, максимальной эффективностью, но не обязательно несмещенным. Достоинством метода максимального правдоподобия является то, что оценки параметров имеют нормальное распределение.

Пусть $x_1, x_2, \dots, x_n$ результат n независимых наблюдений над случайной величиной X , которая может быть как дискретной, так и непрерывной с плотностью распределения $P(a_1, a_2, \dots, a_n / x_1, x_2, \dots, x_n)$, зависящая от неизвестных параметров $a_1, a_2, \dots, a_n$. Требуется оценить эти параметры по результатам n опытов.

Функцией правдоподобия для непрерывной случайной величины называют выражение

$$L(a_1, a_2, \dots, a_n / x_1, x_2, \dots, x_n) = \prod_{i=1}^n f(a_1, a_2, \dots, a_i / x_1, x_2, \dots, x_i) \quad (1.25)$$

для дискретной случайной величины

$$L(a_1, a_2, \dots, a_n / x_1, x_2, \dots, x_n) = \prod_{i=1}^n P(a_1, a_2, \dots, a_i / x_1, x_2, \dots, x_i). \quad (1.26)$$

При выборе в качестве оценок параметров a_i значений, которые максимизируют L при значениях $(x_1, x_2, \dots, x_n)$ более удобно работать с $\ln L$, а не с самой функцией L .

Величину $\ln L$ можно максимизировать относительно a_i , приравнивая к нулю частные производные от $\ln L$ по каждому из параметров

$$\begin{cases} \frac{\partial \ln L}{\partial a_1} = \frac{\partial \sum_{i=1}^n \ln P(x_i, a_1, a_2, \dots, a_n)}{\partial a_1} = 0, \\ \frac{\partial \ln L}{\partial a_2} = \frac{\partial \sum_{i=1}^n \ln P(x_i, a_1, a_2, \dots, a_n)}{\partial a_2} = 0, \\ \dots, \\ \frac{\partial \ln L}{\partial a_n} = \frac{\partial \sum_{i=1}^n \ln P(x_i, a_1, a_2, \dots, a_n)}{\partial a_n} = 0. \end{cases} \quad (1.27)$$

Решение системы уравнений дает искомые оценки $a_1, a_2, \dots, a_n$, которые называют оценками максимального правдоподобия.

Пример 1.8. Получить оценки максимального правдоподобия для нормального закона распределения.

Пусть $x_1, x_2, \dots, x_n$ — n независимых наблюдений над случайной величиной X , подчиняющейся нормальному закону с математическим ожиданием m и дисперсией σ^2 .

Плотность распределения имеет вид

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

Задача заключается в том, чтобы по результатам опыта оценить параметры m и σ^2 .

Функция правдоподобия имеет вид

$$L(m, \sigma^2 / x_1, x_2, \dots, x_n) = \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} \cdot e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - m)^2}.$$

Прологарифмируем функцию правдоподобия

$$\ln L = \ln \left[\frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} \right] - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - m)^2 = -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - m)^2.$$

Для получения оценок параметров m и σ^2 необходимо решить систему уравнений

$$\begin{cases} \left. \frac{\partial(\ln L)}{\partial m} \right|_{\substack{m=\bar{x} \\ \sigma^2=S^2}} = 0, \\ \left. \frac{\partial(\ln L)}{\partial \sigma^2} \right|_{\substack{m=\bar{x} \\ \sigma^2=S^2}} = 0. \end{cases}$$

Вычислим производные

$$\frac{\partial(\ln L)}{\partial \bar{x}} = -\frac{1}{2S^2} \cdot \frac{\partial}{\partial \bar{x}} \left[\sum_{i=1}^n (x_i - \bar{x})^2 \right] = \frac{1}{S^2} \sum_{i=1}^n (x_i - \bar{x}),$$

$$\frac{\partial(\ln L)}{\partial S^2} = -\frac{n}{2} \cdot \frac{2\pi}{2\pi S^2} - \frac{1}{2} \sum_{i=1}^n (x_i - \bar{x})^2 (-1) \frac{1}{S^4} = -\frac{n}{2S^2} + \frac{1}{2S^4} \sum_{i=1}^n (x_i - \bar{x})^2.$$

Приравняем к нулю производные и найдем оценки для параметров $\bar{x}$ и S^2 .

$$\begin{cases} \frac{1}{S^2} \sum_{i=1}^n (x_i - \bar{x}) = 0, \\ -\frac{n}{2S^2} + \frac{1}{2S^4} \sum_{i=1}^n (x_i - \bar{x})^2 = 0. \end{cases}$$

Из первого уравнения системы имеем

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

Отсюда
$$\tilde{m} = \frac{1}{n} \sum_{i=1}^n x_i = \bar{x}.$$

Для определения σ^2 умножим второе уравнение на $2\sigma^2$ и, учитывая, $\tilde{m} = \bar{x}$, получим

$$-n\sigma^2 + \sum_{i=1}^n (x_i - \bar{x}) = 0.$$

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

Убедимся в том, что функция правдоподобия в точке с координатами $m = \bar{x}$ и $\sigma_{\bar{x}}^2 = S$ достигает максимума. Для этого найдем вторые производные

$$\frac{\partial^2(\ln L)}{\partial m^2} = \frac{\partial}{\partial m} \left[\frac{1}{S} \sum_{i=1}^n (x_i - m) \right] = -\frac{n}{S},$$

$$\frac{\partial^2(\ln L)}{\partial m \partial S} = \frac{\partial}{\partial S} \left[\frac{1}{S} \sum_{i=1}^n (x_i - m) \right] = -\frac{1}{S} \sum_{i=1}^n (x_i - m) = -\frac{1}{S} \sum_{i=1}^n \left[\sum_{i=1}^n x_i - nm \right],$$

$$\frac{\partial^2(\ln L)}{\partial S^2} = \frac{\partial}{\partial S} \left[\frac{n}{2S} + \frac{1}{2S^2} \sum_{i=1}^n (x_i - m)^2 \right] = \frac{n}{2S^2} - \frac{1}{S^3} \sum_{i=1}^n (x_i - m)^2.$$

В точке с координатами $m = \bar{x}$, $S = \sigma_{\bar{x}}^2$

$$\frac{\partial^2(\ln L)}{\partial m^2} = -\frac{n}{\sigma_{\bar{x}}^2} < 0, \quad \frac{\partial^2(\ln L)}{\partial m \partial S} = -\frac{1}{\sigma_{\bar{x}}^2} \left[\sum_{i=1}^n (x_i - n\bar{x}) \right] = 0,$$

$$\frac{\partial^2(\ln L)}{\partial S^2} = \frac{n}{2\sigma_{\bar{x}}^4} - \frac{1}{\sigma_{\bar{x}}^6} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{n}{2\sigma_{\bar{x}}^2} - \frac{n\sigma_{\bar{x}}^2}{\sigma_{\bar{x}}^6} = -\frac{n}{\sigma_{\bar{x}}^4} < 0.$$

Таким образом, второй дифференциал в точке с координатами $m = \bar{x}$, $S = \sigma_{\bar{x}}^2$

$$d^2(\ln L) = -\frac{n}{\sigma_{\bar{x}}^2} (dm)^2 - \frac{n}{\sigma_{\bar{x}}^4} (dS)^2 < 0.$$

Следовательно, функция правдоподобия в этой точке достигает максимума.

1.5.2 Метод наименьших квадратов

Пусть на основании опыта требуется установить функциональную зависимость величины y от величины x .

Обычно экспериментальные точки на графике $y = f(x)$ дают некоторый «разброс», т.е. обнаруживают случайные отклонения от видимой общей закономерности. Эти отклонения связаны с неизбежными при всяком опыте ошибками измерения.

В результате проведения опыта получены n независимых наблюдений случайной величины X . Из теоретических или иных соображений выбран вид независимости $y = f(x)$. Функция $y = f(x)$ содержит ряд числовых параметров $a_1, a_2, \dots, a_n$. Требуется так выбрать эти параметры, чтобы кривая $y = f(x)$ наилучшим образом изображала зависимость, полученную в опыте.

Рассмотрим два часто встречающихся на практике случая: когда функция $y = f(x)$ линейна и когда она выражается многочленом второй степени.

В первом случае ($y = a_1x + a_2$) система уравнений принимает вид

$$\begin{cases} \sum_{i=1}^n x_i y_i - a_1 \sum_{i=1}^n x_i^2 - a_2 \sum_{i=1}^n x_i = 0, \\ \sum_{i=1}^n y_i - a_1 \sum_{i=1}^n x_i - a_2 n = 0. \end{cases}$$

Во втором случае ($y = a_1x^2 + a_2x + a_3$) система уравнений имеет следующий вид

$$\begin{cases} \sum_{i=1}^n x_i^2 y_i - a_1 \sum_{i=1}^n x_i^4 - a_2 \sum_{i=1}^n x_i^3 - a_3 \sum_{i=1}^n x_i^2 = 0, \\ \sum_{i=1}^n x_i y_i - a_1 \sum_{i=1}^n x_i^3 - a_2 \sum_{i=1}^n x_i^2 - a_3 \sum_{i=1}^n x_i = 0, \\ \sum_{i=1}^n y_i - a_1 \sum_{i=1}^n x_i^2 - a_2 \sum_{i=1}^n x_i - a_3 n = 0. \end{cases}$$

Пример 1.9. Пусть в результате экспериментального исследования получены четыре значения искомой функции $y = f(x)$ при четырёх значениях аргумента, которые записаны в таблице 1.3.

Таблица 1.3

x	1	2	3	4
y	3	4	2,5	0,5

Будем искать функцию $y = f(x)$ в виде линейной функции $y = a_1x + a_2$. Составим выражение $S(a_1, a_2)$.

$$S(a_1, a_2) = \sum_{i=1}^4 [y_i - (a_1x + a_2)]^2.$$

На основании формул (1.31) составляем систему уравнений для определения коэффициентов a_1, a_2

$$\begin{cases} 21 - 39a_1 - 11a_2 = 0, \\ 10 - 11a_1 - 4a_2 = 0. \end{cases}$$

Решая систему, находим a_1 и a_2 : $a_1 = -\frac{26}{35}$, $a_2 = \frac{159}{35}$. Искомая прямая (рисунок 1.5) определяется уравнением $y = -\frac{26}{35}x + \frac{159}{35}$.

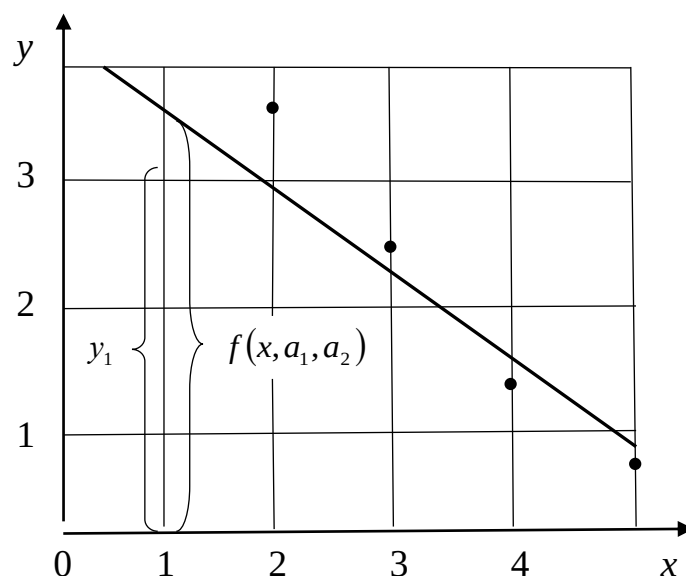


Рис. 1.8

1.6 Доверительный интервал. Доверительная вероятность

1.6.1 Доверительный интервал для выборочного среднего

В ряде задач требуется не только найти для параметра a подходящую оценку $\bar{a}$, но и оценить его точность и надёжность, т.е. указать, к каким ошибкам может привести замена параметра a его оценкой $\bar{a}$. Такого рода задачи особенно актуальны при малом числе наблюдений, когда точечная оценка $\bar{a}$ в значительной мере случайна и приближенная замена a на $\bar{a}$ может привести к серьёзным ошибкам.

Для определения точности и надёжности оценки $\bar{a}$ в математической статистике пользуются доверительными интервалами и доверительными вероятностями.

Пусть для параметра a получена из опыта несмещенная оценка $\bar{a}$.

Оценим возможную при этом ошибку. Выберем некоторую достаточно большую вероятность P такую, что событие с данной вероятностью можно считать практически достоверным, и найдем такое значение ε , для которого

$$P(|\bar{a} - a| < \varepsilon) = \alpha. \quad (1.32)$$

Диапазон практически возможных значений ошибки, возникающей при замене a на $\bar{a}$, равен $\pm \varepsilon$, большие по абсолютной величине ошибки будут появляться с малой вероятностью $\alpha = 1 - P$.

Равенство (1.32) перепишем в виде

$$P(\bar{a} - \varepsilon < a < \bar{a} + \varepsilon) = 1 - \alpha. \quad (1.33)$$

Можно сказать, что случайный интервал

$$J = (\bar{a} - \varepsilon, \bar{a} + \varepsilon) \quad (1.34)$$

накрывает параметр a с вероятностью $P = 1 - \alpha$.

Вероятность P называется *доверительной* вероятностью, величина $\alpha = 1 - P$ - уровнем значимости, интервал J - доверительным интервалом. Границы интервала J : $a_1 = \bar{a} - \varepsilon$ и $a_2 = \bar{a} + \varepsilon$ называются доверительными границами.

Рассмотрим нахождение доверительного интервала для оценки математического ожидания нормального распределения при неизвестном σ , поскольку при решении практических задач обычно значение дисперсии неизвестно.

Закон распределения оценки $\bar{a}$ в общем случае зависит от самих неизвестных параметров величины X . Подберем такую функцию наблюдаемых значений $x_1, x_2, \dots, x_n$, закон распределения которой не зависит от неизвестных параметров, а зависит только от числа опытов и от вида закона распределения величины X . В теории вероятностей доказано, что при нормальном распределении величины X случайная величина

$$t = \sqrt{N} \cdot \frac{\bar{x} - m_x}{S},$$

где

$$\bar{x} = \frac{\sum_{i=1}^N x_i}{N}, \quad S^2 = \frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N-1}$$

имеет распределение Стьюдента, которое определяется параметром N (или, что то же, числом степеней свободы $K = N - 1$) и не зависит от неизвестных параметров a и σ .

Найдём доверительный интервал для математического ожидания. Обозначим через ε половину интервала. Величина ε выбирается из условия

$$P(|\bar{x} - m_x| < \varepsilon) = 1 - \alpha. \quad (1.35)$$

Умножим обе части неравенства $|\bar{x} - m_x| < \varepsilon$ на величину $\frac{\sqrt{N}}{S}$.

$$P\left(\frac{\sqrt{N}|\bar{x} - m_x|}{S} < \frac{\varepsilon}{\frac{S}{\sqrt{N}}}\right) = 1 - \alpha, \quad (1.36)$$

или

$$P\left(|t| < \frac{\varepsilon}{\frac{S}{\sqrt{N}}}\right) = 1 - \alpha. \quad (1.37)$$

Найдём такое число $t_{1-\frac{\alpha}{2}}$, что

$$P(|t| < t_{1-\frac{\alpha}{2}}) = 1 - \alpha. \quad (1.38)$$

Величина $t_{1-\frac{\alpha}{2}}$ находится из условия

$$P(|t| < t_{1-\frac{\alpha}{2}}) = 2 \int_{\sigma}^{1-\alpha} P(t) dt = 1 - \alpha, \quad (1.39)$$

откуда

$$P\left(\bar{x} - t_{1-\frac{\alpha}{2}} \frac{S}{\sqrt{N}} < m < \bar{x} + t_{1-\frac{\alpha}{2}} \frac{S}{\sqrt{N}}\right) = 1 - \alpha. \quad (1.40)$$

Отсюда доверительный интервал для выборочного среднего определяется по выражению

$$\bar{x} - t_{1-\frac{\alpha}{2}} \frac{S}{\sqrt{N}} < m < \bar{x} + t_{1-\frac{\alpha}{2}} \frac{S}{\sqrt{N}} \quad (1.41)$$

Итак, пользуясь распределением Стьюдента, мы нашли доверительный интервал, покрывающий неизвестный параметр μ с надежностью $1-\alpha$. Значение коэффициента $t_{1-\frac{\alpha}{2}}$ приведено в приложении 3.

При неограниченном увеличении числа опытов N распределение Стьюдента стремится к нормальному. Поэтому практически при $N > 30$ можно вместо распределения Стьюдента пользоваться нормальным распределением. Важно подчеркнуть, что при малом числе опытов N ($N < 30$) замена распределения Стьюдента нормальным приводит к грубым ошибкам, а именно, к неоправданному суждению доверительного интервала, т. е. к повышению точности оценки.

Оценка истинного значения измерений величины при проведении независимых равноточных измерений некоторой физической величины производится по выборочному среднему при помощи доверительных интервалов.

Пример 1.10. Проведен ряд измерений диаметра вала. Полученная величина является случайной, распределенной по нормальному закону. Произведено 16 измерений. Получены среднее $\bar{x} = 20,2$ мм и выборочное среднее квадратичное отклонение $\sigma = 0,8$ мм. Оценить неизвестное математическое ожидание диаметра вала при помощи доверительного интервала с надежностью 0,95.

Решение. Для оценки неизвестного выборочного среднего воспользуемся формулой (1.41). Пользуясь приложением 3, найдем коэффициент Стьюдента

$$t_{15;0,975}.$$

Определяем границы доверительного интервала

$$\bar{x} - t_{N-1, 1-\frac{\alpha}{2}} \frac{S}{\sqrt{N}} = 20,20 - 2,13 \frac{0,8}{4} = 19,774 \text{ мм},$$

$$\bar{x} + t_{N-1, 1-\frac{\alpha}{2}} \frac{S}{\sqrt{N}} = 20,20 + 2,13 \frac{0,8}{4} = 20,626 \text{ мм}.$$

Неизвестный диаметр вала с надежностью 0,95 заключен в доверительном интервале

$$19,774 < m_x < 20,626.$$

1.6.2 Доверительный интервал для оценки выборочной дисперсии нормального распределения

Пусть случайная величина X распределена нормально. Требуется оценить неизвестную дисперсию по выборочной дисперсии, т. е. необходимо найти доверительный интервал, покрывающий параметр σ^2 с заданной надежностью $1 - \alpha$.

В теории вероятностей доказано, что случайная величина

$$V = \frac{(N-1)S^2}{\sigma^2} \quad (1.42)$$

Имеет «распределение χ^2 » с $N-1$ степенями свободы. Плотность данного распределения имеет вид

$$P(x, N) = \frac{x^{\frac{N-3}{2}} e^{-\frac{x}{2}}}{2^{\frac{N-1}{2}} \Gamma\left(\frac{N-1}{2}\right)}. \quad (1.43)$$

Это распределение не зависит от оцениваемого параметра σ^2 , а зависит лишь от числа опытов N .

Условимся выбирать интервал $i_{1-\alpha}$ так, чтобы вероятности выхода величина V за пределы интервала вправо и влево были одинаковы и равны $\frac{\alpha}{2}$. Для построения интервала $i_{1-\alpha}$ воспользуемся статистическими таблицами значений χ^2 в зависимости от P и $N-1$ (приложение 4)

$$P(V > \chi^2) = P.$$

По данной таблице (для определённого $N-1$) найдём два значения χ^2 . Одно соответствующее вероятности $P_1 = \frac{\alpha}{2}$, второе вероятности $P_2 = 1 - \frac{\alpha}{2}$, обозначим эти значения χ_1^2 и χ_2^2 . Интервал $i_{1-\alpha}$ слева ограничен χ_2^2 , справа χ_1^2 .

По интервалу $i_{1-\alpha}$ находим искомый доверительный интервал J для дисперсии с границами σ_1^2 и σ_2^2 , который накрывает точку σ^2 с вероятностью $1-\alpha$

$$P(\sigma_1^2 < \sigma^2 < \sigma_2^2) = 1 - \alpha, \quad (1.44)$$

или

$$\frac{S^2(n-1)}{\chi_1^2} < \sigma^2 < \frac{S^2(n-1)}{\chi_2^2}. \quad (1.45)$$

Последнее выражение можно использовать для оценки точности измерений (точности прибора).

Пример 1.11. Производство 20 измерений часового расхода топлива двигателя внутреннего сгорания. Полученная величина является случайной, распределённой по нормальному закону. Получены среднее $\bar{x} = 10,78$, выборочная дисперсия $S^2 = 0,064$. Найти доверительный интервал для дисперсии с доверительной вероятностью 0,8.

По приложению 4 находим значение χ^2 для $(K = 20 - 1 = 19)$

$$P_1 = \frac{\alpha}{2} = 0,1; \quad \chi_1^2 = 27,2;$$

$$P_2 = 1 - \frac{\alpha}{2} = 0,9; \quad \chi_2^2 = 11,65.$$

Определяем границы доверительного интервала

$$\frac{0,064 \cdot 19}{27,2} < \sigma^2 < \frac{0,064 \cdot 19}{11,65};$$

$$0,045 < \sigma^2 < 0,104.$$

В таблице 1.5 приведены формулы для вычисления границ доверительных интервалов математического ожидания, разности математических ожиданий, дисперсии для случаев наиболее часто встречающихся на практике.

1.7 Критерий равенства дисперсий

Часто при обработке экспериментальных данных требуется сравнивать равенство дисперсий нескольких нормально распределённых выборок.

Пусть $S_1, S_2, \dots, S_n$ взаимно независимые статистические оценки дисперсий $\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2$, величины $K_i, \frac{S_i^2}{\sigma_i^2}$ подчиняются χ^2 – распределению с K_i степенями свободы ($K_i = 1, 2, \dots, N$).

Если дисперсии σ_i^2 неизвестны и предлагается, что $\sigma_1^2 = \sigma_2^2 = \sigma_n^2$, то для проверки такой гипотезы (H_0) можно воспользоваться критерием Барлетта

$$M = N \ln \left(\frac{1}{N} K_i S_i^2 \right) - \sum_{i=1}^K K_i \ln S_i^2, \quad N = \sum_{i=1}^K K_i \quad (1.51)$$

Если гипотеза H_0 верна, то отношение

$$\frac{M}{1 + \frac{1}{G(N-1)} \left(\sum_i^n \frac{1}{K_i} - \frac{1}{N} \right)} \quad (1.52)$$

распределено приближённо χ^2 с $N-1$ степени свободы.

Следует отметить, что M критерий Барлетта весьма чувствителен к отклонениям истинных распределений величин $K_i, \frac{S_i^2}{\sigma_i^2}$ от χ^2 – распределения.

В частности, если все оценки S_i^2 найдены по выборкам из совокупностей, распределения которых отличны от нормальных, то M критерий может с большой вероятностью отвергнуть гипотезу H_0 , когда она верна.

Так как распределение статистики M не всегда достаточно хорошо аппроксимируется χ^2 – распределением (например, приближение будет неудовлетворительным, если некоторые из K_i равны 1, 2 и 3), то для отыскания критических значений M предложены специальные таблицы (приложение 5),

Таблица 1.5

Формулы для вычисления границ доверительных интервалов

параметр	условия	Точечная оценка	Доверительный интервал	Распределение
m	σ^2 известна	$\bar{x}$	$\bar{x} - U_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{N}} < m < \bar{x} + U_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{N}}$	Нормальное U
m	σ^2 неизвестна	$\bar{x}$	$\bar{x} - t_{N-1, 1-\frac{\alpha}{2}} \frac{S}{\sqrt{N}} < m < \bar{x} + t_{N-1, 1-\frac{\alpha}{2}} \frac{S}{\sqrt{N}}$	t - Стьюдента
$m_1 - m_2$	$\sigma_1^2 = \sigma_2^2$ известны	$\bar{x}_1 - \bar{x}_2$	$\bar{x}_1 - \bar{x}_2 - U_{1-\frac{\alpha}{2}} \sigma \sqrt{\frac{1}{N_1} + \frac{1}{N_2}} < m_1 - m_2 < \bar{x}_1 - \bar{x}_2 + U_{1-\frac{\alpha}{2}} \sigma \sqrt{\frac{1}{N_1} + \frac{1}{N_2}}$	Нормальное U
$m_1 - m_2$	$\sigma_1^2 = \sigma_2^2$ неизвестны	$\bar{x}_1 - \bar{x}_2$	$\bar{x}_1 - \bar{x}_2 - t_{N_1+N_2-2, 1-\frac{\alpha}{2}} S_p \sqrt{\frac{1}{N_1} + \frac{1}{N_2}} < m_1 - m_2 < \bar{x}_1 - \bar{x}_2 + t_{N_1+N_2-2, 1-\frac{\alpha}{2}} S_p \sqrt{\frac{1}{N_1} + \frac{1}{N_2}}$,где S_p - объединенная оценка дисперсии $S_p = \frac{(N_1-1)S_1^2 + (N_2-1)S_2^2}{N_1 + N_2 - 2}$	t - Стьюдента
σ^2		S^2	$(N-1) \frac{S^2}{\chi^2_{N-1, 1-\frac{\alpha}{2}}} < \sigma^2 < (N-1) \frac{S^2}{\chi^2_{N-1, 1-\frac{\alpha}{2}}}$	χ^2 распределение
$\frac{\sigma_1^2}{\sigma_2^2}$		$\frac{S_1^2}{S_2^2}$	$\frac{S_1^2}{S_2^2} F_{N_2-1, N_1-1, \alpha} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{S_1^2}{S_2^2} F_{N_2-1, N_1-1, \alpha}$	F - Фишера-Снедекора

основанные на уточнении указанной аппроксимации. Эти таблицы позволяют оценить α – процентные точки $m(\alpha)$ статистики M для $\alpha = 0,01$ и $\alpha = 0,05$ в зависимости от N , а также от коэффициентов

$$C_1 = \sum_{i=1}^n \frac{1}{K_i} - \frac{1}{N}, \quad C_3 = \sum_{i=1}^n \frac{1}{K_i^3} - \frac{1}{N}. \quad (1.53)$$

Функция $m(\alpha)$ определяется как решение уравнения $P(M > m) = 0,01\alpha$. Для нахождения $m(\alpha)$ используется приближённая формула

$$m(\alpha) = \frac{1}{\Delta C} [(C_1 - C_3)m_a(\alpha, N, C_1) + (C_3 - C_1)m_b(\alpha, N, C_1)], \quad (1.54)$$

где m_a и m_b – функции, зависящие от α , N и C_1 .

Величины C и ΔC определяется по формулам

$$C = \frac{C_1^3}{N^2}, \quad \Delta C = C_1 - \frac{C_1^3}{N^2}. \quad (1.55)$$

Поскольку в критических случаях $C < C_3 < C + \Delta C$, то искомое решение $m(\alpha)$ заключено между $m_b(\alpha, N, C_1)$ и $m_a(\alpha, N, C_1)$.

В приложении 5 даны значения функций m_a и m_b . При $C_1 = 0$ значение m_a и m_b представляют собой α - процентные точки χ^2 -распределения с $N - 1$ степенями свободы.

Применяя M - критерий Барлетта, рекомендуется пользоваться следующими правилами

1. Вычислить M по формуле (1.51).
2. Сравнить M со значениями функций m_a и m_b в строке « n » приложения 5, если при всех C_1 в данной таблице $m_a \leq M$, то гипотезу равенства дисперсий следует отвергнуть, если при всех C_1 будет $M > m_b$, то гипотеза H_0 принимается.
3. В случае $\max m_a \geq \min m_b$ следует вычислить C_1 и по приложению 5 найти $m_a(\alpha, N, C_1)$ и $m_b(\alpha, N, C_1)$. Если $m_a(\alpha, N, C_1) \leq M$, то гипотеза H_0 отвергается, если $M < m_b(\alpha, N, C_1)$, то гипотеза H_0 принимается.

4. При $m_a(\alpha, N, C_1) > M \geq m_b(\alpha, N, C_1)$ следует найти $m(\alpha)$ по формуле (1.50), если $m(\alpha) \leq M$, то гипотеза H_0 отвергается, если $M < m(\alpha)$, то H_0 принимается.

Для облегчения интерполяции по формуле (1.54) прилагается трёхзначная таблица функций C и ΔC (приложение 5).

Пример 1.14. Для сравнения точности результатов химического анализа зерна в четырёх лабораториях были вычислены значения не смещённых оценок S_i^2 для неизвестных дисперсий σ_i^2 (таблица 1.6)

Таблица 1.6

i	S_i^2	K_i	$\frac{1}{K_i}$	$K_i S_i^2$	$\ln S_i^2$	$K_i \ln S_i^2$
1	1,13	2	0,5000	2,26	0,1222	0,244
2	4,05	7	0,1429	28,35	1,3986	9,791
3	6,38	21	0,0476	133,98	1,8532	8,917
4	2,67	29	0,0345	77,43	0,9821	28,481
	Σ	59	0,7250	242,02	-	77,433

Требуется решить: значимо различие статистических дисперсий или нет (неоднородны или однородны дисперсии)? Воспользуемся критерием Барлетта

$$M = 59 \ln \frac{242,02}{59} - 77,433 = 5,25.$$

Согласно приложения 5 полученное значение M меньше $m_b(0,05, 4, C_1)$ при любых C_1 . Следовательно, с уровнем значимости $\alpha = 0,05$ различие выборочных дисперсий следует признать незначимым, т.е. дисперсии однородны.

Если бы оказалось, что $M = 8,60$, то для проверки гипотезы H_0 нужно было бы вычислить C_1 . В данном случае $C_1 = 0,725 - \frac{1}{59} = 0,708$, поэтому по приложению 5 линейной интерполяцией находим

$$m_a(0,05, 4, 0,708) = 8,24 + 0,42 \cdot 0,39 = 8,40 < 8,60$$

$$m_b(0,05, 4, 0,708) = 8,00 + 0,42 \cdot 0,17 = 8,07.$$

Иными словами, величина $M = 8,60$ свидетельствовала бы о значимости различия между дисперсиями.

Наконец, если бы найденное значение M удовлетворяла неравенствам $8,07 \leq M < 8,40$, то для проверки гипотезы H_0 потребовалось бы вычислить C_3 . В данном случае $C_3 = 0,128$, поэтому, согласно формуле (1.54)

$$m(0,05) = \frac{8,40(0,708 - 0,128) + 8,07(0,128 - 0,022)}{0,686} = 8,35.$$

Критерий Барлетта используется в тех случаях, когда количество измерений в различных выборках неодинаково.

В случае равенства количества измерений в отдельных выборках однородность дисперсий можно проверить по критерию Кохрена. Данный критерий менее мощный по сравнению с критерием Барлетта, но зато более простой. Статистика критерия Кохрена выражается формулой

$$G_p = \frac{S_{\max}^2}{S_1^2 + S_2^2 + \dots + S_n^2}, \quad (1.56)$$

где $S_{\max}^2$ - максимальная дисперсия из $S_1^2, S_2^2, \dots, S_n^2$.

Если выполняется неравенство $G_p < G_\tau$ (G_τ - табличное значение критерия Кохрена), то гипотеза об однородности дисперсий принимается. В противном случае гипотеза об однородности дисперсий отклоняется.

Таблицы процентных точек критерия Кохрена приведены в приложении (приложение 6).

1.8 Исключение резко выделяющихся наблюдений

Часто экспериментальные исследования дают неоднородные данные. Иногда большие отклонения возникают в результате случайного просчёта, неправильного чтения показаний прибора, случайного сдвига запятой в десятичной записи числа и т.д., т. е. в результате действительной ошибки. Иногда же они отражают более тонкие моменты, такие как несоответствие в отдель-

ных точках действительности используемой математической модели, незамеченное исследователем изменение условий эксперимента и т.п. В любом случае с математической точки зрения речь идёт о выделении наблюдений, величина которых не согласуется с распределением основной массы данных. Отклоняющиеся опытные данные следует исключить, пользуясь статистическими критериями, и только в том случае, если они находятся в средней части графика. Существует несколько методов исключения «промаха». Рассмотрим наиболее обоснованный с теоретической точки зрения критерий, предложенный Анскомбом.

Пусть даны результаты опытных данных $x_1, x_2, \dots, x_n$ случайной величины X , распределённой по нормальному закону с параметрами $\bar{x}$ и S .

Вычислим разности $Y_i = X_i - \bar{x}$. При использовании максимальной разности $Y_m = X_m - \bar{x}$ правило обнаружения промаха состоит в следующем: при заданном значении коэффициента C наблюдение X_m отбрасывается, если $|Y_m| > CS$. В противном случае наблюдение X_m оставляется. Величина коэффициента C может изменяться в зависимости от количества опытных данных. Анскомб выразил коэффициент C неявно через критерий Стьюдента

$$\sqrt{\frac{NC^2(k-1)}{k\left(k - \frac{NC^2}{k}\right)}} \approx t_{k, 1-k}^{k-1}, \quad (1.57)$$

где $K = N - 1$.

Пример 1.15. Дан ряд опытных значений часового расхода топлива двигателя внутреннего сгорания 23,2; 23,4; 23,5; 24,1; 25,5. Проверить, является ли последнее значение $X_5 = 25,5$ промахом и следует ли выбросить его из экспериментальных данных.

Решение. Вычислим выборочное среднее $\bar{x} = 23,9$, выборочную дисперсию $S^2 = 0,77$, а затем определим разность $Y_5 = X_5 - \bar{x} = 25,5 - 23,9 = 1,6$. Для $\alpha = 0,05$, $N = 5$, $K = 4$ имеем

$$\sqrt{\frac{5C^2(4-1)}{4\left(4-\frac{5C^2}{4}\right)}} = 2,77C^3.$$

Отсюда находим $C = 1,49$. Согласно критерию имеем

$$|1,6| > 1,49 \cdot 0,77 = 1,05.$$

Таким образом, значение X_5 не удовлетворяет неравенству и его можно исключить из экспериментальных данных, т.е. оно является «промахом».

На практике для исключения резко выделяющихся наблюдений часто используют критерий Стьюдента

$$t_p = \frac{|x^* - \bar{x}|}{s}, \quad (1.58)$$

где x^* - резко выделяющееся наблюдение, $\bar{x}$ - среднее значение, вычисленное без учёта x^* .

Если $t_p > t_T$ (t_T - табличное значение критерия Стьюдента), то результат x^* исключается из дальнейшей обработки, в противном случае результат x^* оставляется для дальнейшей обработки.

Однородность дисперсий различных выборок свидетельствует о том, что в статистических данных нет систематических ошибок, и измерения на всех уровнях переменной X имеют примерную одну и ту же точность.

2 РЕГРЕССИОННЫЙ АНАЛИЗ

Среди методов прикладной математической статистики регрессионный анализ занимает исключительное положение. Это связано не только с тем, что регрессионный анализ является одним из самых распространенных методов обработки результатов наблюдений, но и с тем, что он служит основой ряда других разделов математической статистики (планирование эксперимента, дисперсный анализ). Регрессионный анализ устанавливает количественные характеристики изменения случайной величины.

Целью регрессионного анализа является установление уравнения приближенной регрессии, оценивая тем или иным способом величину и вероятность этой приближенности. Уравнение приближенной регрессии имеет вид $y = f(x)$. Определение коэффициентов данного уравнения (модели), доверительных интервалов, проверка гипотезы о значимости, адекватности модели носит название регрессионного анализа.

Из курса математического анализа известно, что любая непрерывная функция может быть со сколь угодно высокой точностью заменена многочленом, при этом повышение точности достигается за счет повышения степени многочлена. Поэтому уравнение любой нелинейной регрессии (модели), можно представить в виде многочлена.

В качестве уравнения нелинейной регрессии рассмотрим полиномиальную модель m -го порядка с одной переменной

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \dots + \beta_m x^m \quad (2.1)$$

Здесь предполагается, что коэффициенты β_i неизвестны и что при каждом фиксированном значении $x = x_i$ соответствующее значение y_i наблюдается со случайной ошибкой.

В случае, если значения x равностоящие

$$x_2 - x_1 = x_3 - x_2 = \dots = x_i - x_{i-1},$$

то уравнение (2.1) можно написать следующим образом

$$y = \beta_0 P_n^{(0)}(x') + \beta_1 P_n^{(1)}(x') + \beta_2 P_n^{(2)}(x') + \dots + \beta_m P_n^{(m)}(x'), \quad (2.2)$$

где $P_n^j(x')$ - многочлены Чебышева, n - количество измерений на разных уровнях, x' - новая переменная, которая связана с переменной x зависимостью

$$x' = \frac{x - x_1}{x_2 - x_1} - \frac{n-1}{2}.$$

Первые шесть многочленов Чебышева выражаются формулами

$$\left. \begin{aligned} P_n^{(0)}(x') &= 1, \quad P_n^{(1)}(x') = \lambda_n^{(1)}(x'), \quad P_n^{(2)}(x') = \lambda_n^{(2)} \left[(x')^2 - \frac{1}{12}(n^2 - 1) \right], \\ P_n^{(3)}(x') &= \lambda_n^{(3)} \left[(x')^3 - \frac{1}{20}(3n^2 - 7)x' \right], \\ P_n^{(4)}(x') &= \lambda_n^{(4)} \left[(x')^4 - \frac{1}{14}(3n^2 - 13)(x')^2 + \frac{3}{560}(n^2 - 1)(n^2 - 9) \right], \\ P_n^{(5)}(x') &= \lambda_n^{(5)} \left[(x')^5 - \frac{5}{18}(n^2 - 7)(x')^3 + \frac{1}{1008}(15n^4 - 230n^2 + 407)x' \right], \\ P_n^{(6)}(x') &= \lambda_n^{(6)} \left[(x')^6 - \frac{5}{24}(3n^2 - 31)(x')^4 + \frac{1}{176}(5n^4 - 110n^2 + 329)(x')^2 - \right. \\ &\quad \left. - \frac{15}{14784}(n^2 - 1)(n^2 - 9)(n^2 - 25) \right]. \end{aligned} \right\} \quad (2.3)$$

Многочлены Чебышева приведены в приложении 7.

Уравнение (2.2) можно переписать в виде

$$y = \sum_{j=0}^m \beta_j P_n^{(j)}(x'), \quad (j = 0, 1, 2, 3, \dots, m). \quad (2.4)$$

Оценки коэффициентов β_j определяются методом наименьших квадратов по формуле

$$\beta_j = \frac{\sum_{i=1}^n y_i P_n^{(j)}(x'_i)}{\sum_{i=1}^n [P_n^{(j)}(x'_i)]^2} = \frac{\sum_{j=1}^n y_i P_n^{(j)}(x'_i)}{A_{jn}^2} \quad (2.5)$$

Значения $P_n^{(j)}(x'_i)$ и A_{jn}^2 даны в таблицах многочленов Чебышева.

Следующим этапом является проверка воспроизводимости эксперимента, т.е. проверка одного из допущений регрессионного анализа - однородности выборочных дисперсий S_g^2 ($g = 1, 2, 3, \dots, N$) на различных уровнях переменной X . Задача состоит в проверке гипотезы о равенстве генеральных дисперсий $\sigma^2\{y_n\}$. Оценки дисперсий находят по формуле

$$S_g^2 = \frac{1}{n-1} \sum_{i=1}^n (y_{gi} - \bar{y}_g), \quad (2.6)$$

где n - количество измерений на одном уровне переменной X (число параллельных измерений).

Всю группу выборочных дисперсий S_g^2 можно считать оценками для одной и той же генеральной дисперсии $\sigma^2\{y\}$ воспроизводимости эксперимента, ее наилучшая оценка имеет вид

$$S_{\text{ср}}^2\{y\} = \frac{1}{K} \sum_{g=1}^N S_g^2\{y\} \quad (2.7)$$

с числом степеней свободы $K = N(n-1)$.

Если проверка воспроизводимости эксперимента дала отрицательный результат, то при повторении эксперимента следует увеличить число n параллельных опытов.

После определения оценок b_j коэффициентов уравнения регрессии необходимо проверить гипотезу об их значимости. Проверку производят по критерию Стьюдента

$$t_p = \frac{|b_j|}{S\{b\}}, \quad (2.8)$$

где

$$S^2\{b\} = \frac{1}{Nn} S_{\text{ср}}^2\{y\}. \quad (2.9)$$

Если t_p окажется больше критического значения критерия Стьюдента, то оценку коэффициента b_j признают значимой. В противном случае оценка коэффициента b_j считается незначимой, т.е. коэффициент b_j можно исключить из уравнения регрессии. Число степеней свободы для критического значения коэффициента Стьюдента равно $N(n-1)$.

Доверительные интервалы для коэффициента b_j определяются по выражению

$$b_j - t_{N(n-1), 1-\frac{\alpha}{2}} \frac{S\{b\}}{A_{jn}} < \beta_j < b_j + t_{N(n-1), 1-\frac{\alpha}{2}} \frac{S\{b\}}{A_{jn}}. \quad (2.10)$$

Доверительны интервалы для $\hat{y}_g$ при фиксированном x_g находятся по формуле

$$\hat{y}_g - t_{N(n-1), 1-\frac{\alpha}{2}} \frac{S_g}{A_{jn}} < y_g < \hat{y}_g + t_{N(n-1), 1-\frac{\alpha}{2}} \frac{S_g}{A_{jn}}. \quad (2.11)$$

Следующим этапом регрессионного анализа является проверка адекватности полученной математической модели. Чтобы проверить гипотезу об адекватности уравнения регрессии опытным данным, достаточно оценить отклонение предсказанной по полученному уравнению регрессии величины $\hat{y}_g$ от результатов измерений $\bar{y}_g$ на одних и тех же уровнях переменной x . Расстояние результатов измерений вблизи уравнения регрессии, оценивающего истинную функцию $y = f(x)$, можно охарактеризовать с помощью дисперсии адекватности

$$S_{ad}^2 = \frac{n}{N-d} \sum_{g=1}^N \left(\bar{y}_g - \hat{y}_g \right)^2, \quad (2.12)$$

где d – число членов уравнения регрессии.

Число степеней свободы дисперсии адекватности равно $N - d$.

Проверка гипотезы об адекватности состоит в выяснении соотношения между дисперсиями адекватности и оценкой дисперсии воспроизводимости. Если эти оценки дисперсии однородны, то математическое описание адекватно представляет результаты опыта; если же нет, то описание считается неадекватным. Проверку гипотезы об адекватности производят при помощи F – критерия Фишера. Критерий Фишера позволяет проверить гипотезу об однородности двух выборочных дисперсий S_{ad}^2 и $S_{voc}^2\{y\}$. В том случае, если $S_{ad}^2 > S_{voc}^2\{y\}$

$$F = - \frac{S_{ad}^2}{S_{voc}^2\{y\}}. \quad (2.13)$$

Если вычисленное значение F критерия меньше критического F критерия для соответствующих степеней свободы

$$K_{ad} = N - d, \quad K_{zn} = N(n-1)$$

при заданном уровне значимости, то гипотезу об адекватности не отвергают. В противном случае гипотезу об адекватности отвергают и математическая модель $y = f(x)$ признается неадекватной.

Проверка адекватности возможна при $K_{ad} > 0$. Если число N уровней переменной x равно числу всех значимых коэффициентов уравнения регрессии, то для проверки гипотезы об адекватности математического описания степеней свободы не остается ($K_{ad} = 0$). Если же некоторые оценки коэффициентов уравнения регрессии оказались незначительными, то число d членов проверяемого уравнения в этом случае меньше N , и для проверки гипотезы об адекватности остается одна или несколько степеней свободы.

Пример 2.1. При моделировании условий подачи хлебной массы пальчиковой частью шпека жатки к плавающему транспортеру зерноуборочного комбайна были получены следующие положения материальной частицы (табл.2.1) для угловой скорости шпека 10% и угла наклона колена оси пальцев 75° .

Таблица 2.1

x , мм абцисса	250	400	550	700	850
выборка	1	2	3	4	5
y , мм ордината	185	225	219	189	141
	187	234	236	189	132
	190	232	237	189	140
	194	246	241	192	152
	183	252	220	195	150
среднее $\bar{y}_g$	189,80	237,80	230,60	190,80	143,0

Требуется по экспериментальным данным найти уравнение регрессии $y = f(x)$.

Обработка экспериментальных данных проводится в следующей последовательности:

1.Проверка на исключение резко выделяющихся наблюдений по выборкам.

В первой выборке проверим значение $y_{14}^* = 194$.

Среднее значение $\bar{y}_1$ в первой выборке без учета y_{14}^*

$$\bar{y}_1 = \frac{1}{n-1} \sum_{i=1}^{n-1} y_{1i} = \frac{1}{4} (185 + 187 + 190 + 183) = 188,50.$$

Среднее квадратическое отклонение

$$S_1 = \sqrt{\frac{\sum_{i=1}^{n-1} (y_{1i} - \bar{y}_1)^2}{n-1-1}} =$$

$$= \sqrt{\frac{1}{3} [(187 - 188,50)^2 + (190 - 188,50)^2 + (185 - 188,50)^2 + (183 - 188,50)^2]} = 4,65.$$

Расчетное значение критерия Стьюдента

$$t_p = \frac{|y_{14}^* - \hat{y}_1|}{S_1} = \frac{194 - 188,50}{4,65} = 1,18.$$

Из таблицы 1 приложения находим критическое значение критерия Стьюдента при числе степеней свободы $K = n - 1 - 1 = 5 - 1 - 1 = 3$ и уровне значимости $\alpha = 0,05$; $t_T = 3,18$.

Поскольку $t_p < t_T$, то значение y_{14}^* оставляется для дальнейшей обработки.

Аналогично проверяли на исключение и другие, резко выделяющиеся наблюдения:

$$y_{25}^* = 252; \quad \bar{y}_2 = 239,75; \quad S_2 = 9,03; \quad t_{p_2} = 1,36 < t_T;$$

$$y_{34}^* = 246; \quad \bar{y}_3 = 234,75; \quad S_3 = 10,81; \quad t_{p_3} = 1,04 < t_T;$$

$$y_{45}^* = 196; \quad \bar{y}_4 = 190,08; \quad S_4 = 1,50; \quad t_{p_4} = 2,80 < t_T;$$

$$y_{54}^* = 152; \quad \bar{y}_5 = 143,00; \quad S_5 = 9,29; \quad t_{p_5} = 0,97 < t_T.$$

Поскольку все значения y в различных выборках оказались однородными, то они участвуют в дальнейшей обработке. После проверки на однородность определяется среднее значение $\bar{y}_g$ для каждой выборки, средние значения $\bar{y}_g$ даны в последней строке таблицы 2.1.

2. Проверка воспроизводимости эксперимента.

Определим оценки дисперсий выборок по формуле (2.6.)

$$S_g^2 = \frac{1}{n-1} \sum_{i=1}^n (y_{gi} - \bar{y}_g)^2.$$

Оценка дисперсии для первой выборки

$$S_1^2 = \frac{1}{5-1} [(185-189,80)^2 + (187-189,80)^2 + (190-189,80)^2 + (194-189,80)^2 + (183-189,80)^2] = 18,70.$$

Аналогично вычисляем оценки для первой выборки

$$S_2^2 = 104,70; \quad S_3^2 = 106,30; \quad S_4^2 = 7,20; \quad S_5^2 = 66,00.$$

Поскольку количество измерений в каждой выборке одинаково, то проверку на однородность дисперсий производим по критерию Кохрена

$$G_p = \frac{S_{\max}^2}{S_1^2 + S_2^2 + S_3^2 + S_4^2 + S_5^2} = \frac{106,30}{18,70 + 104,70 + 106,30 + 7,20 + 66,00} = 0,351.$$

При степени свободы $n-1=5-1=4$, $N=5$ и уровне значимости $\alpha=0,05$ определяем по таблице 5 приложения табличное значение критерия

$$G_T = 0,629.$$

Поскольку $G_p < G_T$, то гипотеза об однородности выборочных дисперсий принимается.

Оценка генеральной дисперсии воспроизводимости эксперимента определяется по формуле(2.7)

$$S_{\text{соч}}^2 \{y\} = \frac{1}{N} \sum_{g=1}^N S_g^2 \{y\} = \frac{1}{5} (18,70 + 104,70 + 106,30 + 7,20 + 66,00) = 60,58.$$

3. Вычисление оценок коэффициентов уравнения регрессии.

Оценки коэффициентов уравнения регрессии определяются по формуле (2.14)

$$b_j = \frac{\sum_{i=1}^n y_i P_n^{(j)}(x'_i)}{\sum_{i=1}^n [P_n^{(j)}(x'_i)]^2}. \quad (2.14)$$

В таблице многочленов Чебышева (приложение 7) найдем значения $P_5^{(j)}(x')$, соответствующие $n=5$.

Таблица 2.2

Часть таблицы многочленов Чебышева

$\bar{y}_g$	$P_5^{(0)}(x')$	$P_5^{(1)}(x')$	$P_5^{(2)}(x')$	$P_5^{(3)}(x')$	$P_5^{(4)}(x')$
189,80	1	-2	2	-1	1
237,80	1	-1	-1	2	4
230,60	1	0	-2	0	6
190,80	1	1	-1	-2	-4
143,0	1	2	2	1	1
A_{jn}^2	4	10	14	10	70
$\lambda_n^{(j)}$	-	1	1	5/6	35/12

По формуле (2.5.) вычисляем оценки коэффициентов

$$b_0 = \frac{189,8 \cdot 1 + 237,8 \cdot 1 + 230,6 \cdot 1 + 190,8 \cdot 1 + 143,0 \cdot 1}{5} = 198,04;$$

$$b_1 = \frac{189,8 \cdot (-2) + 237,8 \cdot (-1) + 230,6 \cdot 0 + 190,8 \cdot 1 + 143,0 \cdot 2}{10} = -14,42;$$

$$b_2 = \frac{189,8 \cdot 2 + 237,8 \cdot (-1) + 230,6 \cdot (-2) + 190,8 \cdot (-1) + 143,0 \cdot 2}{14} = -16,27;$$

$$b_3 = \frac{189,8 \cdot (-1) + 237,8 \cdot 2 + 230,6 \cdot 0 + 190,8 \cdot (-2) + 143,0 \cdot 1}{10} = 4,54;$$

$$b_4 = \frac{189,8 \cdot 1 + 237,8 \cdot 4 + 230,6 \cdot 6 + 190,8 \cdot (-4) + 143,0 \cdot 1}{70} = 0,0029.$$

Согласно формуле (2.2.) составим уравнение регрессии

$$y = 198,04 - 14,42x' - 16,27[(x')^2 - 2] + 4,54 \cdot \frac{5}{6} \left[(x')^3 - \frac{68}{20}x' \right] +$$

$$+ 0,0029 \cdot \frac{35}{12} [(x')^4 - 4,43(x')^2 + 2,06].$$

Чтобы получить уравнение регрессии в конечном виде, выразим переменную x' через x

$$x' = \frac{x - x_1}{x_2 - x_1} - \frac{N-1}{2} = \frac{x - 250}{400 - 250} - \frac{5-1}{2} = \frac{x - 550}{150}.$$

В уравнении (2.14.), произведя замену x' на $\frac{x - 550}{150}$, получим

$$y = 198,04 - 14,42 \left(\frac{x-550}{150} \right) - 16,27 \left[\left(\frac{x-550}{150} \right)^2 - 2 \right] + 3,37 \left[\left(\frac{x-550}{150} \right)^3 - 3,4 \left(\frac{x-550}{150} \right) \right] + 0,088 \left[\left(\frac{x-550}{150} \right)^4 - 4,43 \left(\frac{x-550}{150} \right)^2 + 2,06 \right]$$

или после преобразований

$$y = -87,410 + 1,563x - 0,0023x^2 + 0,000001x^3 + 0,00000000017x^4.$$

Проверка гипотезы о значимости коэффициентов уравнения регрессии показывает, что оценку β_4 можно считать статистически незначимой, т.е. $\beta_4 = 0$. Тогда конечное уравнение регрессии имеет вид

$$y = -87,410 + 1,563x - 0,0023x^2 + 0,000001x^3. \quad (2.15)$$

График зависимости (2.15.) и разброс экспериментальных данных изображены на рисунке 2.1.

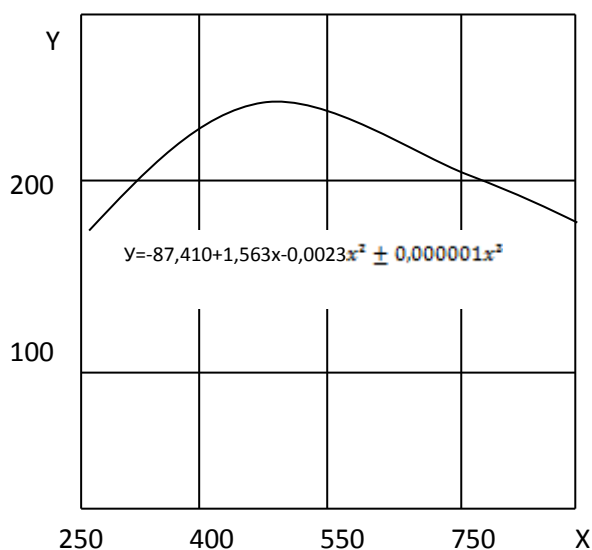


Рис. 2.1

4. Проверка адекватности уравнения регрессии.

Для каждого уровня независимой переменной X определяем значение

$\hat{y}_g$, вычисленное по уравнению регрессии (2.15).

Находим $\hat{y}_1$ для первого уровня

$$\hat{y}_1 = -87,410 + 1,563 \cdot 250 - 0,0023 \cdot 250^2 + 0,000001 \cdot 250^3 = 175,215.$$

Аналогично вычисляем

$$\hat{y}_2 = 233,79; \quad \hat{y}_3 = 242,49; \quad \hat{y}_4 = 222,69; \quad \hat{y}_5 = 193,52.$$

Дисперсию адекватности вычисляем по формуле (2.12.)

$$S_{ad}^2 = \frac{5}{5-4} \left[(189,9 - 175,12)^2 + (237,8 - 233,79)^2 + (230,6 - 242,49)^2 + (190,8 - 222,69)^2 + (143,0 - 193,52)^2 \right] = 19710,98.$$

Средние экспериментальные значения $\bar{y}_g$ взяты из последней строки таблицы 1.2.

Расчетное значение F – критерия Фишера определяем по формуле

$$F = \frac{S_{ad}^2}{S_{\text{вс}}^2 \{y\}} = \frac{19710,98}{60,58} = 325,37. \quad (2.16)$$

При степенях свободы $K_{ad} = N - d = 5 - 4 = 1$, $K_{\text{зн}} = N(n-1) = 5(5-1) = 20$ и при уровне значимости $\alpha = 0,05$ по таблице 7 приложения находим табличное значение f - критерия

$$F_T = 4,35.$$

Поскольку $F_p > F_T$, то гипотезу об адекватности полученной модели отвергаем. Математическое описание исследуемого процесса признаем неадекватным.

Так как гипотеза об адекватности отвергнута, то необходимо переходить к более сложной форме математического описания изучаемого процесса. В соответствии с выше приведенной методикой обработки опытных данных (с использованием многочленов Чебышева) для этого нужно увеличить N – количество уровней варьирования независимой переменной x .

3 МЕТОДИКА АКТИВНОГО ПЛАНИРОВАНИЯ ЭКСПЕРЕМЕНТА

3.1 Общие сведения

Научно-техническая революция ставит перед инженерно-техническими и научными работниками задачу - при минимальных затратах труда, средств и времени получить максимальную отдачу, т.е. найти оптимальный вариант процесса получения продукции. В качестве продукции может быть представлен урожай, химический состав, механические свойства и т.п., т.е. как материальные, так и качественные показатели. На современном этапе развития производительных сил мощным средством повышения производительности труда стал **метод математического моделирования** изучаемого процесса. Прежде чем качать эксперимент, исследователь должен иметь априорные данные об изучаемом процессе. Эти данные устанавливают по литературным источникам, из поисковых опытов и т.п. При этом он должен иметь четкое представление о цели своего исследования. Для этого ему необходимо: во-первых - выяснить, какие из факторов следует фиксировать на тех или иных значениях; во-вторых - установить критерий оптимизации; в-третьих - заранее продумать и спланировать ход эксперимента. Если исследователь имеет полную информацию, полученную в результате реализаций эксперимента, то он составляет **детерминированную модель** процессов в виде системы дифференциальных уравнений.

Реализация эксперимента может быть осуществлена последовательным изучением неизвестных сторон процесса, так называемых факторов. Этот метод дает возможность получить полную информацию о влиянии одного фактора на изучаемый процесс. Для изучения влияния другого фактора нужно фиксировать изученный фактор на лучшем его значении (уровне) и изменять (варьировать) другим фактором. Такая методика исследования получила название **пассивного планирования** или **однофакторного планирования** эксперимента. Стратегия поиска наилучшего результата (поиска параметра

оптимизации) путем последовательного варьирования факторами имеет ряд существенных недостатков. Основными из них являются: необходимость постановки большого числа опытов, чтобы иметь достаточную информацию об исследуемом процессе; значительная продолжительность проведения эксперимента; отсутствие информации о совместном влиянии на параметр оптимизации; отсутствие возможности оценить влияние неконтролируемых факторов и др.

Указанные выше недостатки можно устранить, если заранее составить план эксперимента, в котором предусмотреть варьирование всеми факторами одновременно. Реализация такого плана позволит оценить не только влияние отдельных факторов на оптимизацию процесса, но и их взаимодействие; снизить количество опытов, повысить точность эксперимента.

Такая стратегия постановки опытов получила название **оптимального планирования, активного планирования** или **планирования многофакторного** эксперимента. Планирование многофакторного эксперимента позволяет активно участвовать в исследуемом процессе, оперативно вмешиваться в процесс. Такой вид планирования постановки опытов дает возможность выделить наиболее активные факторы и не исследовать факторы, оказывающие незначительное влияние на процесс. Вследствие специально разработанных планов постановки эксперимента нахождение математической модели исследуемого процесса не требует сложных математических расчетов.

Результаты эксперимента чаще всего выражают математической моделью. Математическая модель $y = f(x_1, x_2, \dots, x_n)$ чаще выражается в виде полинома – отрезка ряда Тейлора, в который разлагается неизвестная функция

$$y_0 = \beta_0 + \sum_i^k \beta_i x_i + \sum_{\substack{i, u=1 \\ u \neq i}}^k \beta_{ui} x_u x_i + \sum_i^k \beta_{ii} x_i^2 + \dots,$$

$$\text{где } \beta_i = \frac{\partial \varphi}{\partial x_i}; \quad \beta_{ui} = \frac{\partial^2 \varphi}{\partial x_u \partial x_i}; \quad \beta_{ii} = \frac{\partial^2 \varphi}{2 \partial x_i^2}.$$

Поскольку в реальном процессе всегда имеют место неуправляемые и неконтролируемые переменные W , изменение величины уносит случайный характер. При обработке экспериментальных данных получают так называемые выборочные коэффициенты уравнения регрессии b_0, b_i, b_{ui}, b_{ii} являющиеся оценками теоретических коэффициентов $\beta_0, \beta_i, \beta_{ui}, \beta_{ii}$.

Уравнение регрессии, полученное на основании данных опыта, будет иметь вид

$$y = b_0 + \sum_i^k b_i x_i + \sum_{\substack{i, u=1 \\ u \neq i}}^k b_{ui} x_u x_i + \sum_i^k b_{ii} x_i^2 + \dots,$$

где b_0 – свободный член уравнения регрессии;

b_{ui} – коэффициент при парном взаимодействии факторов;

b_{ii} – коэффициент при квадратичном взаимодействии факторов;

x – независимые переменные, называемые факторами.

Пространство, ограниченное факторами с координатами x_i , называют **факторным пространством**, а геометрическое изображение функции в факторном пространстве – **поверхностью отклика**.

Следовательно, изменение величины y будет ограничиваться координатами точек x_i и в пределах факторного пространства будут находиться одна или несколько экстремальных точек. Для определения более точного месторасположения экстремальной точки y следует сократить объем факторного пространства путем изменения координат точек x_i .

Экстремальную точку также можно определить, двигаясь к ней по наикратчайшему пути.

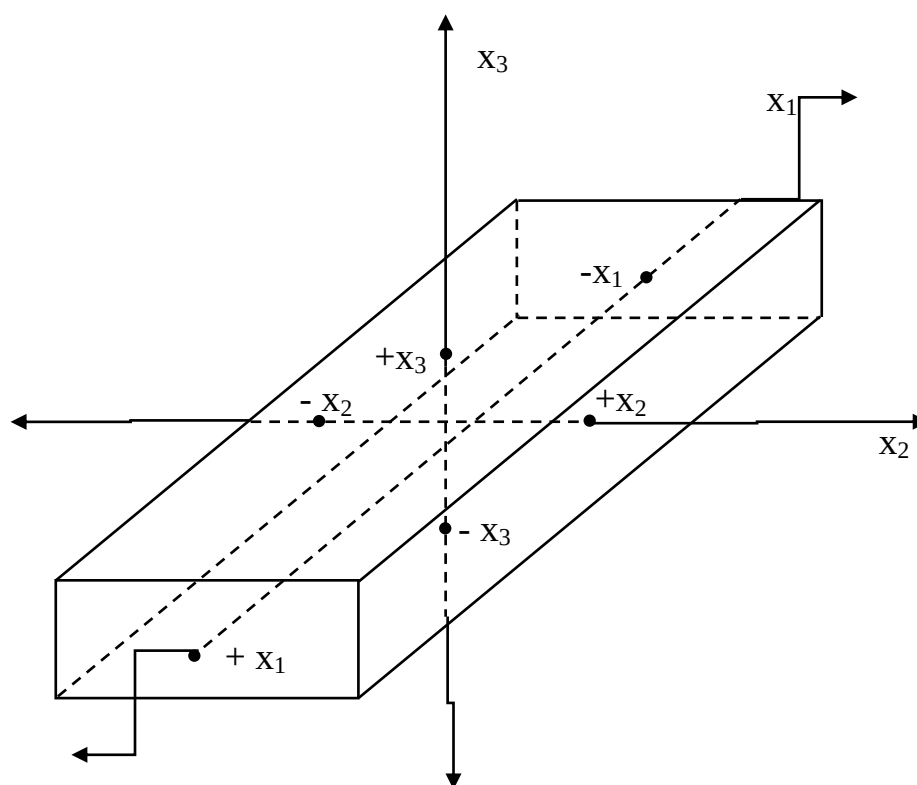


Рис. 3.1. Факторы и факторное пространство

Исходя из вышеизложенного, стратегия поиска экстремальной точки будет состоять из следующих этапов:

1. Выбирают и устанавливают границы факторного пространства. Определяют область (поверхность отклика) с максимальным выходом параметра оптимизации при помощи постановки полного факторного эксперимента (ПФЭ).

В случае большого числа факторов, ограничивающих факторное пространство, вначале производят оценку их влияния на процесс оптимизации путем реализации дробного факторного эксперимента, а затем ставится полный факторный эксперимент.

2. В области с максимальным выходом параметра оптимизации реализуют еще ряд опытов, направленных на улучшение оптимизирующего параметра путем «крутого восхождения по поверхности отклика» к точке y_{opt} .

На каждом этапе реализации эксперимента разрабатывается математическая модель, производится её математическая обработка и принимается решение.

Изучение и реализацию методик активного планирования эксперимента начнем с полного факторного эксперимента.

3.2 Полный факторный эксперимент

Полный факторный эксперимент (ПЭФ) реализует все возможные комбинации факторов на всех выбранных для исследования уровнях. Уровни факторов представляют собой точки факторного пространства. Наиболее часто используются 2 ($+n$ и $-n$), 3 ($+n$, 0, $-n$) и реже 4 уровня. Необходимое количество опытов определяется по формуле

$$N = n^k, \quad (3.1)$$

где N – количество опытов;

n – количество уровней варьирования;

k – количество факторов.

В результате реализации всех опытов получают ряд данных, которые математически обрабатывают. В полученном уравнении регрессии коэффициенты определяют либо составлением и решением системы нормальных уравнений, либо методами матричной алгебры.

Рассмотрим методику определения коэффициентов уравнения регрессии методом матричной алгебры.

В матричной форме система нормальных уравнений запишется следующим образом

$$X^T \cdot X \cdot B = X^T Y \quad (3.2)$$

где X^T – матрица, транспонированная к матрице независимых переменных;

X – матрица независимых переменных;

B – матрица-столбец коэффициентов уравнения;

Y – матрица-столбец вектора наблюдения.

Представим исходный материал в матричной форме и будем называть матрицу

$$X = \begin{vmatrix} x_{01} & x_{11} & x_{21} & \dots & x_{k1} \\ x_{02} & x_{12} & x_{22} & \dots & x_{k2} \\ \dots & \dots & \dots & \dots & \dots \\ x_{0N} & x_{1N} & x_{2N} & \dots & x_{kN} \end{vmatrix} \quad (3.3)$$

матрицей независимых переменных, а матрицу-столбец

$$Y = \begin{vmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{vmatrix} \quad (3.4)$$

вектором наблюдения. Введем матрицу-столбец

$$B = \begin{vmatrix} b_0 \\ b_1 \\ \dots \\ b_k \end{vmatrix} \quad (3.5)$$

и матрицу, транспонированную в X

$$X^T = \begin{vmatrix} x_{01} & x_{02} & \dots & x_{0N} \\ x_{11} & x_{12} & \dots & x_{1N} \\ x_{21} & x_{22} & \dots & x_{2N} \\ \dots & \dots & \dots & \dots \\ x_{k1} & x_{k2} & \dots & x_{kN} \end{vmatrix} \quad (3.6)$$

Перемножим матрицу X на X^T получим матрицу моментов $X^T \cdot X$

$$X^T \cdot X = \begin{vmatrix} \sum_i^N x_{0i}^2 & \sum_i^N x_{0i} x_{1i} & \dots & \sum_i^N x_{0i} x_{ki} \\ \sum_i^N x_{1i} x_{0i} & \sum_i^N x_{1i}^2 & \dots & \sum_i^N x_{1i} x_{ki} \\ \dots & \dots & \dots & \dots \\ \sum_i^N x_{ki} x_{0i} & \sum_i^N x_{ki} x_{1i} & \dots & \sum_i^N x_{ki}^2 \end{vmatrix}$$

Умножив матрицу моментов $X^T \cdot X$ на матрицу-столбец B , будем иметь

$$X^T \cdot X \cdot B = \begin{vmatrix} b_0 \sum_i^N x_{0i}^2 + b_1 \sum_i^N x_{0i} x_{1i} + \dots + b_k \sum_i^N x_{0i} x_{ki} \\ b_0 \sum_i^N x_{1i} x_{0i} + b_1 \sum_i^N x_{1i}^2 + \dots + b_k \sum_i^N x_{1i} x_{ki} \\ \dots \\ b_0 \sum_i^N x_{ki} x_{0i} + b_1 \sum_i^N x_{ki} x_{1i} + \dots + b_k \sum_i^N x_{ki}^2 \end{vmatrix} \quad (3.7)$$

Затем умножим матрицу X^T на вектор наблюдения, получим

$$X^T \cdot Y = \begin{vmatrix} \sum_i^N x_{0i} \cdot y_i \\ \sum_i^N x_{1i} \cdot y_i \\ \dots \\ \sum_i^N x_{ki} \cdot y_i \end{vmatrix}.$$

Из выражения 1 определим матрицу столбец B

$$B = (X^T X)^{-1} X^T Y \quad (3.8)$$

где $(X^T X)^{-1}$ – матрица, обратная матрице моментов.

Таким образом, любой коэффициент линейной части уравнения регрессии может быть определен скалярным произведением матрицы обратной матрицы на матрицу вектор наблюдения

$$B = \begin{vmatrix} \frac{1}{\sum_i^N x_{0i}^2} & \frac{1}{\sum_i^N x_{0i} x_{1i}} & \cdots & \frac{1}{\sum_i^N x_{0i} x_{ki}} \\ \frac{1}{\sum_i^N x_{1i} x_{0i}} & \frac{1}{\sum_i^N x_{1i}^2} & \cdots & \frac{1}{\sum_i^N x_{1i} x_{ki}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{\sum_i^N x_{ki} x_{0i}} & \frac{1}{\sum_i^N x_{ki} x_{1i}} & \cdots & \frac{1}{\sum_i^N x_{ki}^2} \end{vmatrix} \times \begin{vmatrix} \sum_i^N x_{0i} \cdot y_i \\ \sum_i^N x_{1i} \cdot y_i \\ \cdots \\ \sum_i^N x_{ki} \cdot y_i \end{vmatrix}. \quad (3.9)$$

Матрица моментов может значительно быть упрощена, а следовательно, и упрощено определение коэффициентов уравнения, если использовать свойства матрицы:

1. Скалярное произведение двух и более вектор-столбцов должно быть равным нулю (свойство ортогональности матрицы). Ортогональность матрицы обеспечивается тем, что матрица коэффициентов нормальных уравнений $X^T \cdot X$ становится диагональной и её диагональные элементы будут равны числу опытов N в матрице планирования

$$\sum_i^N x_{ui} \cdot x_{ji} = 0; \quad u \neq j; \quad u, j = 0, 1, 2, \dots, k. \quad (3.10)$$

2. Сумма любого вектор-столбца должна быть равна нулю (свойство и симметричного расположения уровней относительно центра эксперимента).

$$\sum_i^N x_{ji} = 0; \quad j = 1, 2, \dots, k; \quad j \neq 0. \quad (3.11)$$

3. Сумма квадрата вектор-столбца должна быть равна числу опытов (свойство нормировки)

$$\sum_i^N x_{ji}^2 = N; \quad j = 0, 1, 2, \dots, k. \quad (3.12)$$

С учетом вышеизложенного составим план полного факторного эксперимента типа 2^3 . Имеем независимые переменные $z_1^{\max}, z_1^{\min}, z_2^{\max}, z_2^{\min}, z_3^{\max}, z_3^{\min}$, определяющие исследуемое факторное пространство.

В соответствии со вторым свойством (4) определим центр плана или основной уровень независимой переменной z_1^0, z_2^0, z_3^0

$$z_i^0 = \frac{z_i^{\max} + z_i^{\min}}{2} \quad (3.13)$$

С интервалом варьирования

$$\Delta z_i = \frac{z_i^{\max} - z_i^{\min}}{2}. \quad (3.14)$$

С целью упрощения составления плана эксперимента и последующей математической обработки закодируем натуральные значения факторов по формулам

$$x_i = \frac{z_i^{\max} - z_i^0}{\Delta z_i}; x_i = \frac{z_i^{\min} - z_i^0}{\Delta z_i}. \quad (3.15)$$

Введем в ПЭФ типа 2^3 столбец так называемой фиктивной переменной $x_0 = +1$ и составим план эксперимента (табл.3.1).

Таблица 3.1

Матрица планирования с фиктивной переменной

Номер 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
5	+1	-1	-1	+1	y_5
6	+1	+1	-1	+1	y_6
7	+1	-1	+1	+1	y_7
8	+1	+1	+1	+1	y_8

Матрица моментов ($X^T \cdot X$) соответствующая табл.3.1, имеет вид

$$X^T \cdot X = \begin{vmatrix} \sum_1^8 x_{0i}^2 & \sum_1^8 x_{0i}x_{1i} & \sum_1^8 x_{0i}x_{2i} & \sum_1^8 x_{0i}x_{3i} \\ \sum_1^8 x_{1i}x_{0i} & \sum_1^8 x_{1i}^2 & \sum_1^8 x_{1i}x_{2i} & \sum_1^8 x_{1i}x_{3i} \\ \sum_1^8 x_{2i}x_{0i} & \sum_1^8 x_{2i}x_{1i} & \sum_1^8 x_{2i}^2 & \sum_1^8 x_{2i}x_{3i} \\ \sum_1^8 x_{3i}x_{0i} & \sum_1^8 x_{3i}x_{1i} & \sum_1^8 x_{3i}x_{2i} & \sum_1^8 x_{3i}^2 \end{vmatrix} \quad (3.16)$$

Учитывая свойство 1 и 3 матрицы, получим

$$X^T \cdot X = \begin{vmatrix} 8 & 0 & 0 & 0 \\ 0 & 8 & 0 & 0 \\ 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 8 \end{vmatrix}$$

Матрица, обратная матрице моментов, получается равной

$$(X^T X)^{-1} = \begin{vmatrix} \frac{1}{8} & 0 & 0 & 0 \\ 0 & \frac{1}{8} & 0 & 0 \\ 0 & 0 & \frac{1}{8} & 0 \\ 0 & 0 & 0 & \frac{1}{8} \end{vmatrix}$$

И матрица-вектор наблюдения

$$X^T \cdot Y = \begin{vmatrix} \sum_1^8 x_{0i} \cdot y_i \\ \sum_1^8 x_{1i} \cdot y_i \\ \sum_1^8 x_{2i} \cdot y_i \\ \sum_1^8 x_{3i} \cdot y_i \end{vmatrix}$$

Таким образом,

$$B = \begin{pmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{8} & 0 & 0 & 0 \\ 0 & \frac{1}{8} & 0 & 0 \\ 0 & 0 & \frac{1}{8} & 0 \\ 0 & 0 & 0 & \frac{1}{8} \end{pmatrix} \times \begin{pmatrix} \sum_{i=1}^8 x_{0i} \cdot y_i \\ \sum_{i=1}^8 x_{1i} \cdot y_i \\ \sum_{i=1}^8 x_{2i} \cdot y_i \\ \sum_{i=1}^8 x_{3i} \cdot y_i \end{pmatrix}. \quad (3.17)$$

Следовательно, процесс определения численных значений коэффициентов уравнения заключается в перемножении столбца y_i на соответствующий столбец x_i и деленным на число опытов в матрице планирования N :

$$b_i = \frac{1}{N} \sum_{i=1}^N x_{ji} \cdot y_i. \quad (3.18)$$

Если по ходу исследования необходимо установить влияние парных, тройных и т.п. взаимодействий, то в матрицу планирования вводят дополнительные столбцы: x_1x_2 ; x_1x_3 ; x_2x_3 ; $x_1x_2x_3$ и т.д.

Для определения значения коэффициента необходимо, аналогично определению коэффициентов линейной части, умножить соответствующий столбик матрицы планирования на столбец y_i .

Рассмотренную методику построения математической модели применим при изучении выхода продукции y , %. На выход влияют три фактора: температура (z_1) в диапазоне $100 \dots 200^\circ \text{C}$; давление (z_2) $2 \dots 6 \times 10^6 \text{ Па}$, время пребывания z_3 $10 \dots 20$ мин.

Находим для первого фактора центр 150°C , интервал варьирования $\Delta z_1 = 50$ и кодовые значения $+1$ и -1 .

Аналогично производим расчеты и для z_2 и z_3 . Результаты заносим в матрицу планирования типа 2^3 . Число возможных комбинаций N из трех факторов на двух уровнях равно 8.

Таблица 3.2

Матрица планирования полного факторного эксперимента 2^3 .

Факторы в натуральном масштабе				Факторы в безразмерной системе				
Номер опыта	z_1	z_2	z_3	x_0	x_1	x_2	x_3	y
1	100	20	10	+1	-1	-1	-1	2
2	200	20	10	+1	+1	-1	-1	3
3	100	60	10	+1	-1	+1	-1	4
4	200	60	10	+1	+1	+1	-1	8
5	100	20	20	+1	-1	-1	+1	10
6	200	20	20	+1	+1	-1	+1	18
7	100	60	20	+1	-1	+1	+1	8
8	200	60	20	+1	+1	+1	+1	12

Первый опыт ставим следующим образом: температура 100^0C , давление 2×10^5 Па, время 10 мин. Выход продукции равен 2.

Шестой опыт: температура 200^0C давление 2×10^5 Па, время 20 мин.

Выход продукции равен 18 и т.д. по всем 8 опытам.

Определяем коэффициенты уравнения

$$b_0 = \frac{\sum_{i=1}^8 x_0 y_i}{8} = \frac{-1 \times 2 + 1 \times 6 + 1 \times 4 - 1 \times 8 + 1 \times 10 + 1 \times 18 + 1 \times 8 + 1 \times 12}{8} = 8,5$$

$$b_1 = \frac{\sum_{i=1}^8 x_1 y_i}{8} = \frac{-1 \times 2 + 1 \times 6 - 1 \times 4 - 1 \times 8 - 1 \times 10 + 1 \times 18 - 1 \times 8 + 1 \times 12}{8} = 2,5$$

$$b_2 = \frac{\sum_{i=1}^8 x_2 y_i}{8} = \frac{-1 \times 2 - 1 \times 6 + 1 \times 4 + 1 \times 8 + 1 \times 10 - 1 \times 18 + 1 \times 8 + 1 \times 12}{8} = -0,5$$

$$b_3 = \frac{\sum_{i=1}^8 x_3 y_i}{8} = \frac{-1 \times 2 - 1 \times 6 - 1 \times 4 - 1 \times 8 + 1 \times 10 + 1 \times 18 + 1 \times 8 + 1 \times 12}{8} = 3,5$$

и т.д. Математическая модель будет иметь вид

$$y = 8,5 + 2,5x_1 - 0,5x_2 + 3,5x_3.$$

С учетом парных и тройных взаимодействий факторов уравнение примет вид:

$$y = 8,5 + 2,5x_1 - 0,5x_2 + 3,5x_3 - 0,5x_1x_2 - 1,5x_2x_3 + 0,5x_1x_3 - 0,5x_1x_2x_3.$$

Коэффициенты парных и тройных взаимодействий получают перемножением столбцов x_i между собой и на y_i :

$$b_{12} = \frac{(-1)(-1) \cdot 2 + (+1)(-1) \cdot 6 + (-1)(-1) \cdot 4 + (-1)(+1) \cdot 8 + (-1)(-1) \cdot 10 + (+1)(-1) \cdot 18 + (-1)(+1) \cdot 8 + (+1)(+1) \cdot 12}{8} \\ b_{12} = -0,5.$$

Аналогично

$$b_{23} = \frac{2 + 6 - 4 - 8 - 10 - 18 + 8 + 12}{8} = -1,5;$$

$$b_{13} = \frac{2 - 6 + 4 - 8 - 10 + 18 - 8 + 12}{8} = 0,5;$$

$$b_{123} = \frac{-2 + 6 + 4 - 8 + 10 - 18 - 8 + 12}{8} = -0,5.$$

Из полученного уравнения можно видеть, что для повышения % выхода продукта следует факторы x_1 и x_3 увеличивать, а фактор x_2 уменьшать.

Из парных взаимодействий наибольшее влияние оказывают факторы x_2 и x_3 .

3.3 Анализ уравнения регрессии

Полученное уравнение регрессии необходимо исследовать с тем, чтобы по возможности упростить и проверить, с какой точностью найденное уравнение описывает исследуемый процесс, а так же наметить стратегию дальнейшего исследования.

Определение значимости коэффициентов уравнения регрессии.

Найденные частные значения коэффициентов уравнения следует проанализировать не только качественно, но и количественно. С этой целью, используя математические приёмы, производят разделение факторов на существенные и несущественные. При дальнейшем планировании эксперимента несущественные факторы исключают, т.к. они в выбранном факторном пространстве не оказывают существенного влияния на процесс. Если нам необходимо знать их влияние, следует сократить факторное пространство по направлению этого фактора.

С целью установления значимости коэффициентов уравнения нужно поставить дополнительно параллельные опыты в центре плана.

В связи с тем, что ковариационная матрица $(X^T; X)^{-1}$ для спланированного эксперимента – матрица диагональная, коэффициенты уравнения некоррелированы между собой.

Следовательно, значимость коэффициентов уравнения регрессии можно проверить для каждого коэффициента в отдельности по критерию **Стьюдента**.

Исключение из уравнения регрессии незначимого коэффициента не скажется на остальных коэффициентах.

Диагональные элементы ковариационной матрицы равны между собой, поэтому все коэффициенты уравнений определяются с одинаковой точностью:

$$S_{bj} = \frac{S_{\text{воспр}}}{\sqrt{N}}. \quad (3.19)$$

Для нахождения среднеквадратичной ошибки опыта поставим в центре плана несколько дополнительных опытов и получим значения y_u , тогда

$$\bar{y}_i^c = \frac{\sum y_u^c}{n} \text{ и } S_{\text{воспр}}^2 = \frac{\sum_i^n (y_u^c - \bar{y}_i^c)^2}{n-1}, \text{ или } S_{\text{воспр}} = \sqrt{S_{\text{воспр}}^2}.$$

Отсюда значимость коэффициентов по критерию Стьюдента

$$t_i = \frac{|b_i|}{S_{bj}}. \quad (3.20)$$

Сравнивая табличные значения критерия Стьюдента с полученными значениями t_i , находим несущественные, т.е. имеющие меньшие значения.

Например, были поставлены в центре плана опыты и получены следующие значения y : $y_1^c = 8$; $y_2^c = 9$; $y_3^c = 8,8$.

$$\bar{y}^k = \frac{8+9+8,8}{3} = 8,6; \quad S_{\text{воспр}}^2 = \frac{(8-8,6)^2 + (9-8,6)^2 + (8,8-8,6)^2}{3-1} = 0,28;$$

$$S_{\text{воспр}} = \sqrt{0,28} = 0,55; \quad S_{bj} = \frac{0,55}{\sqrt{8}} = 0,2.$$

Значимость коэффициентов по критерию Стьюдента

$$t_0 = \frac{8,5}{0,2} = 42,5; \quad t_1 = \frac{2,5}{0,2} = 12,5; \quad t_2 = \frac{0,5}{0,2} = 2,5;$$

$$t_3 = \frac{3,5}{0,2} = 17,5; \quad t_{12} = \frac{0,5}{0,2} = 2,5; \quad t_{13} = \frac{0,5}{0,2} = 2,5;$$

$$t_{23} = \frac{1,5}{0,2} = 7,5; \quad t_{123} = \frac{0,5}{0,2} = 2,5.$$

Табличные значения критерия Стьюдента для уровня значимости $p=0,05$ и числа степеней свободы $f = 3-1=2$ будет $t_{0,05(2)} = 4,3$. Таким образом, коэффициенты b_0, b_1, b_3 и b_{23} значимы и уравнение примет вид

$$y = 8,5 + 2,5x_1 + 3,5x_3 - 1,5x_2x_3.$$

Проверка уравнения регрессии на адекватность.

Проверяют уравнения регрессии на адекватность по критерию **Фишера**:

$$F = \frac{S_{\text{ост.}}^2}{S_{\text{воспр.}}^2}; \quad S_{\text{ост.}}^2 = \frac{\sum_i^N (y_i - \hat{y}_i)^2}{N-l},$$

где l – число значимых коэффициентов в уравнении регрессии;

Если найденное значение F будет меньше табулированного значения критерия Фишера $F_{1-p}(f_1, f_2)$, т.е. $F < F_{1-p}(f_1, f_2)$, то полученное уравнение регрессии адекватно описывает исследуемый процесс.

Например, имеем

$$S_{\text{инд.}}^2 = \frac{(1-2)^2 + (6-6)^2 + (5-4)^2 + (8-8)^2 + (0-10)^2 + (10-18)^2 + (9-9)^2 + (13-12)^2}{8-4} = 2.$$

Тогда $F = \frac{2,0}{0,28} = 7,2$. Табулированное значение критерия Фишера для

$$p = 0,05, f_1 = 4, f_2 = 2, F_{0,05(4,2)} = 19,3 > F = 7,2.$$

Следовательно, исследуемое уравнение адекватно описывает проведенный эксперимент и может быть использовано для интерпретации протекания процесса.

3.4 Дробные реплики факторного планирования

При планировании и реализации опыта перед исследователем возникает проблема – по возможности уменьшить объем работы без снижения ее точности. Особенно эта проблема стоит остро перед исследователем на первых этапах исследования, когда нет информации о значимости факторов. Исследователь вынужден включать в эксперимент все факторы. Так, при 4-факторном эксперименте на двух уровнях требуется реализовать 8 опытов, при 3-факторном – 16, 5-ти – 32 и т.д.

Вместе с тем, при наложении определенных ограничений (использование только линейной части уравнения, отказ от проверки адекватности и некоторых других) можно использовать не полный факторный план, а часть его, так называемую **дробную реплику**.

Для того чтобы дробная реплика соответствовала требованию ортогональности, т.е. чтобы сумма произведений строк и столбцов матрицы планирования была равна нулю, необходимо брать полный факторный план, но для меньшего числа факторов. При этом число опытов должно быть больше или равно числу неизвестных коэффициентов уравнения регрессии

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + b_3x_3. \quad (3.21)$$

Допустим, что нам необходимо описать факторное пространство четырех неизвестных факторов. Полный факторный эксперимент типа 2 будет

содержать 16 опытов. Можно сократить число опытов наполовину, взяв 1/2 реплика полного факторного эксперимента. Однако в этом случае не будет выполнено основное требование факторного планирования – условие ортогональности матрицы, т.к. ее диагональные элементы не будут равны количеству опытов. Поэтому, в данном случае, следует принять полный факторный план типа 2^3 , а фактор X_4 приравнять любому парному или тройному взаимодействию (X_1X_2, X_1X_3, X_2X_3 или $X_1X_2X_3$).

Возникает вопрос – к какому взаимодействию приравнять X_4 ?

Рассмотрим два случая приравнивания X_4 .

1. Переменный фактор x_4 приравняем к парному взаимодействию, например, к X_1X_2 , т.е. $X_4 = X_1X_2$ или к любой другой паре эффектов. Такое соотношение (приравнивание) факторов называется **генерирующим соотношением**.

Умножим обе части генерирующего соотношения на X_4 , будем иметь $X_4^2 = X_1X_2X_4$, т.к. $X^2 = 1$, то получим $1 = X_1X_2X_4$. Полученное произведение называется **определяющим контрастом**.

Воспользуемся определяющим контрастом и получим систему совместных оценок для коэффициентов уравнения регрессии. С этой целью умножим поочередно определяющий контраст на факторы X_1, X_2, X_3, X_4 .

$X_1 = X_2X_4$	$b_1 \rightarrow \beta_1 + \beta_{24}$
$X_2 = X_1X_4$	$b_2 \rightarrow \beta_2 + \beta_{14}$
$X_3 = X_1X_2X_3X_4$	$b_3 \rightarrow \beta_3 + \beta_{1234}$
$X_4 = X_1X_2$	$b_4 \rightarrow \beta_4 + \beta_{12}$
$X_1X_3 = X_2X_3X_4$	$b_{13} \rightarrow \beta_{13} + \beta_{234}$
$X_2X_3 = X_1X_3X_4$	$b_{23} \rightarrow \beta_{23} + \beta_{134}$
$X_3X_4 = X_1X_2X_3$	$b_{34} \rightarrow \beta_{34} + \beta_{123}$

В реальных задачах тройные взаимодействия обычно имеют значения близкими к нулю, поэтому ими пренебрегают и при анализе уравнения не учитывают. Линейные эффекты смешаны с двойными взаимодействиями, кроме X . Следовательно, принятая дробная реплика не дает возможности получить несмешанную оценку линейным эффектам.

2. Переменный фактор приравняем к тройному взаимодействию $X_4 = X_1 X_2 X_3$.

Определяющий контраст $1 = X_1 X_2 X_3 X_4$. Система совместных оценок для коэффициентов уравнения регрессии будет иметь вид

$$\begin{aligned}
 X_1 &= X_2 X_3 X_4 & b_1 &\rightarrow \beta_1 + \beta_{234} \\
 X_2 &= X_1 X_3 X_4 & b_2 &\rightarrow \beta_2 + \beta_{134} \\
 X_3 &= X_1 X_2 X_4 & b_3 &\rightarrow \beta_3 + \beta_{124} \\
 X_4 &= X_1 X_2 X_3 & b_4 &\rightarrow \beta_4 + \beta_{123} \\
 X_1 X_2 &= X_3 X_4 & b_{12} &\rightarrow \beta_{12} + \beta_{34} \\
 X_1 X_3 &= X_2 X_4 & b_{13} &\rightarrow \beta_{13} + \beta_{24} \\
 X_1 X_4 &= X_2 X_3 & b_{14} &\rightarrow \beta_{14} + \beta_{23}
 \end{aligned}$$

В приведенной дробной реплике обеспечиваются несмешанные оценки для линейных эффектов. В таблице 3 и таблице 4 приведены рассмотренные полуреплики.

Таблица 3.3

Полуреплика от ПФЭ 2^4 с генерирующим соотношением $X_4 = X_1 X_2$.

Номер опыта	X_0	X_1	X_2	X_3	X_4	Номер опыта	X_0	X_1	X_2	X_3	X_4
1	+1	+1	+1	+1	+1	5	+1	+1	+1	-1	+1
2	+1	-1	-1	+1	+1	6	+1	-1	-1	-1	+1
3	+1	-1	+1	+1	-1	7	+1	-1	+1	-1	-1
4	+1	+1	-1	+1	-1	8	+1	+1	-1	-1	-1

Таблица 3.4

Полуреплика от ПФЭ 2^4 с генерирующим соотношением $X_4 = X_1 X_2 X_3$.

Номер опыта	X_0	X_1	X_2	X_3	X_4	Номер опыта	X_0	X_1	X_2	X_3	X_4
1	+1	+1	+1	+1	+1	5	+1	+1	+1	-1	+1
2	+1	-1	-1	+1	+1	6	+1	-1	-1	-1	-1
3	+1	-1	+1	+1	-1	7	+1	-1	+1	-1	+1
4	+1	+1	-1	+1	-1	8	+1	+1	-1	-1	+1

Дробную реплику, в которой p линейных эффектов приравнены к эффектам взаимодействия, обозначают 2^{k-p} . Для четверть реплики, например, в

планировании для $K=5$ тип 2^{5-2} могут быть заданы генерирующие соотношения

$$X_4 = X_1 X_2 X_3, X_5 = X_1 X_2.$$

Число опытов в дробной реплике должно удовлетворять неравенству

$$K+1 \leq N \leq 2^K,$$

где K – число факторов для получения несмешанных оценок линейных эффектов.

Если число опытов N равно $K+1$ – числу определяемых коэффициентов в линейном уравнении регрессии, то дробная реплика представляет собой **насыщенный линейный ортогональный план**.

Поскольку число опытов в насыщенных планах равно или близко числу определяемых коэффициентов, число степеней остаточной дисперсии равно нулю. Для проверки адекватности линейного уравнения, полученного по насыщенному плану, требуется дополнительный эксперимент.

Представленный выше план факторного эксперимента не требует повторности опытов. Для анализа полученного уравнения регрессии достаточно поставить несколько опытов в центре плана.

При планировании факторного эксперимента с M – повторностями схема составления плана будет следующая.

Составим матрицу планирования с m – повторностями (таблица 3.5).

Таблица 3.5

Матрица планирования и результаты измерения

Номер опыта	X_0	X_1	...	X_k	y_{iu}	$\bar{y}_i$	S_i^2
1	+1	+1	...	+1	$y_{11}, y_{12}, \dots, y_{1m}$	$\bar{y}_1$	S_1^2
2	+1	-1	...	+1	$y_{21}, y_{22}, \dots, y_{2m}$	$\bar{y}_2$	S_2^2
3	-1	+1	...	+1	$y_{31}, y_{32}, \dots, y_{3m}$	$\bar{y}_3$	S_3^2
.	.	.	...	.		.	.
.	.	.	...	.		.	.
.	.	.	...	.		.	.
N	+1	-1	...	-1	$y_{N1}, y_{N2}, \dots, y_{Nm}$	$\bar{y}_N$	S_N^2

Находим среднее значение измеряемой величины $\bar{y}_i$ по m параллельным опытам

$$\bar{y}_i = \frac{\sum_{i=1}^m y_{im}}{m}, \quad i = 1, 2, \dots, N, \quad (3.22)$$

и дисперсия

$$S_i^2 = \frac{\sum_{u=1}^m (y_{iu} - \bar{y}_i)^2}{m-1}, \quad i = 1, 2, \dots, N. \quad (3.23)$$

Затем проверяется однородность выборочных дисперсий по критерию Кохрана. С этой целью составляется отношение максимальной дисперсии к сумме всех дисперсий:

$$G = \frac{S_{\max}^2}{\sum_{i=1}^N S_i^2}. \quad (3.24)$$

Полученное отношение сравнивается с табличным: $G_{1-p}(f_1, f_2)$ при $p=0,05$; где p – уровень значимости; f_1 и f_2 степени свободы равные $f_1 = m-1$, $f_2 = N$. Если $G < G_{1-p}(f_1, f_2)$, дисперсии однородны.

Тогда в качестве оценки дисперсии воспроизводимости можно использовать среднюю дисперсию

$$S_{\hat{\sigma}}^2 = \frac{\sum_{i=1}^N S_i^2}{f_{\hat{\sigma}}}. \quad (3.25)$$

С числом степеней свободы $f_{\hat{\sigma}} = N(m-1)$.

Коэффициенты уравнения регрессии определяются по формуле

$$b_i = \frac{\sum_{i=1}^N x_{ij} \bar{y}_i}{N} \quad (3.26)$$

Дисперсия коэффициентов $\bar{y}_i$ определяется следующим образом

$$S_{b_j}^2 = \frac{S_{\hat{\sigma}}^2}{N \cdot m} \quad (3.27)$$

Значимость коэффициентов проверяется по критерию Стьюдента. С этой целью для всех коэффициентов уравнения составляется t_j – отношение

$$t_j = \frac{|b_j|}{S_{bj_j}}, \quad (3.28)$$

которое сравнивается с табличным $t_{1-p(f)}$ уровня значимости p и числа степеней свободы $f = N(m-1)$.

Адекватность уравнения регрессии эксперименту проверяется по критерию Фишера. С этой целью составляется дисперсионное отношение

$$F = \frac{S_{ay}^2}{S_{воспр}^2}, \quad (3.29)$$

где S_{ay}^2 – дисперсия адекватности, определяемая формулой

$$S_{ay}^2 = \frac{m \sum_{i=1}^N (\bar{y}_i - \hat{y}_i)^2}{N - l}, \quad (3.30)$$

где l – число значимых коэффициентов в уравнении регрессии.

Если полученное дисперсионное отношение будет меньше табличного:

$$F < F_{i-p(f_1, f_2)} \quad (3.31)$$

где f_1 – число степеней свободы дисперсии адекватности, $f_1 = N - l$;

f_2 – число степеней свободы дисперсии, $f_2 = N(m-1)$, уравнение регрессии адекватно исследуемому процессу.

Если $F > F_{i-p(f_1, f_2)}$, то для адекватного описания процесса необходимо увеличить порядок аппроксимирующего полинома, т.е. использовать квадратичные члены.

3.5 Оптимизация методом крутого восхождения

по поверхности отклика

Полученное уравнение регрессии дает математическую модель протекания изучаемого процесса. Из уравнения мы можем сделать вывод о влиянии

янии того или иного фактора на процесс как с качественной (знак коэффициента уравнения), так и с количественной (численное значение коэффициента уравнения) сторон. Вместе с тем, в большинстве случаев ставится более узкая задача - установить оптимальные значения переменных независимых с тем, чтобы максимальный (экстремальный) выход продукции

$$y_{\max} = f(x_1^{onm}, x_2^{onm}, \dots, x_k^{onm}) \quad (3.32)$$

При традиционной постановке эксперимента экстремальные точки находят путем последовательного перемещения опыта по координатным осям (рисунок 3.2).

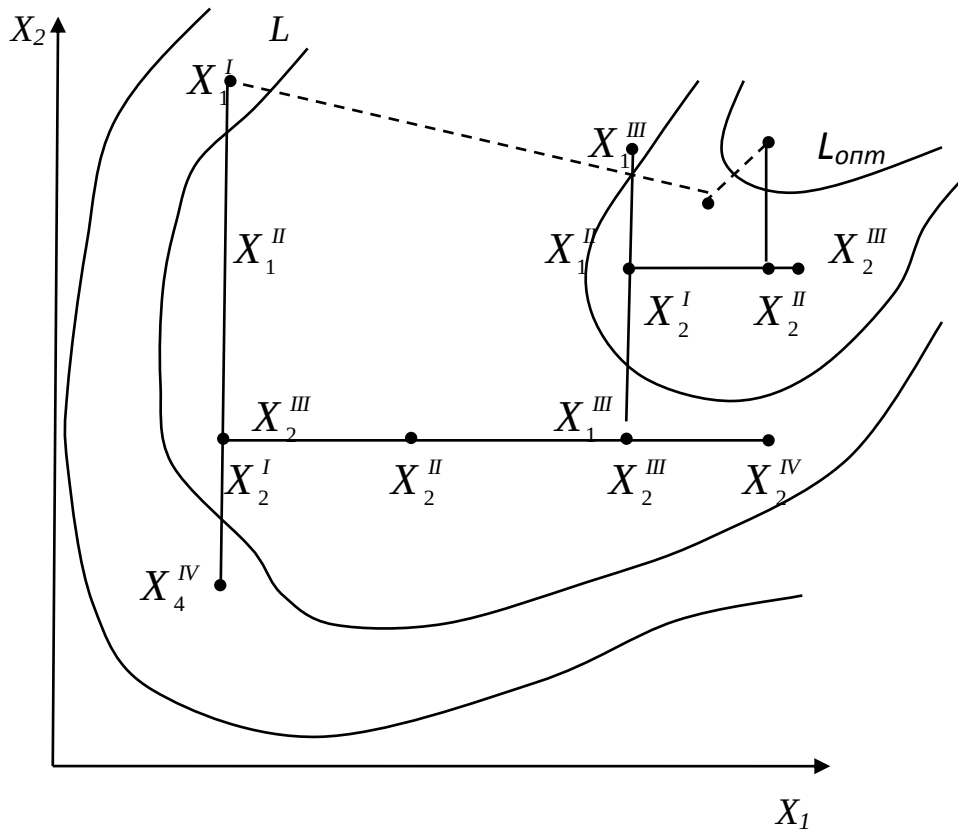


Рис. 3.2. Движение по поверхности отклика к экстремуму в однофакторном эксперименте (сплошная линия) и в методе крутого восхождения (пунктирная линия)

Эту же точку $L_{\text{опт}}$ можно достичь, если двигаться перпендикулярно линиям $y=\text{const}$, т.е. осуществить движение по градиенту

$$\text{grad}f = \frac{\partial f}{\partial x_i} \vec{i} + \frac{\partial f}{\partial x_j} \vec{j} + \dots + \frac{\partial f}{\partial x_k} \vec{k}, \quad (3.33)$$

где $\vec{i}, \vec{j}, \vec{k}$ – единичные векторы в направлении координатных осей. При этом предполагается, что функция f непрерывна, однозначна и не имеет особых точек.

Из приведенного рисунка видно, что для достижения экстремальной точки L_{opt} при традиционной постановке опыта требуется значительно больше опытов, чем при крутом восхождении.

Для нахождения экстремальной точки L_{opt} используют линейную часть уравнения регрессии,

$$y = b_0 + b_1 x_1 + \dots + b_k x_k.$$

Далее двигаются по поверхности отклика в направлении градиента линейного приближения

$$\frac{\partial \hat{y}}{\partial x_1} = b_1; \frac{\partial \hat{y}}{\partial x_2} = b_2; \dots \frac{\partial \hat{y}}{\partial x_k} = b_k. \quad (3.34)$$

Величина изменения переменного при движении (величина шага) должна быть пропорциональна произведению коэффициента уравнения b_j на интервал варьирования Δb_j , уменьшенному или увеличенному в n раз.

Полученное произведение $b_j \Delta b_j$, уменьшенное или увеличенное должно быть таким, чтобы можно было поставить 4...5 опытов в пределах факторного пространства.

Для большинства процессов линейное приближение поверхности отклика адекватно эксперименту только при небольших интервалах варьирования. Факторы, оказывающие незначительное влияние на процесс, в шаговом эксперименте не участвуют. Их фиксируют на нулевом (основном) уровне. Движение по градиенту начинают от основного уровня в направлении знака коэффициента уравнения и прекращают при переходе точки максимального выхода продукции.

При шаговом движении по градиенту можно часть опытов не реализовывать. Это так называемые «мысленные» опыты. Если же оптимум находится в зоне «мысленного» опыта, то его следует реализовать.

Рассмотрим вышеизложенную методику планирования эксперимента на следующем примере.

Требуется определить оптимальные условия получения максимальной сплошности покрытия в % при электроимпульсном наращивании. На процесс электроимпульсного наращивания, априорно установлено, влияют 7 параметров (таблица 3.6) Не указанные в таблице 6 параметры относятся, предположительно, к шумовому эффекту.

Таблица 3.6

Независимые переменные и уровни их варьирования.

Кодовое название	Наименование фактора	Уровни варьирования			Интервал варьирования ΔZ_i
		Нижний $Z_i^{\min}$	Основной $Z_i^{\text{осн}}$	Верхний $Z_i^{\max}$	
X_1	Частота вибрации электрода, Гц	40	70	100	30
X_2	Скорость перемещения детали, мм/с	1,8	2,3	2,8	0,5
X_3	Напряжение наращивания, В	4	12	20	8
X_4	Высота слоя жидкости в ванне, мм	50	70	90	20
X_5	Отношение периода холостого хода к полному циклу	1/4	2/4	3/4	1/4
X_6	Межэлектродный зазор, мм	0,5	1,1	1,7	0,6
X_7	Содержание хромового ангидрида и серной кислоты в охлаждающей жидкости, кг/м ²	100+1	200+2	300+3	100+1

Постановка полного факторного эксперимента типа 2^7 при 3 – кратной повторности потребовало бы $128 \times 3 = 384$ опыта. Поэтому на первом этапе планируем использовать дробную реплику от полного факторного эксперимента. В результате реализации дробной реплики мы можем установить, какое влияние оказывают на процесс взятые на «подозрение» факторы, подразделить их на значимые и незначимые. Затем, на 2-м этапе эксперимента спланировать полный факторный эксперимент при меньшем числе независимых переменных.

При составлении плана эксперимента на первом этапе была использована 1/16 реплика от ПФЭ 2^7 с генерирующими соотношениями

$$X_4 = X_1 X_2, X_5 = X_2 X_3, X_6 = X_1 X_3, X_7 = X_1 X_2 X_3.$$

В связи с тем, что 1/16 реплика от ПФЭ 2^7 представляет собой насыщенный линейный ортогональный план, все линейные эффекты смешаны с эффектами взаимодействия, что затрудняют анализ уравнения.

При составлении матрицы планирования произвели переход от натуральных значений независимых переменных к кодовым значениям.

$$x_1^{\min} = \frac{40 - 70}{30} = -1; \quad x_1^{\max} = \frac{100 - 70}{30} = 1.$$

Аналогично производим перевод других факторов и результаты заносим в таблица 3.7.

Таблица 3.7

Матрица плана 1/16 ПФЭ 2^7 с результатами опытов

№ Опы та	Независимые переменные								Сплошность, %			Среднее, $\bar{y}$
	X_0	X_1	X_2	X_3	X_4	X_5	X_6	X_7	Y_1	Y_2	Y_3	
1	+	2	+	-	+	-	+	-	4	8	3	5
2	+	+	+	-	-	+	-	-	25	30	26	27
3	+	-	-	-	-	+	+	+	6	15	9	10
4	+	+	-	-	+	-	-	+	28	30	38	32
5	+	-	+	+	-	-	-	+	49	58	61	56
6		+	+	+	+	+	+	+	75	79	86	80
7	+	-	-	+	+	+	-	-	49	57	56	54
8	+	+	-	+	-	-	+	-	80	83	90	86

43,75 -1,75 -1,0 1,5
12,5 25,25 1,0 -0,75

Выполнение опытов в эксперименте было рандомизировано, повторность трехкратная. Рандомизацию производили по таблице случайных чисел.

Анализ результатов эксперимента начинаем с проверки однородности дисперсии опытных данных по критерию Кохрена

$$G = \frac{\sum_i^m S_{\max}^2}{\sum_i \sum_i S_i^2}, \quad (3.35)$$

где $S_{\max}^2$ - максимальная дисперсия i -го опыта по m повторностям; $\sum_i^N \sum_i^m S_i^2$ - дисперсия эксперимента по N опытам и m повторностям.

$$\sum_i^m S_{\max}^2 = (80-86)^2 + (83-80)^2 + (95-80)^2 = 126$$

$$\sum_i^N \sum_i^m S_{ij}^2 = (4-5)^2 + (8*5)^2 + (3*5)^2 + \dots + (80-86)^2 + (83-80)^2 + (95-80)^2 = 420$$

$$G = \frac{126}{420} = 0,2857$$

Табличное значение критерия Кохрена для уровня значимости $p = 0,05$ и чисел степеней свободы $f_1 = m - 1 = 3 - 1 = 2$, $f_2 = N = 8$

$$G_{1,05(2,8)} = 0,5157 > G = 0,2857,$$

и следовательно, дисперсии однородны. Все опытные данные принадлежат одной генеральной выборке.

Коэффициенты уравнения регрессии определяем по формуле
 $b_0=43,75$; $b_1=12,5$; $b_2=-1,75$; $b_3=25,25$; $b_4=-1,0$; $b_5=1,0$; $b_6=1,5$; $b_7=-0,75$.

Оценим значимость полученных коэффициентов уравнения по критерию Стьюдента. Для этого определим ошибку коэффициентов уравнения

$$S_{bj} = \sqrt{\frac{S_{\text{воспр}}^2}{N}}; \quad S_{\text{воспр}}^2 = \frac{\sum_i \sum_i S_{ij}}{N(m-1)} = \frac{420}{16} = 26,25; \quad S_{bj}=1,81$$

где $S_{\text{воспр}}^2$ - дисперсия воспроизводимости; f - число степеней свобода дисперсии воспроизводимости, равное $f_{\text{воспр}} = N(m-1) = 8(3-1) = 16$.

Составим t - отношение для всех коэффициентов уравнения

$$t_0 = \frac{43,75}{1,81} = 24,17; \text{ аналогично } t_1 = 6,91; \quad t_2 = 0,97; \quad t_3 = 13,95; \quad t_4 = 0,55; \quad t_5 = 0,55; \\ t_6 = 0,3; \quad t_7 = 0,41.$$

Табличное значение критерия Стьюдента $t_{0,05(3)} = 2,31$ Следовательно, коэффициенты b_2, b_4, b_5, b_6, b_7 в исследуемом факторном пространстве незначимы, т.к. составленные для них t - отношения меньше табличного.

Уравнение регрессии будет иметь вид

$$Y=43,75+12,5X_1+25,25X_3.$$

Проверим адекватность полученного уравнения регрессии эксперименту по критерию Фишера. С этой целью по полученной формуле уравнения определим сплошность покрытия в каждом опыте.

$$\hat{y}_1=43,75+ 12,5 (-1) + 25,25 (-1) = 6,0$$

$$\hat{y}_2=3,75 + 12,5 (+1) + 25,25 (-1) = 31,0$$

$$\hat{y}_3=43,75 + 12,5 (-1) + 25,25 (-1) = 6,0$$

$$\hat{y}_4=43,75 + 12,5 (+1) + 25,25 (-1) = 31,0$$

$$\hat{y}_5=43,75 + 12,5 (-1) + 25,25 (+1) = 56,5$$

$$\hat{y}_6=43,75 + 12,5 (+1) + 25,25 (+1) = 81,5$$

$$\hat{y}_7=43,75 + 12,5 (-1) + 25,25 (+1) = 56,5$$

$$\hat{y}_8=43,75 + 12,5 (+1) + 25,25 (+1) = 81,5$$

Затем определяем остаточную дисперсию

$$\frac{(5-6)+(27-31)+(10-6)+(32-31)+(56-56,5)+(80-81,5)+(54-56,5)+(86-81,5)}{8-3} = 4,2.$$

Тогда F – отношение равно

$$F = \frac{4,2}{16,87} = 0,248.$$

Табличное значение критерия Фишера для $p = 0,05$; $f_1 = 5$; $f_2 = 16$

$$F_{0.05(5,n)} = 2,9 > F = 0,248.$$

Следовательно, представленная математическая модель адекватно описывает исследуемый процесс и может быть использована для крутого восхождения по поверхности отклика при нахождении оптимального режима наращивания.

При крутом восхождении в «почти стационарную область» на процесс оптимизации оказывают существенное влияние два фактора: частота вибрации электрода и напряжение процесса наращивания.

В качестве шаров приняты величины произведения $b_i \cdot \Delta Z_i$ в 10 раз меньше (таблица 3.8.).

Таблица 3.8

Планирование и реализация «шагового» эксперимента

Показатели	Z_1	Z_3	y_i
Основной уровень	70	12	
Интервал варьирования	30	8	
Произведение	375	202	
Шаг опыта	4	0,12	
Опыт 9	47	14	Не реализован
Опыт 10	78	16	89
Опыт 11	82	18	91
Опыт 12	86	20	79

Остальные независимые переменные были зафиксированы на основных уровнях. Из 4-х планируемых опытов были реализованы 3. Наибольшую сплошность покрытия получили во II опыте.

Таким образом, оптимальным режимом электроимпульсного наращивания с максимальной сплошностью покрытия следует считать:

1. Частоту вибрации электрода 82 Гц;
2. Скорость перемещения детали 2,3 мм/с;
3. Напряжение наращивания 18 В;
4. Высоту слоя жидкости в ванне 70 мм;
5. Отношение периода холостого хода к полному циклу 2/4;
6. Межэлектродный зазор 1,1 мм;
7. Содержание хромового ангидрида и серной кислоты в охлаждающей жидкости 200+2,0 кг/м³.

Уравнение регрессии в натуральных значения переменных имеет вид

$$C_n = 43,75 + 12,5 \left(\frac{v - 70}{-30} \right) + 25,25 \left(\frac{u - 2}{8} \right) = -23,29 + 0,42v + 3,16u$$

Таким образом, был найден оптимальный режим электроимпульсного наращивания путем постановки 27 опытов, представлена математическая модель процесса и произведен математический анализ.

ПРИЛОЖЕНИЯ

Приложение 1

Таблица значений функции $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{z^2}{2}} dx$

целые и де- сятые доли x	Сотые доли x									
	0	1	2	3	4	5	6	7	8	9
0,0	0,0000	00399	00798	01197	01595	01994	02392	02790	03188	03586
0,1	03983	04380	04776	05172	05567	05962	06356	06749	07142	07535
0,2	07926	08317	08706	09095	09483	09871	10257	10642	11026	11409
0,3	11791	12172	12552	12930	13307	13683	14058	14431	14803	15173
0,4	15542	15910	16276	16640	17003	17364	17724	18082	18439	18793
0,5	19146	19497	19847	20194	20540	20884	21226	21566	21904	22240
0,6	22575	22907	23237	23565	23891	24215	24537	24857	25175	25490
0,7	25804	26115	26424	26730	27035	27337	27637	27935	28230	28524
0,8	28814	29103	29389	29673	29955	30234	30511	30785	31057	31327
0,9	31594	31859	32121	32381	32639	32894	33147	33398	33646	33891
1,0	34134	34375	34614	34850	35083	35314	35543	35769	35993	36214
1,1	36433	36650	36864	37076	37286	37493	37698	37900	38100	38298
1,2	38493	38686	38877	39065	39251	39455	39617	39796	39973	40147
1,3	40320	40490	40658	40824	40988	41149	41309	41466	41621	41774
1,4	41924	42073	42220	42364	42507	42647	42786	42922	43056	43189
1,5	43319	43448	43574	43699	43822	43943	44062	44179	44295	44408
1,6	44520	44630	44738	44845	44950	45053	45154	45254	45352	45449
1,7	45543	45637	45728	45818	45907	45994	46080	46164	46246	46327
1,8	46407	46485	46562	46638	46712	46784	46856	46926	46995	47062
1,9	47128	47193	47257	47320	47381	47441	47500	47558	47615	47670
2,0	47725	47778	47831	47882	47932	47982	48030	48077	48124	48169
2,1	48214	48257	48300	48341	48382	48422	48461	48500	48537	48574
2,2	48610	48645	48679	48713	48745	48778	48809	48840	48870	48899
2,3	48928	48956	48983	49010	49036	49061	49086	49111	49134	49158
2,4	49180	49202	49224	49245	49266	49286	49305	49324	49343	49361
2,5	49379	49396	49413	49430	49446	49461	49477	49492	49506	49520
2,6	49534	49547	49560	49573	49585	49598	49609	49621	49632	49643
2,7	49653	49664	49674	49683	49693	49702	49711	49720	49728	49736
2,8	49744	49752	49760	49767	49774	49781	49788	49795	49801	49807
2,9	49815	49819	49825	49831	49836	49841	49846	49851	49856	49861

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
3,0	0,49865	3,5	0,49977	4,0	0,499968	4,5	0,4999966
3,1	0,49903	3,6	0,49984	4,1	0,499979	4,6	0,4999979
3,2	0,49931	3,7	0,49989	4,2	0,499987	4,7	0,4999987
3,3	0,49952	3,8	0,499993	4,3	0,499991	4,8	0,4999992
3,4	0,49966	3,9	0,499995	4,4	0,499995	4,9	0,4999995

Процентные точки F – распределения

$k_2 \backslash k_1$	1	2	3	4	5	6	7	8	9	10	11
1	161	200	216	225	230	234	237	239	241	242	243
2	18,51	19	19,16	19,25	19,3	19,33	19,36	19,37	19,38	19,39	19,4
3	10,13	9,55	9,28	9,12	9,01	8,94	8,88	8,84	8,81	8,78	8,76
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,93
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,78	4,74	4,70
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,03
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,63	3,60
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,34	3,31
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,13	3,10
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,97	2,94
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,86	2,82
12	4,75	3,88	3,49	3,26	3,11	3,00	2,92	2,85	2,80	2,76	2,72
14	4,60	3,74	3,34	3,11	2,96	2,85	2,77	2,70	2,65	2,60	2,56
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,45
20	4,35	3,49	3,10	2,87	2,71	2,60	2,52	2,45	2,40	2,35	2,31
24	4,26	3,40	3,01	2,78	2,62	2,51	2,43	2,36	2,30	2,26	2,22
30	4,17	3,32	2,92	2,69	2,53	2,42	2,34	2,27	2,21	2,16	2,12
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,07	2,04
50	4,03	3,18	2,79	2,56	2,40	2,29	2,20	2,13	2,07	2,02	1,98
60	3,98	3,13	2,74	2,50	2,35	2,23	2,14	2,07	2,01	1,97	1,93
70	3,94	3,09	2,70	2,46	2,30	2,19	2,10	2,03	1,97	1,92	1,88
100	3,89	3,04	2,65	2,41	2,26	2,14	2,05	1,98	1,92	1,87	1,83
200	3,86	3,02	2,62	2,39	2,23	2,12	2,03	1,96	1,90	1,85	1,81
400	3,84	2,99	2,60	2,37	2,21	2,09	2,01	1,94	1,88	1,83	1,79
∞											

Процентные точки t – распределения Стьюдента

N	P	0.750	0.900	0.950	0.975	0.990	0.995	0.999	0.9995
1		1.000	3.078	6.314	12.706	31.821	63.657	318.31	636.62
2		0.816	1.886	2.920	4.303	6.965	9.925	22.326	31.598
3		0.765	1.638	2.353	3.182	4.541	5.841	10.213	12.924
4		0.741	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5		0.727	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6		0.718	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7		0.711	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8		0.706	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9		0.703	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10		0.700	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11		0.697	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12		0.695	1.356	1.782	2.179	2.681	3.015	3.930	4.318
13		0.694	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14		0.692	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15		0.691	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16		0.690	1.337	1.746	2.120	2.583	2.981	3.686	4.015
17		0.689	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18		0.688	1.330	1.743	2.101	2.552	2.878	3.610	3.922
19		0.688	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20		0.687	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21		0.686	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22		0.686	1.321	1.721	2.074	2.508	2.819	3.505	3.792
23		0.685	0.319	0.714	2.096	2.500	2.807	3.485	3.767
24		0.685	1.318	1.711	2.064	2.492	2.797	3.427	3.745
25		0.684	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26		0.684	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27		0.684	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28		0.683	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29		0.683	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30		0.683	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40		0.681	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60		0.679	1.296	1.671	2.000	2.390	2.660	3.232	3.460
120		0.677	1.289	1.658	1.980	2.358	2.617	3.160	3.373
∞		0.674	1.282	1.645	1.960	2.326	2.576	3.090	3.291

Приложение 4

Процентные точки χ^2 - распределения

n/α	0.005	0.010	0.025	0.050	0.100	0.250	0.500
1	0.00004	0.00016	0.00098	0.00393	0.01579	0.1015	0.4549
2	0.0100	0.0201	0.0506	0.1026	0.2107	0.5754	1.386
3	0.0717	0.1148	0.2158	0.3518	0.5844	1.213	2.366
4	0.2070	0.2971	0.4844	0.7107	1.064	1.923	3.357
5	0.4117	0.5543	0.8312	1.145	1.610	2.675	4.351
6	0.6757	0.8721	1.2373	1.635	2.204	3.455	5.348
7	0.9893	1.239	1.690	2.167	2.833	4.255	6.346
8	1.344	1.646	2.180	2.733	3.490	5.071	7.344
9	1.735	2.088	2.700	3.325	4.168	5.899	8.343
10	2.156	2.558	3.247	3.940	4.865	6.737	9.342
11	2.603	3.053	3.816	4.575	5.578	7.584	10.34
12	3.074	3.571	4.404	5.226	6.304	8.438	11.34
13	3.565	4.107	5.009	5.892	7.041	9.299	12.34
14	4.075	4.660	5.629	6.571	7.790	10.17	13.34
15	4.601	5.229	6.262	7.261	8.547	11.04	14.34
16	5.142	5.812	6.908	7.962	9.312	11.91	15.34
17	5.697	6.408	7.564	8.672	10.09	12.79	16.34
18	6.265	7.015	8.231	9.390	1.86	13.68	17.34
19	6.844	7.633	8.907	10.12	11.65	14.56	18.34
20	7.434	8.260	9.591	10.85	12.44	15.45	19.34
21	8.034	8.897	10.28	11.59	13.24	16.34	20.34
22	8.643	9.542	10.98	12.34	14.04	17.24	21.34
23	9.260	10.20	11.69	13.09	14.85	18.14	22.34
24	9.866	10.86	12.40	13.85	15.66	19.04	23.34
25	10.52	11.52	13.12	14.61	16.47	19.94	24.34
26	11.16	12.20	13.84	15.38	17.29	20.84	25.34
27	11.81	12.88	14.57	16.15	18.11	21.75	26.34
28	12.46	13.56	15.31	16.93	18.94	22.66	27.34
29	13.12	14.26	16.05	17.71	19.77	23.57	28.34
30	13.79	14.95	16.79	18.49	20.60	24.480	29.34
40	2.71	22.16	24.43	26.51	29.05	33.66	39.34
50	27.99	29.71	32.36	34.76	37.69	42.94	49.33
60	35.53	37.48	40.48	43.19	46.46	52.29	59.33
70	43.28	45.44	48.76	51.74	55.33	61.70	69.33
80	51.17	53.54	57.15	60.39	64.28	71.14	79.33
90	59.20	61.75	65.65	69.13	73.29	80.62	89.33
100	67.33	70.06	74.222	77.93	82.36	90.13	99.33

n/α	0.750	0.900	0.950	0.975	0.990	0.995	0.999
1	1.323	2.706	3.841	5.024	6.635	7.879	10.83
2	2.773	4.605	5.991	7.378	9.210	10.60	13.82
3	4.108	6.251	4.815	9.348	11.34	12.84	16.27
4	5.385	7.779	9.488	11.14	13.28	14.86	18.47
5	6.626	9.236	11.07	12.83	15.09	16.75	20.52
6	7.841	10.64	12.59	14.45	16.81	18.55	22.46
7	9.037	12.02	14.07	16.01	18.48	20.28	24.32
8	10.22	13.36	15.51	17.53	20.09	21.96	26.12
9	11.39	14.68	16.92	19.02	21.67	23.59	27.88
10	12.55	15.99	18.31	20.48	23.21	25.19	29.59
11	13.70	17.28	19.68	21.92	24.72	26.76	31.26
12	14.85	18.55	21.03	23.34	26.22	28.30	32.91
13	15.98	19.81	22.36	24.74	27.69	29.82	34.58
14	17.12	21.06	23.68	26.12	29.14	31.32	36.12
15	18.25	22.31	25.00	27.49	30.58	32.80	37.70
16	19.37	23.54	26.30	28.85	32.00	34.27	39.25
17	20.49	24.77	27.59	30.19	33.41	35.72	40.79
18	21.60	25.99	28.87	31.53	34.81	37.16	42.31
19	22.72	27.20	30.14	32.85	36.19	38.58	43.82
20	23.83	28.41	31.41	34.17	37.57	40.00	45.32
21	24.93	29.62	32.67	35.48	38.93	41.40	46.80
22	26.04	30.81	33.92	36.78	40.29	42.80	48.27
23	27.14	32.01	35.17	38.08	41.64	44.18	49.73
24	28.24	33.20	36.42	39.36	42.98	45.56	51.18
25	29.34	34.38	37.65	40.65	44.31	46.93	52.62
26	30.43	35.56	38.89	41.92	45.64	48.29	54.05
27	31.53	36.74	40.11	43.19	46.96	49.64	55.48
28	32.62	37.92	41.34	44.46	48.28	50.99	56.89
29	33.71	39.09	42.56	45.72	49.59	52.34	58.30
30	34.80	40.26	43.77	46.98	50.89	53.67	59.70
40	45.62	51.80	55.76	59.34	63.69	66.77	73.40
50	56.33	63.17	67.50	71.42	76.15	79.49	86.66
60	66.98	74.40	79.08	83.30	88.38	91.95	99.61
70	77.58	85.53	90.53	95.02	100.4	104.2	112.3
80	88.13	96.58	101.9	106.6	112.3	116.3	124.8
90	98.65	107.6	113.1	118.1	124.1	128.3	137.2
100	109.1	118.5	124.3	129.6	135.8	140.2	149.4

Критерий Бартлетта. Процентные точки $m(\alpha)$ статистики M . $\alpha=1\%$

$k \backslash c_1$	0.0	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	6.0	7.0	8.0	9.0	10.0	12.0	14.0
3(a)	9.21	9.92	10,47	10,78	10,81	10,50	9,83											
(b)	9.21	9.29	9,38	9,48	9,59	9,71	9,83											
4(a)	11.34	11.95	12,46	12,86	13,11	13,18	13,03	12,65	12,03									
(b)	11.34	11.40	11,46	11,54	11,63	11,72	11,82	11,92	12,03									
5(a)	13.28	13.81	14,30	14,71	15,03	15,25	15,34	15,28	15,06	14,66	14,07							
(b)	13.28	13.33	13,39	13,45	13,53	13,61	13,69	13,78	13,87	13,97	14,07							
6(a)	15.09	15.58	16,03	16,44	16,79	17,07	17,27	17,37	18,37	17,24	16,08	16,03						
(b)	15.09	15.14	15,20	15,26	15,33	15,41	15,48	15,57	15,05	15,74	15,04	16,03						
7(a)	16.81	17.27	17,70	18,10	18,46	18,77	19,02	19,21	19,32	19,35	19,28	18,84	17,92					
(b)	16.81	16.87	16,93	16,99	17,06	17,14	17,21	17,29	17,37	17,46	17,55	17,73	17,92					
8(a)	18.48	18.91	19,32	19,71	20,07	20,39	20,67	20,90	21,08	21,20	21,25	21,13	20,64	19,76				
(b)	18.48	18.54	18,60	18,67	18,74	18,81	18,88	18,96	19,04	19,13	19,21	19,39	19,57	19,76				
9(a)	20.09	20.50	20,90	21,28	21,64	21,97	22,26	22,52	22,74	22,91	23,03	23,10	22,91	22,41	21,56			
(b)	20.09	20.15	20,22	20,29	20,36	20,44	20,51	20,59	20,67	20,75	20,84	21,01	21,19	21,37	21,56			
10(a)	21.67	22.06	22,45	22,82	23,17	23,50	23,80	24,08	24,32	24,54	24,69	24,90	24,90	24,66	24,15	23,33		
(b)	21.67	21.73	21,80	21,88	21,95	22,02	22,10	22,18	22,26	22,34	22,42	22,60	22,77	22,95	23,14	23,33		
11(a)	23.21	23.59	23,97	24,33	24,067	25,00	25,231	25,59	25,85	26,08	26,28	26,57	26,70	26,65	26,38	25,86		
(b)	23.21	23.28	23,35	23,43	23,59	23,58	23,66	23,74	23,82	23,90	23,98	24,15	24,33	24,51	24,69	24,88		
12(a)	24.72	25.10	25,46	25,81	26,15	26,48	26,79	27,08	27,35	27,59	27,81	28,16	28,39	28,46	28,37	28,07	26,79	
(b)	24.72	24.80	24,87	24,95	25,023	25,11	25,18	25,27	25,35	25,43	25,51	25,68	25,86	26,04	26,22	26,41	26,79	
13(a)	26.22	26.58	26,93	27,28	27,62	27,94	28,25	28,54	28,81	27,07	29,30	29,70	29,99	30,16	30,19	30,06	29,22	
(b)	26.22	26.29	26,37	26,45	26,53	26,66	26,69	26,77	26,85	26,94	27,02	27,19	27,37	27,55	27,73	27,91	28,29	
14(a)	27.69	28.04	28,39	28,73	29,06	29,38	29,69	29,98	30,26	30,52	30,77	31,49	31,53	31,77	31,89	31,88	31,39	30,16
(b)	27.69	27.77	27,85	27,93	28,01	28,09	28,17	28,28	28,24	28,42	28,51	28,68	28,86	29,03	29,22	29,40	29,77	30,16
15(a)	29.14	29.49	29,83	30,16	30,49	30,80	30,11	31,40	31,68	31,95	32,20	32,66	33,03	33,32	33,51	33,57	33,37	32,52
(b)	29.14	29.22	29,30	29,38	29,47	29,55	29,63	29,72	29,80	29,89	29,97	30,15	30,32	30,50	30,69	30,87	31,24	31,62

Критерий Бартлетта. Процентные точки $t(\alpha)$ статистики M . $\alpha=5\%$

$k \backslash c_1$	0.0	0.5	1.0	1.5	2.0	2.5	3.0	4.0	4.5	5.0	6.0	7.0	8.0	9.0	10.0	12.0	14.0
3(a)	5.99	6.47	6.89	7.20	7.38	7.39	7.22										
(b)	5.99	6.22	6.43	6.64	6.84	7.03	7.22										
4(a)	7.81	8.24	8.63	8.96	9.21	9.38	9.43	9.37	9.18								
(b)	7.81	8.00	8.17	8.35	8.52	8.69	8.85	9.02	9.18								
5(a)	9.49	9.88	10.24	10.57	10.86	11.08	11.24	11.32	11.31	11.21	11.02						
(b)	9.49	9.65	9.80	9.96	10.11	10.27	10.42	10.57	10.72	10.87	11.02						
6(a)	11.07	11.43	11.78	12.11	12.40	12.65	12.86	13.01	13.11	13.14	13.10	12.78					
(b)	11.07	11.22	11.36	11.51	11.65	11.79	11.94	12.08	12.22	12.36	12.50	12.78					
7(a)	12.59	12.94	13.27	13.59	13.88	14.15	14.38	14.58	14.73	14.83	14.88	14.81	14.49				
(b)	12.59	12.73	12.87	13.00	13.14	13.27	13.41	13.55	13.68	13.82	13.95	14.22	14.49				
8(a)	14.07	14.40	14.72	15.03	15.32	15.60	15.84	16.06	16.25	16.40	16.51	16.60	16.49	16.16			
(b)	14.07	14.20	14.33	14.46	14.49	14.72	14.85	14.98	15.11	15.25	15.38	15.64	15.90	16.16			
9(a)	15.51	15.83	16.14	16.73	16.73	17.01	17.26	17.49	17.70	17.88	18.03	18.22	18.26	18.12	17.79		
(b)	15.51	15.63	15.76	15.89	16.02	16.14	16.27	16.40	16.52	16.65	16.78	17.03	17.29	17.54	17.79		
10(a)	16.92	17.23	17.54	17.83	18.12	18.39	18.65	18.89	19.11	19.31	19.48	19.75	19.89	19.89	19.73	19.40	
(b)	16.92	17.04	17.17	17.29	17.41	17.54	17.66	17.79	17.91	18.04	18.16	18.41	18.66	18.91	19.16	19.40	
11(a)	18.31	18.61	18.91	19.20	19.48	19.76	20.02	20.26	20.49	20.70	20.89	21.21	21.42	21.52	21.49	21.32	
(b)	18.31	18.43	18.55	18.67	18.79	18.91	19.04	19.16	19.28	19.40	19.52	19.77	20.01	20.26	20.50	20.75	
12(a)	19.68	19.97	20.26	20.55	20.83	21.10	21.36	21.61	21.84	22.06	22.27	22.62	22.88	23.06	23.12	22.56	
(b)	19.68	19.79	19.91	20.03	20.15	20.27	20.39	20.51	20.63	20.75	20.87	21.12	21.36	21.60	21.84	22.56	
13(a)	21.03	21.14	21.26	21.38	21.50	21.62	21.74	21.85	21.97	22.09	22.21	22.45	22.69	22.92	23.16	23.88	
(b)	21.03	21.14	21.26	21.38	21.50	21.62	21.74	21.85	21.97	22.09	22.21	22.45	22.69	22.92	23.16	23.88	
14(a)	22.36	22.65	22.93	23.21	23.48	23.75	24.01	24.26	24.50	24.73	24.95	25.34	25.95	26.14	26.25	26.17	25.66
(b)	22.36	22.48	22.60	22.71	22.83	22.95	23.06	23.18	23.30	23.42	23.53	23.77	24.00	24.24	24.71	25.19	25.66
15(a)	23.68	23.97	24.24	24.52	24.79	25.05	25.31	25.56	25.80	26.04	26.26	26.67	27.03	27.33	27.56	27.80	27.50
(b)	23.68	23.80	23.92	24.03	24.15	24.26	24.38	24.50	24.61	24.73	24.85	25.08	25.31	25.55	26.01	26.48	26.95

Критерий Бартлетта. Процентные точки $m(\alpha)$ статистики М. Вспомогательная таблица (С и ΔС)

$k \backslash c_1$	0,5	1	1,5	2	2,5	3	3,5	4	4,5	5,0	6,0	7,0	8,0	10,0	12,0	14,0
3C	0,014	0,111	0,375	0,889	1,736	3,000										
ΔC	0,486	0,889	1,125	1,111	0,764	0,000										
4C	0,008	0,062	0,211	0,500	0,977	1,688	2,680	4,000								
ΔC	0,492	0,938	1,289	1,500	1,523	1,312	0,820	0,000								
5C	0,005	0,040	0,135	0,320	0,625	1,080	1,715	2,560	3,645	5,000						
ΔC	0,495	0,960	1,365	1,680	1,875	1,920	1,785	1,440	0,855	0,000						
6C	0,003	0,028	0,094	0,222	0,434	0,750	1,191	1,778	2,531	3,472	6,000					
ΔC	0,497	0,972	1,406	1,778	2,066	2,250	2,309	2,222	1,969	1,528	0,000					
7C	0,003	0,020	0,069	0,163	0,319	0,551	0,875	1,306	1,860	2,551	4,408	7,000				
ΔC	0,497	0,980	1,431	1,837	2,181	2,449	2,625	2,694	2,640	2,449	1,592	0,000				
8C	0,002	0,016	0,053	0,125	0,244	0,422	0,670	1,000	1,424	1,953	3,375	5,359	8,000			
ΔC	0,498	0,984	1,447	1,875	2,256	2,578	2,830	3,000	3,076	3,047	2,625	1,641	0,000			
9C	0,002	0,012	0,042	0,099	0,193	0,333	0,529	0,790	1,125	1,543	2,667	4,235	6,321			
ΔC	0,498	0,988	1,458	1,901	2,307	2,667	2,971	3,021	3,375	3,457	3,333	2,765	1,679			
10C	0,001	0,010	0,034	0,080	0,156	0,270	0,429	0,640	0,911	1,250	2,160	3,430	5,120	10,000		
ΔC	0,499	0,990	1,466	1,920	2,344	2,730	3,071	3,360	3,589	3,750	3,840	3,570	2,880	0,000		
11C	0,001	0,008	0,028	0,066	0,129	0,223	0,354	0,529	0,753	1,033	1,785	2,835	4,231	8,264		
ΔC	0,499	0,992	1,472	1,934	2,371	2,777	3,146	3,471	3,747	3,967	4,215	4,165	3,769	1,736		
12C	0,001	0,007	0,023	0,056	0,109	0,188	0,298	0,444	0,633	0,868	1,500	2,382	3,556	6,944	12,000	
ΔC	0,499	0,993	1,477	1,944	2,391	2,812	3,202	3,556	3,867	4,132	4,500	4,618	4,444	3,056	0,000	
13C	0,001	0,006	0,020	0,047	0,092	0,160	0,254	0,379	0,539	0,740	1,278	2,030	3,030	5,917	10,225	
ΔC	0,499	0,994	1,480	1,953	2,408	2,840	3,246	3,621	3,961	4,260	4,722	4,970	4,970	4,083	1,775	
14C	0,001	0,994	0,017	0,041	0,080	0,138	0,219	0,327	0,465	0,638	1,102	1,750	2,612	5,102	8,816	14,000
ΔC	0,499	0,005	1,483	1,959	2,420	2,862	3,281	3,673	4,035	4,362	4,898	5,250	5,388	4,898	3,184	0,000
15C	0,001	0,004	0,115	0,036	0,069	0,120	0,191	0,284	0,405	0,556	0,960	1,524	2,276	4,444	7,680	12,196
ΔC	0,499	0,996	1,485	1,964	2,431	2,880	3,309	3,716	4,095	4,444	5,040	5,476	5,724	5,556	4,320	1,804

Приложение 6

Критерий Кохрена. Верхние 1% процентные критические значения α (1%) для статистики $G = \frac{s_{\max}^2}{s_1^2 + s_2^2 + \dots + s_n^2}$,
 построенной по n независимым оценкам дисперсии, каждая из которых обладает k_2 степенями свободы

$k_1 \backslash k_2$	1	2	3	4	5	6	7	8	9	10	16	36	144	∞
2	0.9999	0.9950	0.9794	0.9586	0.9373	0.9172	0.8988	0.8823	0.8674	0.8539	0.7949	0.7067	0.6062	0.5000
3	0.9933	0.9423	0.8831	0.8335	0.7933	0.7606	0.7335	0.7107	0.6912	0.6743	0.6059	0.5153	0.4230	0.3333
4	0.9676	0.8643	0.7814	0.7212	0.6761	0.6410	0.6129	0.5897	0.5702	0.5536	0.4884	0.4057	0.3251	0.2500
5	0.9279	0.7885	0.6957	0.6329	0.5875	0.5531	0.5259	0.5037	0.4854	0.4697	0.4094	0.3351	0.2644	0.2000
6	0.8828	0.7218	0.6258	0.5653	0.5195	0.4866	0.4608	0.4401	0.4229	0.4084	0.3529	0.2858	0.2229	0.1667
7	0.8376	0.6644	0.5685	0.5080	0.4659	0.4347	0.4105	0.3911	0.6751	0.3616	0.3105	0.2494	0.1929	0.1429
8	0.7945	0.6152	0.5209	0.4927	0.4226	0.3932	0.3704	0.3522	0.3373	0.3248	0.2779	0.2214	0.1700	0.1250
9	0.7544	0.5727	0.4810	0.4251	0.3870	0.3592	0.3378	0.3207	0.3067	0.2950	0.2514	0.1992	0.1521	0.1111
10	0.7175	0.5358	0.4469	0.3934	0.3572	0.3308	0.3106	0.2945	0.2813	0.2704	0.2297	0.1811	0.1379	0.1000
12	0.6528	0.4751	0.3919	0.3428	0.3099	0.2861	0.2680	0.2535	0.2419	0.2320	0.1961	0.1535	0.1157	0.0833
15	0.5747	0.4069	0.3317	0.2882	0.2593	0.2386	0.2228	0.2104	0.2002	0.1918	0.1612	0.1251	0.0934	0.0667
20	0.4799	0.3297	0.2654	0.2288	0.2048	0.1877	0.1748	0.1646	0.1567	0.1501	0.1248	0.0960	0.0709	0.0500
24	0.4247	0.2871	0.2295	0.1970	0.1759	0.1608	0.1495	0.1406	0.1338	0.1283	0.1060	0.0810	0.0595	0.0417
30	0.3632	0.2412	0.1913	0.1635	0.1454	0.1327	0.1232	0.1157	0.1100	0.1054	0.0867	0.0658	0.0480	0.0333
40	0.2940	0.1915	0.1508	0.1281	0.1135	0.1033	0.0957	0.0898	0.0853	0.0816	0.0668	0.0503	0.0363	0.0250
60	0.2151	0.1371	0.1069	0.0902	0.0796	0.0722	0.0668	0.0625	0.0594	0.0567	0.0461	0.0344	0.0245	0.0167
120	0.1225	0.0759	0.0585	0.0489	0.0429	0.0387	0.0357	0.0334	0.0316	0.0302	0.0242	0.0178	0.0125	0.0083
∞	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

продолжение приложения 6

Критерий Кохрена. Верхние 5% процентные критические значения α (5%) для статистики $G = \frac{s_{\max}^2}{s_1^2 + s_2^2 + \dots + s_n^2}$,
построенной по n независимым оценкам дисперсии, каждая из которых обладает k_2 степенями свободы.

$k_1 \backslash k_2$	1	2	3	4	5	6	7	8	9	10	16	36	144	∞
2	0.9985	0.9750	0.9392	0.9057	0.8772	0.8534	0.8332	0.8159	0.8010	0.7880	0.7341	0.6602	0.5812	0.5000
3	0.9669	0.8709	0.7977	0.7457	0.7071	0.6771	0.6530	0.6333	0.6167	0.6025	0.5466	0.4748	0.4031	0.3333
4	0.9065	0.7679	0.6841	0.6287	0.5895	0.5598	0.5365	0.5175	0.5017	0.4884	0.4366	0.3720	0.3093	0.2500
5	0.8412	0.6838	0.5981	0.5440	0.5063	0.4783	0.4564	0.4387	0.4241	0.4118	0.3645	0.3066	0.2513	0.2000
6	0.7808	0.6161	0.5321	0.4803	0.4447	0.4184	0.3980	0.3817	0.3682	0.3568	0.3135	0.2612	0.2119	0.1667
7	0.7271	0.5612	0.4800	0.4307	0.3974	0.3729	0.3535	0.3384	0.3259	0.3154	0.2756	0.2278	0.1833	0.1429
8	0.6798	0.5157	0.4377	0.3910	0.3595	0.3362	0.3185	0.3043	0.2926	0.2829	0.2462	0.2022	0.1616	0.1250
9	0.6385	0.4775	0.4027	0.3584	0.3286	0.3067	0.2901	0.2768	0.2659	0.2568	0.2226	0.1820	0.1446	0.1111
10	0.6020	0.4450	0.3733	0.3311	0.3029	0.2823	0.2666	0.2541	0.2439	0.2353	0.2032	0.1655	0.1308	0.1000
12	0.5410	0.3924	0.3264	0.2880	0.2624	0.2439	0.2299	0.2187	0.2098	0.2020	0.1737	0.1403	0.1100	0.0833
15	0.4709	0.3346	0.2758	0.2419	0.2195	0.2034	0.1911	0.1815	0.1736	0.1671	0.1429	0.1144	0.0889	0.0667
20	0.3894	0.2705	0.2205	0.1921	0.1735	0.1602	0.1501	0.1422	0.1357	0.1303	0.1108	0.0879	0.0675	0.0500
24	0.3434	0.2354	0.1907	0.1656	0.1493	0.1374	0.1286	0.1216	0.1160	0.1113	0.0942	0.0743	0.0567	0.0417
30	0.2929	0.1980	0.1593	0.1377	0.1237	0.1137	0.1061	0.1002	0.0958	0.0921	0.0771	0.0604	0.0457	0.0333
40	0.2370	0.1576	0.1259	0.1082	0.0968	0.0887	0.0827	0.0780	0.0745	0.0713	0.0595	0.0462	0.0347	0.0250
60	0.1737	0.1131	0.0895	0.0765	0.0682	0.0623	0.0583	0.0552	0.0520	0.0497	0.0411	0.0316	0.0234	0.0167
120	0.0998	0.0632	0.0495	0.0419	0.0371	0.0337	0.0312	0.0292	0.0279	0.0266	0.0218	0.0165	0.0120	0.0083
∞	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Ортогональные многочлены Чебышева

В таблице даны точные значения многочленов Чебышева

$$P_n^{(1)}(x) = \lambda_n^{(1)} x, \quad P_n^{(2)}(x) = \lambda_n^{(2)} \left[x^2 - \frac{1}{12}(n^2 - 1) \right],$$

$$P_n^{(3)}(x) = \lambda_n^{(3)} \left[x^3 - \frac{1}{20}(3n^2 - 7)x \right],$$

$$P_n^{(4)}(x) = \lambda_n^{(4)} \left[x - \frac{1}{14}(3n^2 - 13)x^2 + \frac{3}{560}(n^2 - 1)(n^2 - 9) \right],$$

$$P_n^{(5)}(x) = \lambda_n^{(5)} \left[x^5 - \frac{5}{18}(n^2 - 7)x^3 + \frac{1}{1098}(15n^4 - 230n^2 + 407)x \right],$$

$$P_n^{(6)}(x) = \lambda_n^{(6)} \left[x^6 - \frac{5}{24}(3n^2 - 31)x^4 + \frac{1}{176}(5n - 110n^3 + 329)x^2 - \frac{5}{14784}(n^2 - 1)(n^2 - 9)(n^2 - 25) \right].$$

Нулевой многочлен $P_n^{(0)}(x)$ тождественно равен единице. Аргумент $x_t = t - (n+1)/2$ ($t = 1, 2, \dots, n$) явно не указан, поэтому в качестве табличных значений аргумента следует пользоваться таблицей линейного многочлена $P_n^{(1)}(x) = \lambda_n^{(1)} x_t$, где $\lambda_n^{(1)} = 1$, если n нечетное число и $\lambda_n^{(1)} = 2$, если n четное. В предпоследней строке указаны суммы квадратов $A_{t,n}^2 = \sum_{t=1}^n \{P_n^{(t)}(x_t)\}^2$, а в последней строке даны коэффициенты $\lambda_n^{(t)}$.

продолжение приложения 7

n=3		n=4			n=5				n=6					n=7					
$p\frac{(1)}{n}$	$p\frac{(2)}{n}$	$p\frac{(1)}{n}$	$p\frac{(2)}{n}$	$p\frac{(3)}{n}$	$p\frac{(1)}{n}$	$p\frac{(2)}{n}$	$p\frac{(3)}{n}$	$p\frac{(4)}{n}$	$p\frac{(1)}{n}$	$p\frac{(2)}{n}$	$p\frac{(3)}{n}$	$p\frac{(4)}{n}$	$p\frac{(5)}{n}$	$p\frac{(1)}{n}$	$p\frac{(2)}{n}$	$p\frac{(3)}{n}$	$p\frac{(4)}{n}$	$p\frac{(5)}{n}$	$p\frac{(6)}{n}$
-1 0 1	1 -2 1	-3 -1 1 3	1 -1 -1 1	-1 3 -3 1	-2 -1 0 1 2	2 -1 -2 -1 2	-1 2 0 -2 1	1 -4 6 -4 1	-5 -3 -1 1 3 5	5 -1 -4 -4 -1 5	-5 7 4 2 -7 5	1 -3 2 2 -3 1	-1 5 10 10 -5 1	-3 -2 -1 0 1 2	5 0 -3 -4 -3 0 5	-1 1 1 0 -1 1	3 -7 6 1 1 3	-1 4 0 6 1 -7 1	1 -6 15 -20 15 -6 1
2 1	6 3	20 2	4 1	20 $\frac{10}{3}$	10 1	14 1	10 $\frac{5}{6}$	70 $\frac{35}{12}$	70 2	84 $\frac{3}{2}$	180 $\frac{5}{3}$	28 $\frac{7}{12}$	252 $\frac{21}{10}$	3 28	84 1	6 $\frac{1}{6}$	154 $\frac{7}{12}$	84 $\frac{7}{20}$	924 $\frac{77}{60}$
n=8							n=9						n=10						
$P_n^{(1)}$	$P_n^{(2)}$	$P_n^{(3)}$	$P_n^{(4)}$	$P_n^{(5),(6)}$		$P_n^{(1)}$	$P_n^{(2)}$	$P_n^{(3)}$	$P_n^{(4)}$	$P_n^{(5)}$	$P_n^{(6)}$	$P_n^{(1)}$	$P_n^{(2)}$	$P_n^{(3)}$	$P_n^{(4)}$	$P_n^{(5)}$	$P_n^{(6)}$		
-7 -5 -3 -1	7 1 -3 -5	-7 5 7 3	7 -13 -3 9	-7 23 -17 -15	1 -5 9 -5	-4 -3 -2 -1	28 7 -8 -17	-14 7 13 9	14 -21 -11 9	-4 11 -4 -9	4 -17 22 1	-9 -7 -5 -3 -1	6 2 -1 -3 -4	-42 14 35 -17 31 12	18 -22 -17 3 18	-6 14 -1 -11 -6	3 -11 10 6 -8		
1 3 5 7	5 -3 1 7	-3 -7 -5 7	9 -3 -13 7	15 17 -23 7	-5 9 -5 1	0 1 2 3 4	-20 -17 -8 7 28	0 -9 -13 -7 14	18 9 9 -11 14	0 9 4 -11 4	-20 1 22 -17 4	1 3 5 7 9	-4 -3 -1 2 6	-12 -31 -35 -17 -14 42	18 3 -17 -22 -14 18	6 11 -1 -14 6	-8 6 10 -11 3		
168 2	168 1	2644 $\frac{2}{3}$	6166 $\frac{7}{12}$	22184 $\frac{7}{10}$	246 $\frac{11}{60}$	60 11	227722 3	990 $\frac{5}{6}$	2002 $\frac{7}{12}$	468 $\frac{3}{20}$	1980 $\frac{11}{60}$	330 2	132 $\frac{1}{2}$	8580 $\frac{5}{3}$	2860 $\frac{5}{12}$	780 $\frac{1}{10}$	660 $\frac{11}{240}$		

$n=11$						$n=12$					
$P_n^{(1)}$	$P_n^{(2)}$	$P_n^{(3)}$	$P_n^{(4)}$	$P_n^{(5)}$	$P_n^{(6)}$	$P_n^{(1)}$	$P_n^{(2)}$	$P_n^{(3)}$	$P_n^{(4)}$	$P_n^{(5)}$	$P_n^{(6)}$
-5	15	-30	6	-3	15	-11	55	-33	33	-33	11
-4	6	6	-6	6	-48	-9	25	3	-27	57	-31
-3	-1	22	-6	1	29	-7	1	21	-33	21	11
-2	-6	23	-1	-4	36	-5	-17	25	-13	-29	25
-1	-9	14	4	-4	-12	-3	-29	19	12	-44	4
						-1	-35	7	28	-20	-20
0	-10	0	6	0	-40						
						1	-35	-7	28	20	-20
						3	-29	-19	12	44	4
1	-9	-14	4	4	-12	5	-17	-25	-13	29	25
2	-6	-23	-1	4	36	7	1	-21	-33	-21	11
3	-1	-22	-6	-1	29	9	25	-3	-27	-57	-31
4	6	-6	-6	-6	-48	11	55	33	33	33	11
5	15	30	6	3	15						
110	858	4290	286	156	11220	572	12012	5148	8008	15912	4488
1	1	$\frac{5}{6}$	$\frac{1}{12}$	$\frac{1}{40}$	$\frac{11}{120}$	2	3	$\frac{2}{3}$	$\frac{7}{24}$	$\frac{3}{20}$	$\frac{11}{360}$

$n=13$						$n=14$					
$P_n^{(1)}$	$P_n^{(2)}$	$P_n^{(3)}$	$P_n^{(4)}$	$P_n^{(5)}$	$P_n^{(6)}$	$P_n^{(1)}$	$P_n^{(2)}$	$P_n^{(3)}$	$P_n^{(4)}$	$P_n^{(5)}$	$P_n^{(6)}$
-6	22	-11	99	-22	22	-13	13	-143	143	-143	143
-5	11	0	-64	33	-55	-11	7	-11	-77	187	-319
-4	2	6	-96	18	8	-9	2	66	-132	132	-11
-3	-5	8	-54	-11	43	-7	-2	98	-92	-28	227
-2	-10	7	11	-26	22	-5	-5	95	-13	-139	185
-1	-13	4	64	-20	-20	-3	-7	67	63	-145	-25
0	-14	0	84	0	-40	-1	-8	24	108	-60	-200
182	2002	572	68068	6188	14212	910	728	97240	136136	235144	497420
1	1	$\frac{1}{6}$	$\frac{7}{12}$	$\frac{7}{120}$	$\frac{11}{360}$	2	$\frac{1}{2}$	$\frac{5}{3}$	$\frac{7}{12}$	$\frac{7}{30}$	$\frac{77}{720}$

$n=15$						$n=16$					
$P_n^{(1)}$	$P_n^{(2)}$	$P_n^{(3)}$	$P_n^{(4)}$	$P_n^{(5)}$	$P_n^{(6)}$	$P_n^{(1)}$	$P_n^{(2)}$	$P_n^{(3)}$	$P_n^{(4)}$	$P_n^{(5)}$	$P_n^{(6)}$
-7	91	-91	1001	-1001	143	-15	35	-455	273	-143	65
-6	52	-13	-429	1444	-286	-13	21	-91	-91	143	-117
-5	19	35	-869	972	-55	-11	9	143	-221	143	-39
-4	-8	58	-704	44	176	-9	-1	267	-201	33	59
-3	-29	61	-249	-751	197	-7	-9	301	-101	-77	87
-2	-44	49	251	-1000	50	-5	-15	265	23	-131	45
-1	-53	27	621	-675	-125	-3	-19	179	129	-115	-25
0	-56	0	756	0	-200	-1	-21	63	189	-45	-75
280		39780		10581480		1360	5712	1007760	470288	201552	77520
	37128		6466460		426360						
						2	1	$\frac{10}{3}$	$\frac{7}{12}$	$\frac{1}{10}$	$\frac{1}{600}$
1	3	$\frac{5}{6}$	$\frac{35}{12}$	$\frac{21}{20}$	$\frac{11}{180}$						

СПИСОК ЛИТЕРАТУРЫ

3. Вентцель Е.С., Овчаров Л.А. Теория вероятностей и ее инженерные приложения / Вентцель Е.С., Овчаров Л.А. – М.: Высшая школа, 2007. – 491с.
4. Гмурман В. Е. Руководство к решению задач по теории вероятностей и математической статистике / Гмурман В. Е. – М.: Высшая школа, 2004.— 404 с.
5. Гмурман В. Е Теория вероятностей и математическая статистика / Гмурман В. Е. – М.: Высшая школа, 2003.— 479 с.
6. Джашеев К.А.-М., Джашеева З.А.-М. Монограммный метод анализа результатов многофакторного эксперимента // Успехи современного естествознания. – 2008. – № 8 – С. 19-28.
7. Кобзарь А.И. Прикладная математическая статистика. (Для инженеров и научных работников) / Кобзарь А.И. – М.: Физматлит, 2006.— 816 с.
8. Маслов Г.Г., Дидманидзе О.Н., Цибулевский В.В. Оптимизация параметров и режимов работы машин методами планирования эксперимента: Учебн. пособие для сельскохозяйственных вузов. – М.:УМЦ «Триада», 2007. – 292 с.
9. Славутский Л.А. Основы регистрации данных и планирования эксперимента. Учебное пособие: Изд-во ЧГУ, Чебоксары, 2006, 200 с.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	3
1 ОСНОВНЫЕ ПОНЯТИЯ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ	4
1.1 Статистическая функция распределения. Гистограмма	4
1.2 Числовые характеристики статистического распределения	7
1.3 Оценки для математического ожидания и дисперсии (точечная оценка параметров распределения)	8
1.4 Основные законы распределения	10
1.4.1 Нормальный закон распределения и его параметры	10
1.4.2 Вероятность попадания случайной величины в заданный интервал. Функция Лапласа	12
1.4.3 Распределение χ^2 «хи квадрат»	17
1.4.4 Распределение Стьюдента	18
1.4.5 Распределение F Фишера-Снедекора	19
1.5 Методы оценивания параметров	20
1.5.1 Метод максимального правдоподобия	20
1.5.2 Метод наименьших квадратов	23
1.6 Доверительный интервал. Доверительная вероятность	26
1.6.1 Доверительный интервал для выборочного среднего	26
1.6.2 Доверительный интервал для оценки выборочной дисперсии нормального распределения	30
1.7 Критерий равенства дисперсий	31
1.8 Исключение резко выделяющихся наблюдений	36
2 РЕГРЕССИОННЫЙ АНАЛИЗ	38
3 МЕТОДИКА АКТИВНОГО ПЛАНИРОВАНИЯ ЭКСПЕРЕМЕНТА	49
3.1 Общие сведения	49
3.2 Полный факторный эксперимент	53
3.3 Анализ уравнения регрессии	61
3.4 Дробные реплики факторного планирования	64
3.5 Оптимизация методом крутого восхождения по поверхности отклика	69
СПИСОК ЛИТЕРАТУРЫ	91

*Емельянов Александр Михайлович,
Кидяева Наталья Петровна,
Подолько Евгения Александровна,
Шилёв Евгений Михайлович*

СТАТИСТИЧЕСКИЕ МЕТОДЫ ОБРАБОТКИ,
ПЛАНИРОВАНИЯ ИНЖЕНЕРНОГО ЭКСПЕРИМЕНТА

Учебное пособие

В редакции составителей

Лицензия ЛР 020427 от 25.04.1997 г.
Подписано к печати 13.01.2015 г. Формат 60×90/16.
Уч.-изд.л. – 4,2. Усл.-п.л. – 6,0.
Тираж 100 экз. Заказ 7.

Отпечатано в отделе оперативной полиграфии издательства ДальГАУ
675005, г. Благовещенск, ул. Политехническая, 86

