

Федеральное агентство по образованию

*ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ СИСТЕМ  
УПРАВЛЕНИЯ И РАДИОЭЛЕКТРОНИКИ (ТУСУР)*

Кафедра радиоэлектронных технологий и экологического  
мониторинга  
(РЭТЭМ)

**Г.В.Смирнов**

## **СТАТИСТИЧЕСКИЕ МЕТОДЫ ОБРАБОТКИ**

Учебное методическое пособие для специальностей

*020801 (013100) «Экология»  
280101 «Безопасность жизнедеятельности в техносфере»  
210201 «Проектирование и технология  
радиоэлектронных средств»*

Томск, 2012

## 1 ОЦЕНКА НЕИЗВЕСТНЫХ ПАРАМЕТРОВ

### 1.1 Основные понятия и определения

Обработка результатов эксперимента производится, чтобы получить исходные данные для любых последующих расчётов, в том числе и вероятностных.

Методы анализа результатов эксперимента и определение по ним вероятностей событий и характеристик случайных величин даёт *математическая статистика (МС)*- общий раздел современной теории вероятностей.

Исходным материалом для применения статистических методов являются экспериментальные или статистические данные, под которыми понимаются сведения о числе объектов, обладающих теми или иными признаками. Например, статистическими данными являются:  $d_1, d_2, d_3, \dots, d_i, \dots, d_n$ -отклонения размеров диаметров  $n$  однотипных бактерий сферической формы от среднего размера ( $d_i$ - отклонение диаметра  $i$ -ой бактерии);

$x_1, x_2, x_3, \dots, x_i, \dots, x_n$ -отклонения размеров  $n$  однотипных деталей от номинального размера ( $x_i$ - отклонение размера  $i$ -ой детали);

$x_1, x_2, x_3, \dots, x_i, \dots, x_n$ -значения предельной прочности  $n$  некоторого материала;

$t_1, t_2, t_3, \dots, t_i, \dots, t_n$ -продолжительность функционирования до отказа (долговечности)  $n$  образцов прибора.

Перечисленные данные являются числовыми характеристиками массовых случайных явлений (размеров бактерий, сортности деталей, прочности материала, долговечности прибора и т.д.) Поэтому предметом математической статистики служат случайные явления, а её основной задачей – количественный и качественный анализ этих явлений.

Напомним кратко исходные понятия МС и те задачи, которые она решает. Совокупность всех возможных значений изучаемого признака  $X$  (другими словами, случайной величины  $X$ ) называется *генеральной совокупностью (ГС)*, конечной или бесконечной. Те значения признака  $X$ , которые зафиксированы в опыте на  $n$  исследуемых образцах, случайным образом отобранных от из генеральной совокупности называются выборкой объёма  $n$  и обозначаются  $x_1, x_2, x_3, \dots, x_i, \dots, x_n$  (малыми буквами).

Учебное пособие написано в рамках инновационной образовательной программы Томского государственного университета систем управления и радиоэлектроники. В учебном пособии изложены основные методы статистической обработки результатов измерений, которые дополнены оригинальными примерами, полученными в собственных исследованиях автора. Учебное пособие написано в соответствии с учебным планом специальностей 210201 «Проектирование и технология радиоэлектронных средств», 020801-«Экология» по дисциплине «Статистическая обработка данных».

Пособие может быть использовано также в курсах лекций дисциплин метрология, математическое моделирование технологических процессов, специальностей 280101 «Безопасность жизнедеятельности в техносфере» и 210201 «Проектирование и технология радиоэлектронных средств», а также в в дисциплинах, связанных с обработкой статистических данных, других специальностей.

Если опыт будет многократно повторяться (хотя бы мысленно!), то мы будем получать различные конкретные выборки. В МС постулируется, что множество конкретных выборок образует *n*-мерный случайный вектор  $x_1, x_2, x_3, \dots, x_i, \dots, x_n$ , называемый случайной выборкой из ГС, где, во-первых, все компоненты  $x_i$  независимы, во-вторых, каждый компонент  $x_i$  есть случайная величина с тем же законом распределения (вообще говоря, неизвестным), которому подчиняется генеральная совокупность  $X$ .

Таким образом, выборочные данные  $x_1, x_2, x_3, \dots, x_i, \dots, x_n$ , полученные в результате эксперимента по изучению случайной величины  $X$ , выступают в двух ролях одновременно:

- 1.) в роли конкретных чисел  $x_1, x_2, x_3, \dots, x_i, \dots, x_n$ , с которыми мы оперируем для вычисления значений некоторых характеристик;
- 2.) в роли случайных величин  $x_1, x_2, x_3, \dots, x_i, \dots, x_n$ , если мы мысленно процесс получения  $n$  выборочных значений «проигрываем» многократно.

Последнее нужно для того, чтобы можно было сделать выводы о свойствах тех или иных правил обработки результатов наблюдений.

Основой любых статистических выводов о свойствах генеральной совокупности является *выборочный метод*, заключающийся в том, что свойства генеральной совокупности  $X$  (случайной величины  $X$ ), устанавливаются путем изучения тех же свойств на случайной выборке объема  $n$ . При этом сами выводы носят вероятностный характер (т.е. справедливы лишь в среднем при многократном повторении эксперимента), а их точность тем выше, чем больше объем выборки  $n$  и чем больше объем априорной информации имеется вначале исследования.

*Замечание.* Наряду с одномерной ГС, можно рассматривать многомерную ГС  $(X, Y, \dots, Z)$ ;  $(X_1, Y_1, \dots, Z_1); \dots; (X_n, Y_n, \dots, Z_n)$ .

Условно можно выделить следующие основные классы задач, решаемых методами МС:

1. Оценка неизвестных параметров в статистических моделях (например, оценка параметров  $\theta_1, \theta_2, \theta_3, \dots, \theta_n$ , от которых зависит функция распределения  $F(x, \theta_1, \theta_2, \theta_3, \dots, \theta_n)$  случайной величины, вид которой известен).
2. Проверка статистических гипотез (в том числе гипотез о предполагаемом виде неизвестного закона распределения случайной величины)

3. Установления формы и степени связи между несколькими случайными переменными.
  4. Классификация объектов по измерению вектора признаков  $(Y_1, Y_2, \dots, Y_m)$ , т.е. распределение объектов на классы по какому-то критерию близости «родственных» объектов.
- Первые три группы задач будут кратко рассмотрены ниже.

## 1.2 Основные выборочные характеристики

Пусть исследуется, некоторый, случайный признак  $X$  (например, для конкретности, долговечность детали), закон распределения которого (функция распределения)  $F(x)$ , вообще говоря, неизвестен.

Обозначим через  $(X_1, X_2, \dots, X_n)$  случайную выборку объема  $n$  из ГС  $X$ . с функцией распределения  $F(x)$ , а через  $(x_1, x_2, \dots, x_n)$  - реализацию этой случайной выборки в данном эксперименте (т.е.  $x_i$ - долговечность  $i$ -го образца из  $n$  поставленных на испытания). Поставим вопрос: «Какие свойства или характеристики ГС  $X$  можно хотя бы приближенно оценить на основе выборочных данных?»

1. Вариационный ряд. Одним из самых простых преобразований выборочных значений является их упорядочение по величине, т.е. расположение их в так называемый *вариационный ряд*:

$$X_1 \leq X_2 \leq \dots \leq X_k \leq \dots \leq X_n \quad (1.1)$$

где  $X_k$  -  $k$ -й член вариационного ряда, или  $k$ -й порядковый член выборки.

Закон распределения случайной величины  $X_k$  легко выражается через  $F(x)$  - закон распределения, ГС  $X$ . Например, закон распределения левого крайнего члена  $X_1$   $F(X_1) = 1 - [1 - F(X)]^n$ . Это позволяет решать обратную задачу: по наблюдениям «слабейших» образцов, т.е. по реализациям случайной величины  $X_1$  находить закон распределения ГС  $F(X)$ . На основе порядковых членов выборки основаны многие критерии проверки различных свойств ГС.

2. Эмпирическая функция распределения. Приближенное представление о неизвестной функции распределения ГС  $F(X)$ , определяющей в точке  $X$  вероятность события  $(X < x)$ , дает так называемая *эмпирическая*, или *выборочная функция распределения*

$$F_n(X) = \frac{v(x)}{n}, \quad (1.2)$$

где  $v(x)$ - число выборочных значений, лежащих левее точки  $x$  на числовой оси.

из определения (1.2) видно, что  $\frac{v(x)}{n}$  есть частота события

$(X < x)$ , которая при больших значениях величины  $n$ , близка к вероятности этого события  $F(X)$ , т.е.  $F_n(X) \cong F(X)$ . Доказано, что при  $n \rightarrow \infty$   $F_n(X) \rightarrow F(X)$  для любых  $x$ .

График функции  $F_n(X)$  представляет собой, ступенчатую функцию со скачками  $\frac{1}{n}$  в каждой точке  $x_i, i = 1, 2, \dots, n$  (рис.1.1).

Зная вид функции  $F_n(X)$ , можно решать задачу обоснованного выбора закона распределения ГС  $F(X)$ , используя так называемые критерии согласия, а также вероятностную бумагу.

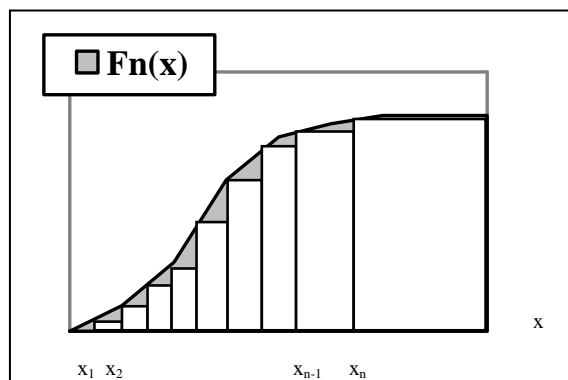


Рис. 1.1. Графики функций  $F_n(X)$  и  $F(X)$

3. Эмпирическая плотность распределения  $f(x)$ . Если число наблюдений  $n$  велико, то выборочные данные распределяют по разрядам, разбивая занятую ими область на интервалы  $(J_1, J_2, \dots, J_m)$  одинаковой длины  $\Delta$  и подсчитывая частоты  $n_k/n$  попадания в эти интервалы.

Табличное оформление этих данных носит название интервального статистического ряда:

$J_1$	$J_2$	.....	$J_k$	.....	$J_m$
$n_1/n$	$n_2/n$	.....	$n_k/n$	.....	$n_m/n$

$$\sum_{k=1}^m \left(\frac{n_k}{n}\right) = 1$$

Разделив частоты  $p_k = n_k/n$  на длину интервалов  $\Delta$ , получают относительные плотности выборочных значений  $f_k = n_k/n \times \Delta, k=1, 2, \dots,$

$m$ , которые и определяют эмпирическую, или выборочную, плотность распределения  $f_k(x) = f_k$  при  $x \in J_k, (k=1, 2, \dots, m)$ . Ясно, что  $f_k(x) \cong f_k$  - плотности распределения ГС  $X$ . Графическое изображение функции  $f_k(x)$  носит название гистограммы (рис. 1.2). При вычислении гистограммы рекомендуется интервалы  $J_k$  выбирать так, чтобы в каждом из них было не менее 10 точек.

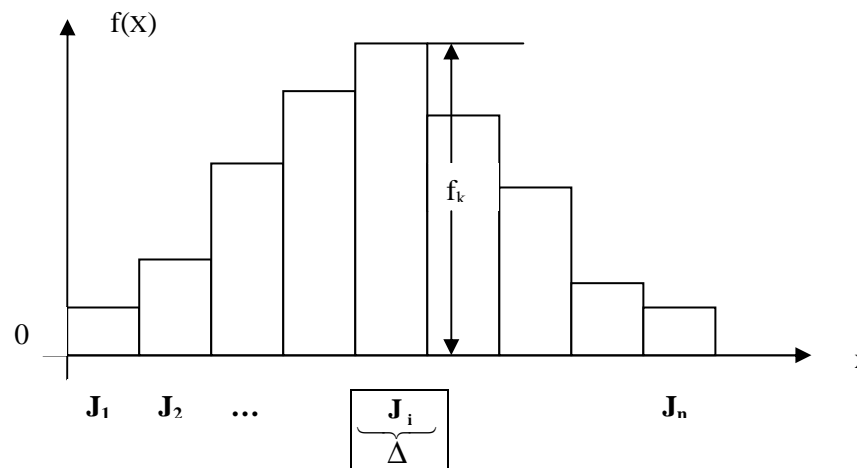


Рис.1.2 Гистограмма

4. Эмпирические (выборочные) числовые характеристики - это приближенные значения соответствующих числовых характеристик ГС, определяемые по выборочным значениям  $x_1, x_2, \dots, x_n$ .

Представление о значении *математического ожидания* или *среднем значении* ГС дает *выборочное среднее значение*

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_k, \tag{1.3}$$

для дисперсии ГС ФТ приближенным значением служит *выборочная дисперсия*

$$D(x) = S^2 = \frac{1}{n} \sum_{i=1}^n (X_k - \bar{X})^2, \tag{1.4}$$

Величина  $\sqrt{S^2} = \bar{S}$  определяет *выборочное среднее квадратическое* (или стандартное) отклонение ГС  $\sigma = \sqrt{S^2}$ .

Приближенным значением ковариации (или корреляционного момента) двух случайных величин  $X$  и  $Y$  является выборочная ковариация

$$k_{xy} = \frac{1}{n} \sum_{i=1}^n (X_k - \bar{X})(Y_k - \bar{Y}), \quad (1.5)$$

Для приближенного значения коэффициента корреляции  $\rho$  между  $X$  и  $Y$  вычисляют выборочный коэффициент корреляции

$$r = k_{xy} / \bar{S}_x \times \bar{S}_y = \sum_{i=1}^n (X_k - \bar{X})(Y_k - \bar{Y}) / \sqrt{\sum_{i=1}^n (X_k - \bar{X})^2 \sum_{i=1}^n (Y_k - \bar{Y})^2}, \quad (1.6)$$

При больших  $n$  от исходных выборочных данных  $(x_1, x_2, \dots, x_n)$  переходят к группированным выборкам  $(x_i, n_i, i=1, 2, \dots, n, \sum n_i = n)$ ,

где  $x_i$  – середина интервала  $J_i$ , на который попадают  $n_i$  выборочных значений  $x_k$ ; тогда формулы (1.3)-(1.5) очевидным образом меняются на следующие:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n \rho_k X_k, \quad S^2 = \frac{1}{n} \sum_{i=1}^n \rho_k (X_k - \bar{X})^2,$$

$$k_{xy} = \frac{1}{n} \sum_{i=1}^n \rho_{kj} (X_k - \bar{X})(Y_j - \bar{Y}), \quad (1.7)$$

где  $\rho_k = n_k/n$ ;  $\rho_{kj} = n_{kj}/n$ .

Заметим, что при вычислении выборочных характеристик по результатам опыта в формулы (1.3) ÷ (1.6) вместо случайных величин  $X_k$  подставляются их конкретные числовые значения  $x_k$ , полученные в опыте.

### 1.3 Точечные оценки параметров

Во многих случаях бывает заранее известно, что функция распределения  $F(X)$  ГС принадлежит к определенному классу функций, зависящих от одного или нескольких параметров:  $F(X) = F(x, \theta_1, \theta_2, \theta_3, \dots, \theta_n)$ . В этом случае отыскание закона распределения ГС  $X$  сводится к оценке неизвестных параметров  $\theta_1, \theta_2, \theta_3, \dots, \theta_n$  по выборке  $x_1, x_2, \dots, x_n$ .

Например, пусть  $X$  есть результат измерения некоторой физической постоянной  $a$ , т.е.  $x = a + \varepsilon$ , где  $\varepsilon$  – погрешность измерения, которую в силу действия многих независимых факторов можно считать распределенной по нормальному закону с математическим ожиданием  $M\varepsilon = 0$  (если отсутствует систематическая погрешность) и дисперсией  $S_\varepsilon^2 = \sigma^2$ , определяющей класс точности измерительного прибора. Тогда закон распределения  $X$  является также нормальным, с параметрами  $Mx = a$  и  $S_x^2 = S_\varepsilon^2 = \sigma^2$ , которые и нужно определить по результатам  $n$  измерений:  $x_1, x_2, \dots, x_n$ .

Подчеркнем, что ни при каком конечном объеме выборки  $n$  нельзя определить по выборке точное значение неизвестного параметра  $\theta$ , можно лишь найти его приближенное значение, которое называется оценкой параметра по выборочным данным.

Существуют два вида оценок неизвестного параметра  $\theta$ :

1) точечная оценка  $\theta^*$  под которой понимается любая функция выборочных значений  $\theta^* = f(x_1, x_2, \dots, x_n)$ , приближенно равная  $\theta$ :  $\theta^* = f(x_1, x_2, \dots, x_n) \cong \theta$

2) интервальная оценка  $(\underline{\theta}, \bar{\theta})$  в виде интервала, границы которого  $\underline{\theta}$  и  $\bar{\theta}$  определяются по выборочным данным таким образом, что интервал  $(\underline{\theta}, \bar{\theta})$  покрывает неизвестное значение  $\theta$  с заданной вероятностью  $\gamma$  называемой доверительной и близкой к 1), т.е.

$$P(\underline{\theta} < \theta < \bar{\theta}) = \gamma.$$

Рассмотрим вначале точечные оценки параметров. Прежде всего заметим, что любая оценка  $\theta^*$  являясь функцией случайных выборочных значений  $x_1, x_2, \dots, x_n$ , сама является случайной величиной. Поэтому приближенное равенство  $\theta^* \cong \theta$  следует понимать в вероятностном смысле: для любого  $\varepsilon > 0$  вероятность того, что  $|\theta^* - \theta| < \varepsilon$ , близка к 1 при больших значениях  $n$ .

Оценка  $\theta^*$  параметра  $\theta$  считается «хорошей», когда она является:

1) несмещенной (иначе - не дает систематической погрешности) - это означает, что  $M\theta^* = \theta$ . (1.8)

2) состоятельной, если сходится по вероятности к  $\theta$ , т.е.

$$P(|\theta^* - \theta| < \varepsilon) \rightarrow 1 \quad \text{при } n \rightarrow \infty; \quad (1.9)$$

3) *эффективной* по сравнению с другими оценками  $\theta_k$ , если обладает наименьшей среди них дисперсией, т.е.

$$D\theta^* = \min D\theta_k. \quad (1.10)$$

Существуют различные методы отыскания точечных оценок: метод моментов, метод максимального правдоподобия, метод наименьших квадратов и др. [15, 16]. Оценки максимального правдоподобия обладают перечисленными выше свойствами (кроме, может быть, первого, нарушение которого обычно легко исправить) и, кроме того, при больших значениях  $n$  исследуемые величины распределены по нормальному закону [18].

Примеры точечных оценок. 1. Пусть параметр  $\theta = Mx$  - среднее значение ГС  $x$ . Тогда независимо от вида закона распределения  $X$  несмещенной, состоятельной оценкой  $\theta$  является выборочное среднее

$$\bar{\theta} = \bar{x} = \frac{1}{n} \sum_{k=1}^n X_k.$$

Оценка  $\theta_1 = [x_1 + x_n]/2$  - является несмещенной и состоятельной, но не является эффективной. Если известно дополнительно, что величина  $X$  распределена нормально, т.е.  $X \sim N(\mu, \sigma)$ , то  $\bar{x} = \frac{1}{n} \sum_{k=1}^n X_k$  является оценкой максимального правдоподобия.

2. Пусть  $\theta = Dx$  - дисперсия ГС  $X$ . Тогда можно показать, что выборочная дисперсия  $S_{выб}^2 = \frac{1}{n} \sum_{k=1}^n (X_k - \bar{X})^2$ , является состоятельной и эффективной оценкой, но не является несмещенной, так как  $M S_{выб}^2 = \frac{n-1}{n} \sigma^2$ .

Нетрудно видеть, что «подправленная» выборочная дисперсия

$$D(x) = S S_{выб}^2 = \frac{n-1}{n} \sum_{k=1}^n (X_k - \bar{X})^2 \quad (1.11)$$

будет уже *несмещенной оценкой*. Эту оценку мы и будем использовать в дальнейшем.

3. Пусть  $\sigma = p$  - параметр *биномиального закона* распределения, имеющий смысл вероятности появления «успеха» в любом из  $n$  независимых испытаний.

Оценкой *максимального правдоподобия* это го параметра является частота «успеха» в  $n$  опытах:  $p \cong$

$\theta^* = \frac{m}{n}$ , где  $m$  - число «успехов» в  $n$  испытаниях. Эта оценка является *несмещенной и состоятельной*.

Другие примеры и техника расчета точечных оценок приведены в [16,18].

Недостаток точечных оценок состоит в том, что при малом объеме выборки  $n$  вычисленное значение оценки может сильно отличаться от неизвестного значения параметра  $\theta^*$ , и о значении этого отклонения у нас нет никакой информации.

Можно было бы наряду с оценкой  $\theta^*$  указывать её дисперсию  $D\theta^*$  или её оценку  $X$ , что позволило бы судить о точности  $\theta^*$ . Однако более удобным инструментом для этой цели являются интервальные оценки, или доверительные интервалы.

#### 1.4 Интервальные оценки параметров

Определение интервальной оценки, или доверительного интервала (ДИ)  $(\underline{\theta}, \bar{\theta})$ , уровня  $\gamma$  было дано выше. Добавим, что  $\underline{\theta}$  называется *нижней доверительной границей (НДГ)*, а  $\bar{\theta}$  - *верхней доверительной границей (ВДГ)*; ширина интервала  $\bar{\theta} - \underline{\theta} = 2r$  определяет точность оценки, а доверительная вероятность  $\gamma$  - достоверность её. Смысл величины  $\gamma$  состоит в следующем: при большом числе повторений эксперимента, по результатам которого определяется ДИ  $(\underline{\theta}, \bar{\theta})$ , лишь в  $(1 - \gamma)$  100% случаев истинное значение  $\theta$  не будет покрыто ДИ. Так, если  $\gamma = 0,9$  и опыт повторен  $N = 100$  раз, то лишь в 10 случаях мы не гарантируем покрытия доверительным интервалом неизвестного значения параметра  $\theta$ . Обычно рассматривают либо симметричные доверительные интервалы вида  $\theta - \varepsilon, \theta + \varepsilon$  (в случае симметричного распределения оценки  $\theta^*$ ), где  $\theta^*$  - некоторая точечная оценка параметра  $\theta$ , а  $\varepsilon$  подбирается из условия  $P(|\theta^* - \theta| < \varepsilon) = \gamma$ , либо интервалы, у которых вероятность нарушения условия  $(\underline{\theta} < \theta < \bar{\theta})$  в ту и другую сторону одинакова и равна

$$(1 - \gamma)/2 = \alpha/2. \text{ т.е. } P(\theta < \underline{\theta}) = P(\theta > \bar{\theta}) = \alpha/2$$

Иногда интерес представляет лишь НДГ  $\underline{\theta}$  или лишь ВДГ  $\bar{\theta}$  (например, если  $\theta$  есть средний срок службы изделия, то достаточно указать НДГ  $\underline{\theta}$ , если  $\theta$  - интенсивность отказов, то достаточно определить ВДГ  $\bar{\theta}$ ). Тогда говорят об одностороннем доверительном интервале уровня  $\gamma$ . Если найдены односторонние ДГ  $\underline{\theta}$  и  $\bar{\theta}$  уровня  $\gamma$  каждая, то тем самым определен двусторонний ДИ  $(\underline{\theta}, \bar{\theta})$  уровня  $\gamma' = 2\gamma - 1$ . Методы построения точных ДИ для неизвестного параметра изложены в [5,7]. Общее правило определения ДИ для любого параметра  $\theta$  (которое является приближенным) основано на центральной предельной теореме Ляпунова, согласно которой оценка  $\theta$  при больших значениях  $n$  ( $n > 50$ ) имеет нормальный закон распределения со средним  $M\theta^* = \theta$  (если  $\theta$  - несмещенная оценка) и дисперсией  $D\theta^*$ .

Обозначим через  $Z_{1-\alpha/2}$  квантиль нормального распределения уровня  $1-\alpha/2$ , где  $\alpha=1-\gamma$ , т.е. такое значение аргумента функции Лапласа  $\Phi(x)$ , при котором  $\Phi(Z_{1-\alpha/2}) = 1-\alpha/2$ . Тогда

$P(\theta^* - \theta < Z_{1-\alpha/2} \sqrt{D\theta^*}) \cong 2\Phi(Z_{1-\alpha/2}) - 1 = \gamma$ , т.е. ДИ для параметра  $\theta$  (будем его обозначать  $J_\gamma(\theta)$ ) имеет вид

$$J_\gamma(\theta) = \theta^* \pm Z_{1-\alpha/2} \sqrt{D\theta^*}. \quad (1.12)$$

Таким образом, чтобы применить формулу (1.12), нужно лишь знать оценку  $\theta^*$  и её дисперсию  $D\theta^*$ . Например, если  $\theta = \mu = Mx$ , то  $\theta^* = \bar{X}$  и  $D\theta^* = s^2/n$ , а если  $\theta = \sigma = \sqrt{Dx}$ , то  $\theta^* = S$  и  $D\theta^* = 2s^2/(n-1)$  где  $\bar{X}$  и  $S$  - выборочные оценки, которые определяются по формулам (1.3), (1.11). Если  $\theta = r$  - коэффициент корреляции между случайными величинами

$X$  и  $Y$ , то  $\theta^* = r$  и  $D\theta^* \cong \frac{(1-r^2)^2}{n}$ , где  $r$  - определяется по формуле (1.6).

### 1.5 Точные доверительные интервалы для параметров нормального закона

На практике наиболее распространены нормально распределенные случайные величины (в силу действия центральной предельной теоремы). Ноли исследуемый признак  $X$  имеет нормальный закон распределения с неизвестными параметрами  $\mu = Mx$  и  $\sigma = \sqrt{Dx}$ , (коротко это записывается в виде  $X \sim n(\mu, \sigma)$ ), то точные ДИ уровня  $\gamma$  для  $\mu$  и  $\sigma$  определяются следующим образом. Предварительно вычисляются точечные оценки параметров  $\mu$  и  $\sigma$  (если  $\sigma$  неизвестно) по формулам

(1.3), (1.11), затем по заданной доверительной вероятности  $\gamma$  находят  $\alpha/2 = (1-\gamma)/2$  и  $1-\alpha/2$ , после чего с учетом объема выборки  $n$  по соответствующим таблицам [13,17,18] находят квантили уровней  $1-\alpha/2$  и  $\alpha/2$ -соответствующих распределений.

Возможны следующие случаи:

1. Параметр  $\sigma$  известен. Оценка  $X \sim n(\mu, \sigma/\sqrt{n})$  и, следовательно,

$P(|\bar{X} - \mu| < Z_{1-\alpha/2} \sigma/\sqrt{n} = \gamma)$ , т.е. доверительный интервал

$$J_\gamma(\mu) = \bar{X} \pm Z_{1-\alpha/2} \sigma/\sqrt{n} \quad (1.13)$$

2. Параметр  $\sigma$  неизвестен. Случайная величина  $u(\bar{X} - \mu) \sqrt{n}/s$ , имеет распределение Стьюдента с  $n-1$  степенью свободы. По таблицам этого распределения при заданных значениях  $n-1$  и  $1-\alpha/2$  можно найти квантиль  $t_{1-\alpha/2}$  уровня  $1-\alpha/2$ , при котором

$P((\bar{X} - \mu) \sqrt{n}/s < t_{1-\alpha/2}) = \gamma$ , откуда следует, что доверительный интервал  $J_\gamma(\mu) = \bar{X} \pm t_{1-\alpha/2} S/\sqrt{n}$  (1.14)

При  $n > 20$   $t_{1-\alpha/2} \cong Z_{1-\alpha/2}$  и вместо таблиц распределения Стьюдента можно пользоваться таблицами нормального распределения.

3. Параметры  $\mu$  и  $\sigma$  неизвестны. Нужно найти ДИ  $J_\gamma(\sigma^2)$  и  $J_\gamma(\sigma)$ . Воспользуемся тем, что случайная величина  $s^2(n-1)/S^2$  имеет  $\chi^2$ -распределение ("хи-квадрат" распределение) с  $n-1$  степенью свободы. По таблицам этого распределения [13, 17, 18] при заданных значениях  $n-1$  и  $\alpha=1-\gamma$  можно найти значения  $\chi_{\alpha/2}^2$  и  $\chi_{1-\alpha/2}^2$  - квантили этого распределения уровней  $\alpha/2$  и  $1-\alpha/2$ , для которых

$$P(\chi_{\alpha/2}^2 < \frac{(n-1)\sigma^2}{S^2} < \chi_{1-\alpha/2}^2) = 1-\alpha = \gamma;$$

откуда следует, что искомые ДИ имеют вид

$$J_\gamma(\sigma^2) = (S^2 \frac{n-1}{\chi_{1-\alpha/2}^2}, S^2 \frac{n-1}{\chi_{\alpha/2}^2}), \quad (1.15)$$

$$J_\gamma(\sigma) = (S \sqrt{\frac{n-1}{\chi_{1-\alpha/2}^2}}, S \sqrt{\frac{n-1}{\chi_{\alpha/2}^2}}) \quad (1.16)$$

## 1.6 Обработка результатов измерений (две основные задачи теории ошибок)

При измерении некоторого физического параметра  $a$  результат  $i$ -го измерения  $X$ , как уже отмечалось выше, можно представить в виде  $x_i = a + \varepsilon$ ,  $i=1,2,3,\dots, n$ , где ошибку измерения  $\varepsilon_i$  - полагают нормально распределенной случайной величиной, т.е.  $\varepsilon_i \sim N(0, \sigma^2)$ . Тогда  $x_i \sim N(a, \sigma^2)$ , причем измерения предполагаются независимыми. Таким образом, результаты измерения  $x_1, x_2, \dots, x_n$  оказываются независимыми нормально распределенными с параметрами  $a$  и  $\sigma$  случайными величинами.

Две основные задачи теории ошибок как раз и состоят в оценке «истинного значения»  $a$  измеряемой величины и средней квадратической ошибки  $\sigma$  измерения. Иными словами, требуется по выборке  $x_1, x_2, \dots, x_n$  из нормальной генеральной совокупности  $X$  найти точечные  $a^*$ ,  $S = \sigma^*$  оценки параметров  $a$  и  $\sigma$  и их ДИ  $J_\gamma(a)$  и  $J_\gamma(\sigma)$  заданного уровня  $\gamma$ . Решение этих задач уже рассмотрено выше.

Пример 1.1. По результатам испытаний на прочность 4 образцов стальных стержней получены следующие значения разрывных усилий (в ньютонах):

$$x_1 = 300, x_2 = 310, x_3 = 330, x_4 = 340.$$

Найти ДИ уровня  $\gamma = 0,9$  для средней прочности  $\mu = Mx$  и среднее квадратическое отклонение  $\sigma = \sqrt{Dx}$ , считая закон распределения  $x$  нормальным.

Решение. По формулам (1.3) и (1.11) находим точечные оценки  $\mu$  и  $\sigma$

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i = 320, S = \sqrt{\frac{1}{n-1} \sum (x_i - \bar{x})^2} = 18,25$$

По таблицам распределения Стьюдента [13,18] для  $\alpha = 1 - \gamma = 0,1$  и  $n-1=3$  находим  $t_{1-0,05} = 2,35$ , а по таблицам  $\chi^2$  распределения при  $n-1 = 3$ ,  $\alpha/2 = 0,05$  и  $1-\alpha/2 = 0,95$  находим  $\chi_{0,05}^2 = 0,35$  и  $\chi_{0,95}^2 = 7,8$ .

Подставляя найденные значения в формулы (1.14) и (1.16), получим  $J_{\gamma=0,9}(\mu) = 320 \pm 21,4$ ;  $J_{\gamma=0,9}(\sigma) = (11,3; 53,3)$

## 2 ПРОВЕРКА СТАТИСТИЧЕСКИХ ГИПОТЕЗ 2.1 Понятия критерия и правило проверки гипотез

*Статистической гипотезой* называют любое предположение о законе распределения или его параметрах одной или нескольких случайных величин. Например, предположения типа  $\mu = Mx = \mu_0$ , где  $\mu_0$  - известное значение;  $Dx_1 = Dx_2$ ,  $F(x) = F_0(x)$ , где  $F_0(x)$  - известная функция распределения и т.д. являются статистическими гипотезами.

Задача состоит в том, чтобы на основании выборочных наблюдений (полученных в эксперименте) дать ответ на вопрос «можно ли принять выдвинутую гипотезу или нет» или «можно ли считать, что данные эксперимента согласуются с выдвинутой гипотезой или они ей противоречат?»

Ответ на этот вопрос невозможно дать абсолютно достоверно, как бы ни был велик объем выборочных наблюдений, так как всегда есть некоторая вероятность  $\alpha$ , называемая уровнем значимости ошибочно отклонить исходную (или как говорят, нулевую) гипотезу  $H_0$ , и, следовательно, принять некоторую альтернативную, или конкурирующую, гипотезу  $H_1$ . Уровень значимости  $\alpha$  есть вероятность ошибки 1-го рода - отклонить гипотезу  $H_0$ , когда она верна. Наряду с этим неизбежна ошибка 2-го рода: принять  $H_0$ , когда она неверна, а, следовательно, верна конкурирующая гипотеза  $H_1$ . Вероятность ошибки 2-го рода обозначаемой через  $\beta$ .

Для проверки любой статистической гипотезы подбирается некоторая функция выборочных наблюдений, зависящая от  $H_0$ ,

$$T = T(x_1, x_2, \dots, x_n | H_0), \quad (2.1)$$

закон распределения которой, в предположении что  $H_0$  верна, полностью известен (является одним из табличных). Любая такая функция носит название *критерия проверки данной гипотезы*. Ясно, что для проверки одной и той же гипотезы могут существовать различные критерии и, следовательно, нужно уметь как-то сравнивать их. Вопросы выбора наилучших критериев является достаточно сложным [18] и здесь не рассматривается. Для проверки стандартных гипотез, наиболее часто встречаются на практике наилучшие критерии, как правило, уже известны. Основные из этих критериев рассмотрены ниже.



*Правило проверки гипотезы*  $H_0$  при выбранном критерии  $T$  состоит в следующем. Область возможных значений критерия  $T$  (т.е. всю числовую ось  $R$ ) разбивают на две области:

- критическую область  $K$ , куда возможные значения могут попасть лишь с малой вероятностью (не превышающей допустимого уровня  $\alpha$ ), если  $H_0$  верна;

- область допустимых значений  $D=R/K$ , являющуюся дополнением  $K$ .

Если значение критерия  $T=T_0$ , вычисленное по конкретным выборочным данным, попадает в область  $D$ , то гипотеза  $H_0$  принимается на уровне значимости  $\alpha$  (т.е. делается вывод о том, что она не противоречит опытным данным). В противном случае, т.е. при  $T_0 \in K$ , гипотеза  $H_0$  отклоняется и принимается конкурирующая о ней гипотеза  $H_1$ .

Для применения этого правила нужно:

1) выбрать в соответствии с общими рекомендациями подходящий для данного случая критерий  $T$ ;

2) задать, приемлемый уровень значимости  $\alpha$  ( $\alpha = 0,01 \div 0,2$ ) и по нему определить (пользуясь таблицей для выбранного критерия  $T$ ) критическую область  $K$  так, чтобы вероятность ошибки 1-го рода не превышала  $\alpha$ :

$$P(T \in K | H_0) \leq \alpha, \quad (2.2)$$

а величина ошибки 2-го рода  $\beta$  была минимальной;

3) сравнить вычисленное по результатам конкретных наблюдений  $x_1, x_2, \dots, x_n$  значение критерия  $T=T_0$ , с критической областью  $K$ , отклоняя  $H_0$ , если  $T_0 \in K$ , и принимая  $H_0$ , если  $T_0 \in \bar{K} = D$ .

**Замечание 1.1.** В п. 3 часто вместо условия  $T_0 \in K$  пишут

$$T > T_{кр},$$

где  $T_{кр}$  - критическое значение критерия  $T$ , определяющее границу области  $K$ , которое находят из соответствующих таблиц по заданному уровню значимости  $\alpha$  и объему выборки  $n$ .

**Замечание 1.2.** Другой способ сравнения  $T_0$  с граничным значением  $T_{кр}$  который иногда бывает удобнее, состоит в следующем. Сравнивается вероятность  $P(T > T_0 | H_0)$  того, что возможные при гипотезе  $H_0$  значения  $T$  превысят наблюдаемое в опыте значение  $T_0$  с допустимым уровнем значимости. Если  $P(T > T_0 | H_0) > \alpha$ , то это означает, что значения критерия  $T$ , превышающие  $T_0$ , вполне вероятны, т.е. результаты опыта подтверждают гипотезу  $H_0$ .

## 2.2 Правило наилучшего выбора критической области

Условие (2.2), обеспечивающее приемлемый уровень ошибки 1-го рода, определяет критическую область  $K$  (для выбранного критерия  $T$ ) неоднозначно, т.е. при одном и том же уровне  $\alpha$  область  $K$  может быть различной (рис. 2.1). Но при этом вид области  $K$  существенно влияет на значение ошибки 2-го рода  $\beta$ .

*Правило Неймана-Пирсона* предписывает такой выбор  $K$ , из множества возможных, определяемых условием (2.2), при котором  $\beta$  минимальна или, что все равно, максимальна дополнительная вероятность  $W = 1 - \beta$  отклонения  $H_0$ , когда она неверна (верна конкурирующая гипотеза  $H_1$ ). Вероятность  $W$  называется *мощностью критерия* и, очевидно, зависит от конкурирующей гипотезы

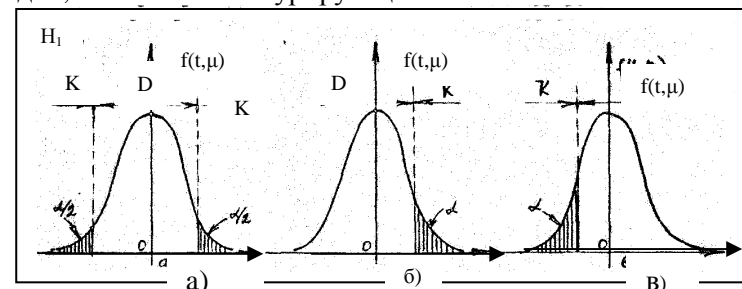


Рис. 2.1. Различные крайние типы критических областей:

а - двусторонняя; б - правая односторонняя; в - левая односторонняя. Проиллюстрируем правило наилучшего выбора  $K$  на примере проверки следующей гипотезы.

**Задача 1.1.** (проверка гипотезы о среднем значении). Требуется проверить на уровне значимости  $\alpha$  гипотезу  $H_0, \mu = \mu_0$ , где  $\mu = Mx$ , а  $x_i \sim N(\mu, \sigma)$ , т.е.  $X$  имеет нормальный закон распределения с параметрами  $\mu$  и  $\sigma$ ;  $\sigma$  - известна. Исходные данные - выборка  $x_1, x_2, \dots, x_n$  объема  $n$  из ГС  $X$ .

Рассмотрим возможные конкурирующие гипотезы:

- а)  $H_1: \mu \neq \mu_0$  - двусторонняя альтернатива;
- б)  $H_2: \mu > \mu_0$  - правая односторонняя альтернатива;
- в)  $H_3: \mu < \mu_0$  - левая односторонняя альтернатива.

Альтернативы «б» и «в» предполагают наличие некоторой дополнительной информации.

Критерий проверки  $H_0$  естественно строить, исходя из выборочной оценки среднего  $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ , которая близка к  $\mu_0$ , если  $H_0$  верна.

Примем в качестве критерия величину

$$T = \frac{\bar{X} - \mu_0}{\sigma} \sqrt{n}, \quad (2.3)$$

которая имеет стандартное нормальное распределение  $N(0,1)$ , если  $H_0$  верна. Рассмотрим следующие 3 типа критической области  $K$ , являющиеся крайними случаями при выполнении условия (2.2) (см. рис. 2.1, где  $f(t, \mu)$  - распределение  $T$ , при условии, что  $\mu = \mu_0$ ).

Можно показать, что ошибка  $\beta$  будет минимальной, если критическая область  $K$  выбирается в области тех значений  $T$ , которые соответствуют конкурирующей гипотезе. Это значит, что если  $H_1$  - двусторонняя, то  $K$  тоже должна быть двусторонней и т.д.

Найдем критическое значение критерия  $T_{кр}$  (по таблицам нормального распределения), определяющее границу критической области  $K$  для каждого из случаев «а», «б», «в». Значение  $T = T_{кр}$  называют также *табличным значением критерия*.

Так как  $T \sim N(0,1)$ , то для двусторонней альтернативы

$$\alpha = P(|T| > T_{кр}) = 1 - P(|T| < T_{кр})$$

откуда  $T_{кр} = Z_{1-\alpha/2}$  - квантиль нормального распределения уровня  $1-\alpha/2$

При правой односторонней альтернативе  $T_{кр}$  выбирается из условия

$$\alpha = P(T > T_{кр}) = 1 - P(T < T_{кр}),$$

откуда  $T_{кр} = Z_{1-\alpha}$  - квантиль нормального закона уровня  $1-\alpha$ .

Наконец, для левой односторонней альтернативы в силу симметричности нормального распределения  $T_{кр} = -Z_{1-\alpha} = Z_{\alpha}$ .

Зная  $T_{кр}$ , с учетом замечания 1 делаем вывод о том, что  $H_0$  отклоняется, если:

$$|T| > T_{кр} \text{ в случае «а»};$$

$$T > T_{кр} \text{ в случае «б»};$$

$$T < T_{кр} \text{ в случае «в»}.$$

**Замечание 1.3.** Если  $\sigma$  - неизвестно, то в качестве критерия принимают величину

$$T = \frac{\bar{X} - \mu_0}{S} \sqrt{n}, \quad (2.4)$$

которая называется *критерием Стьюдента*, так как имеет распределение Стьюдента с  $n-1$  степенью свободы. При этом  $T_{кр} = t_{1-\alpha/2}$  в случае «а»,

$T_{кр} = t_{1-\alpha}$  - в случае «б»,  $T_{кр} = -t_{1-\alpha/2}$  в случае «в», где  $t_{\varphi}$  - квантиль распределения Стьюдента с  $n-1$  степенью свобода уровня  $\varphi$ .

### 2.3 Проверку стандартных гипотез

Рассмотренная выше задача 1 является примером проверки стандартной гипотезы. Рассмотрим ещё ряд стандартных гипотез.

**Задача 1.2** (проверка гипотезы о равенстве двух средних значений).

Требуется проверить на уровне значимости  $\alpha$ . гипотезу  $H_0: \mu = \mu_0$  - против  $H_1: \mu \neq \mu_0$ , имея  $n_1$  выборочных значений  $x_1, x_2, \dots, x_{n_1}$  случайной величины  $x_i \sim N(\mu_1, \sigma_1)$  и  $n_2$  выборочных значений  $y_1, y_2, \dots, y_{n_2}$  случайной величины  $y_i \sim N(\mu_2, \sigma_2)$ .

**Случай 1-й:**  $\sigma_1$  и  $\sigma_2$  - известны. Тогда в качестве критерия принимают

$$T = (\bar{x} - \bar{y}) / \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \sim N(0,1), \text{ если } H_0 \text{ верна} \quad (2.5).$$

$$\text{Здесь } \bar{x} = \frac{1}{n_1} \sum_{i=1}^{n_1} X_i; \quad \bar{y} = \frac{1}{n_2} \sum_{i=1}^{n_2} Y_i$$

При выбранной альтернативе  $H_1$  (двусторонней)  $T_{кр} = Z_{1-\alpha/2}$ .

**случай 2-й:**  $\sigma_1$  и  $\sigma_2$  - неизвестны, но известно, что  $\sigma_1 = \sigma_2$ . Тогда в качестве критерия принимают

$$T = (\bar{x} - \bar{y}) / S \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \quad (2.5a)$$

где  $s^2 = \frac{\nu_1}{\nu_1 + \nu_2} S_x^2 + \frac{\nu_2}{\nu_1 + \nu_2} S_y^2$ ;  $\nu_1 = n_1 - 1$ ;  $\nu_2 = n_2 - 1$

где  $s^2 = \frac{V_1}{V_1 + V_2} S_x^2 + \frac{V_2}{V_1 + V_2} S_y^2$ ;  $V_1 = n_1 - 1$ ;  $V_2 = n_2 - 1$

$S_x^2, S_y^2$  — выборочные дисперсии  $x$  и  $y$ .

Критерий (2.5а) имеет распределение Стьюдента с  $V = V_1 + V_2$  степенями свободы, если  $H_0$  верна. Следовательно,  $T_{кр} = t_{1-\alpha/2}$  — квантиль распределения Стьюдента с  $V = n_1 + n_2 - 2$ , степенями свободы уровня  $1-\alpha/2$ .

Когда  $\sigma_1$  и  $\sigma_2$  неизвестны и, быть может  $\sigma_1 \neq \sigma_2$ , точного решения не существует (это так называемая *проблема Беренса-Фишера* [18]), то неплохим приближением может служить критерий (2.5), в котором неизвестные значения,  $\sigma_1$  и  $\sigma_2$  следует заменить их выборочными оценками  $S_1$  и  $S_2$ .

Задача 1.3 (проверка гипотезы о равенстве двух дисперсий).

Требуется проверить на уровне значимости  $\alpha$  гипотезу  $H_0: \sigma_1^2 = \sigma_2^2$  против  $H_1: \sigma_1^2 \neq \sigma_2^2$ , имея выборки  $x_1, x_2, \dots, x_n$  и  $y_1, y_2, \dots, y_n$  — из двух генеральных совокупностей  $X$  и  $Y$ , распределенных по нормальному закону.

Для проверки указанной гипотезы наилучшим является *критерий Фишера*, равный отношению двух выборочных дисперсий

$$S_1^2 = \frac{1}{n_1 - 1} \sum_{i=1}^n (X_k - \bar{X})^2$$

$$S_2^2 = \frac{1}{n_2 - 1} \sum_{i=1}^n (Y_k - \bar{Y})^2$$

$T = S_1^2 / S_2^2$  при  $S_1^2 > S_2^2$  (или  $T = S_2^2 / S_1^2$  при  $S_2^2 > S_1^2$ ) (2.6), который имеет

F-распределение Фишера с  $V_1 = n_1 - 1$ ;  $V_2 = n_2 - 1$  степенями свободы (с  $V_1$  и  $V_2$ , при  $S_1^2 > S_2^2$ ), если  $H_0$  верна.

Поскольку всегда  $T > 1$ , то принимаем в качестве критической области  $K = \{T > F_{1-\alpha/2}\}$ , где  $F_{1-\alpha/2}$  — квантиль F-распределения с  $V_1$  и  $V_2$  степенями свободы уровня  $1-\alpha/2$ , определяемый по таблицам [13,18].

Задача 1.4 (проверка однородности дисперсий). Требуется проверить на уровне значимости  $\alpha$  гипотезу  $H_0: \sigma_i^2 = \sigma^2, i=1,2,\dots, r$

против  $H_1: \sigma_i^2 \neq \sigma^2$ , где  $\sigma_i^2$  — дисперсия случайной величины

$x_i \sim N(\mu_i, \sigma_i)$ , если выборочные дисперсии  $S_1^2, S_2^2, \dots, S_r^2$  найденные по выборкам объемов  $n_1, n_2, \dots, n_r$  соответственно.

Если  $r=i$ , т.е. проверяется гипотеза о равенстве дисперсии заданному значению  $\sigma^2$ , то для её проверки используется

критерий «*хи-квадраткритерий, или критерий Пирсона*»

$$T = (n-1)S_1^2 / S_2^2, \quad (2.7)$$

имеющий  $\chi^2$ -распределение с  $n_1-1$  степенью свободы, если  $H_0$  верна.

В качестве критической области (при  $H_1: \sigma_i^2 \neq \sigma^2$ ) принимают  $K\{\chi_{\alpha/2}^2 < T < \chi_{1-\alpha/2}^2\}$  где  $\chi_{\alpha/2}^2$  и  $\chi_{1-\alpha/2}^2$  — квантили  $\chi^2$ -распределения с числом степеней свободы  $n_1-1$  уровня  $\alpha$  и  $1-\alpha/2$ .

Если  $r > i$ , но все  $n_i = n$  (объемы выборок одинаковы), то для проверки однородности дисперсий  $r$  случайных величин используется *критерий Кохрена*

$$T = \max (S_i^2) / \sum_{i=1}^r S_i^2, \quad (2.8)$$

таблицы распределения которого, зависящего от двух параметров  $r$  и  $n$ , приведены, например, в [13,17]. По ним, при заданных значениях  $\alpha$ ,  $r$  и  $n$  можно найти критическое значение  $T_{кр}$ .

Если  $r > i$  и  $n_i \neq n$  (объемы выборок различны), то вместо *критерия Кохрена* используется *критерий Бартлета*, имеющего приближенно  $\chi^2$ -распределение. Вид этого критерия можно найти в [1, 2].

Мы оставляем в стороне за недостатком места многие другие критерии проверки «стандартных» гипотез (таких, как проверка однородности двух и более распределений, отбраковка аномальных наблюдений, гипотез о виде закона распределения и др.). Подробное изложение этих вопросов можно найти в [13, 16, 18].

### 3 ЭЛЕМЕНТЫ ТЕОРИИ ОПРЕДЕЛИТЕЛЕЙ

Знакомство с использованием определителей качнем с простейшего случая решения и исследования системы двух уравнений с двумя неизвестными.

Пусть дана система:

$$\begin{cases} a_1x + b_1y = c_1 \\ a_2x + b_2y = c_2 \end{cases} \quad (3.1)$$

Для отыскания решения этой системы, т.е. совокупности таких значений  $x=x_0$ ,  $y=y_0$ , которые обращают в тождества оба уравнения системы, преобразуем (3.1) в такую систему (3.2), где каждое из уравнений содержит лишь одно неизвестное. Для этого умножим первое из уравнений (3.1) на  $b_2$ , второе - на  $b_1$ , и сложим, получим тогда новую систему

$$\begin{cases} (a_1b_2 - a_2b_1)x = c_1b_2 - a_2c_1 \\ (a_1b_2 - a_2b_1)y = a_1c_2 - a_2c_1 \end{cases} \quad (3.2)$$

Заметим, что если  $a_1b_2 - a_2b_1 \neq 0$ , то от системы (3.2) можно аналогичным преобразованием вернуться обратно к системе (3.1), для этого умножаем первое уравнение системы (3.2) на  $a_1$ , второе - на  $b_1$ , и складываем; затем умножаем первое уравнение на  $a_2$ , второе - на  $b_2$  и снова складываем, тогда получим:

$$\begin{cases} (a_1b_2 - a_2b_1)(a_1x + b_1y) = c_1(a_1b_2 - a_2b_1) \\ (a_1b_2 - a_2b_1)(a_2x + b_2y) = c_2(a_1b_2 - a_2b_1) \end{cases} \quad (3.3)$$

Сокращая на  $a_1b_2 - a_2b_1$ , приходим к исходной системе (3.1). Отсюда следует, что системы (3.1) и (3.2) равносильны: каждое ранение системы (3.1) является решением системы (3.2), поскольку (3.2) есть следствие (3.1) и наоборот [поскольку (3.1) есть следствие системы (3.1)].

Из системы (3.2) получаем единственное решение системы (3.1)

$$x_0 = \frac{c_1b_2 - c_2b_1}{a_1b_2 - a_2b_1} \quad y_0 = \frac{a_1c_2 - a_2c_1}{a_1b_2 - a_2b_1} \quad (3.4)$$

Выражения, которые являются коэффициентами системы (3.2) и фигурируют в правых частях формул (3.4), определяющих решение системы (3.1), получили название *определителей второго порядка*.

Для их обозначения вводятся следующая символическая запись:

$$a_1b_2 - a_2b_1 = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \quad (3.5)$$

В этих обозначениях числители формулы (3.4) запишутся в виде

$$\begin{aligned} c_1b_2 - c_2b_1 &= \begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix} \\ a_1c_2 - a_2c_1 &= \begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix} \end{aligned}$$

а сами формулы имеют вид

$$x_0 = \frac{\begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}} ; \quad y_0 = \frac{\begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}} \quad (3.7)$$

и именуются *формулами Крамера*.

Определитель  $\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}$  составленный из коэффициентов при

неизвестных в уравнения системы (3.1), называется *определителем этой системы*: в первом горизонтальном ряде (так называемой *первой строке*) определителя стоят коэффициенты при  $x$  и  $y$  первого уравнения, во второй строке - второго, в первом вертикальном ряде (так называемом *первом столбце*) определителя стоят коэффициенты при  $x$ , во втором - коэффициенты при  $y$ .

Определитель системы часто для краткости обозначают одной буквой  $\Delta$ , а для определителей (3.6) вводятся обозначения  $\Delta_1$  и  $\Delta_2$ :

$$\Delta = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}, \quad \Delta_1 = \begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix}, \quad \Delta_2 = \begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}$$

Тогда формулы (3.7) запишутся так:

$$x_0 = \frac{\Delta_1}{\Delta} ; \quad y_0 = \frac{\Delta_2}{\Delta}$$

Заметим, что определители  $\Delta_1$  к  $\Delta_2$ , получаются из определителя системы  $\Delta$  заменой соответственно первого или второго столбца столбцом свободных членов уравнений (3.1).

Определитель третьего порядка есть число, определяемое следующим равенством

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1 \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - b_1 \begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix} + c_1 \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix} \quad (3.8)$$

Существуют и другие изображения первой части формулы, т.е. другие правила вычисления определителя третьего порядка, приводящие, однако, к тому же результату.

Условимся называть *минором* некоторого элемента данного определителя третьего порядка тот определитель второго порядка, который получится, если из определителя третьего порядка вычеркнуть столбец и строку, содержащие данный элемент. Так минором элемента  $b_1$  будет определитель

$$\begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix}$$

*Минор* данного элемента, взятый со знаком «плюс», если сумма номеров строки и столбца, содержащих этот элемент, - четная, и со знаком «минус», - если сумма эта нечетная, называется *алгебраическим дополнением данного элемента*.

Алгебраические дополнения элементов условимся обозначать теми же буквами и с теми же индексами, что и сами элементы, но прописными. Так алгебраическим дополнением элемента  $a_1$  будет  $A_1$

$$A_1 = + \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix}, \text{ алгебраическим дополнением } b_1 \text{ будет } B_1 = - \begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix},$$

$$\text{алгебраическим дополнением } c_1 \text{ будет } C_1 = + \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}.$$

Равенство (3.3) переписется теперь так:

$$\Delta = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1 A_1 + b_1 B_1 + c_1 C_1 \quad (3.9)$$

Равенство (3.9) дает, как говорят, разложение определителя третьего порядка по элементам первой строки.

Можно доказать, что определитель третьего порядка может быть аналогичным способом разложен по элементам любой его строки и любого столбца, иными словами, определитель третьего порядка равен сумме произведений элементов любой его строки или столбца на их алгебраические дополнения, откуда получим в дополнение (3.9) еще пять следующих равенств.

$$\Delta = a_2 A_2 + b_2 B_2 + c_2 C_2; \quad \Delta = a_3 A_3 + b_3 B_3 + c_3 C_3; \quad \Delta = a_1 A_1 + a_2 A_2 + a_3 A_3; \\ \Delta = b_1 B_1 + b_2 B_2 + b_3 B_3; \quad \Delta = c_1 C_1 + c_2 C_2 + c_3 C_3 \quad (3.10)$$

Сумма произведений элементов какой либо строки (или столбца) определителя на алгебраические дополнения соответствующих элементов другой строки (другого столбца) равно нулю.

Иными словами, справедливы 12 равенств следующего вида.

$a_1 A_2 + b_1 B_2 + c_1 C_2 = 0$ и еще 5 других, аналогичных равенств	$a_1 B_1 + a_2 B_2 + a_3 B_3 = 0$ (3.11) и еще 5 других, аналогичных равенств
---	--

Используя изложенное выше, рассмотрим решение системы трех линейных уравнений с тремя неизвестными

$$\begin{cases} a_1 x + b_1 y + c_1 z = d_1 \\ a_2 x + b_2 y + c_2 z = d_2 \\ a_3 x + b_3 y + c_3 z = d_3 \end{cases} \quad (3.12)$$

Определитель системы равен

$$\Delta = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \quad (3.13)$$

Преобразуем эту систему в такую, где каждое уравнение содержит лишь одно неизвестное. Для этого умножим уравнения (3.12) сначала на  $A_1, A_2, A_3$  и сложим; затем на  $B_1, B_2, B_3$  и снова сложим, наконец, на  $C_1, C_2, C_3$  и опять сложим ( $A_1 \dots C_3$ , как и раньше, - алгебраические дополнения элементов определителя), используя результаты (3.10) и (3.11), придем к новой системе уравнений:

$$\begin{cases} \Delta x = A_1 d_1 + A_2 d_2 + A_3 d_3 \\ \Delta y = B_1 d_1 + B_2 d_2 + B_3 d_3 \\ \Delta z = C_1 d_1 + C_2 d_2 + C_3 d_3 \end{cases} \quad (3.14)$$

Если  $\Delta \neq 0$ , то система (3.14) равносильна исходной, чтобы в этом убедиться, достаточно сложить уравнения (3.14), сначала умножив их на  $a_1, v_1, c_1$ , затем на  $d_2, v_2, c_2$ , и, наконец, на  $d_3, v_3, c_3$ .

Найдем

$$\Delta(a_1 x + v_1 y + c_1 z) = d_1(A_1 a_1 + B_1 v_1 + C_1 c_1) + d_2(A_2 a_1 + B_2 v_1 + C_2 c_1) + d_3(A_3 a_1 + B_3 v_1 + C_3 c_1)$$

или используя (3.10) и (3.11)

$$\begin{cases} \Delta(a_1 x + v_1 y + c_1 z) = \Delta d_1 \\ \Delta(a_2 x + v_2 y + c_2 z) = \Delta d_2 \\ \Delta(a_3 x + v_3 y + c_3 z) = \Delta d_3 \end{cases} \quad (3.15)$$

Поскольку  $\Delta \neq 0$ , то, сокращая на  $\Delta$ , получим исходную систему (3.12).

Таким образом, *равносильность* (уравнений) выражений (3.12) и (3.14) доказана, поскольку каждая из них есть следствие другой и они обе могут иметь, следовательно, только одни и те же решения.

Выражения в правых частях уравнений (3.14) можно записать, используя (3.12) в виде определителей, получаемых из определителя  $\Delta$  заменой его столбцов поочередно столбцом свободных членов:

$$A_1 d_1 + A_2 d_2 + A_3 d_3 = \begin{vmatrix} d_1 & v_1 & c_1 \\ d_2 & v_2 & c_2 \\ d_3 & v_3 & c_3 \end{vmatrix},$$

$$B_1 d_1 + B_2 d_2 + B_3 d_3 = \begin{vmatrix} a_1 & d_1 & c_1 \\ a_2 & d_2 & c_2 \\ a_3 & d_3 & c_3 \end{vmatrix},$$

$$C_1 d_1 + C_2 d_2 + C_3 d_3 = \begin{vmatrix} a_1 & v_1 & d_1 \\ a_2 & v_2 & d_2 \\ a_3 & v_3 & d_3 \end{vmatrix}.$$

Вводя для этих определителей обозначения  $\Delta_1, \Delta_2, \Delta_3$ , запишем систему (3.14) в виде

$$\begin{cases} \Delta x = \Delta_1 \\ \Delta y = \Delta_2 \\ \Delta z = \Delta_3 \end{cases} \quad (3.16)$$

Отсюда находим единственное решение системы (3.14)  $x = x_0, y = y_0, z = z_0$  и равносильной ей системы (3.12), в виде

$$x_0 = \frac{\Delta_1}{\Delta}; \quad y_0 = \frac{\Delta_2}{\Delta}; \quad z_0 = \frac{\Delta_3}{\Delta}. \quad (3.17)$$

Аналогично будут решаться системы линейных уравнений, имеющие 4, 5 и более неизвестных. Пользуясь изложенными выше правилами, в этом случае, выражаем определители системы через алгебраические дополнения третьего, четвертого и т.д. порядков, а затем полученные результаты, через алгебраические дополнения второго порядка.

*Транспонированной* по отношению к матрице  $A$  называют такую матрицу  $A^T$ , которую получают путем изменения мест столбцов и строк матрицы  $A$ , например

$$A = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}; \quad A^T = \begin{vmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \end{vmatrix}. \quad (3.18)$$

*Умножение матриц.* Две матрицы  $A=(a_{ij})_{m \times n}$  и  $B=(b_{ki})_{p \times q}$  можно умножать друг на друга только тогда, когда число столбцов матрицы, стоящей первым сомножителем, равно числу строк матрицы, стоящей вторым сомножителем.

Таким образом, для вышеприведенных матриц  $A$  и  $B$  произведение  $A \times B$  можно вычислить только тогда, когда  $n=p$ , а произведение  $B \times A$ , только тогда, когда  $q=m$ .

Пусть теперь даны две матрицы:  $A=(a_{ij})_{m \times n}$  и  $B=(b_{ki})_{n \times p}$ . За их произведение  $A \times B$  принимается по определению матрица  $C=(c_{ij})_{m \times p}$ , элементы которой  $c_{ij}$  определяются следующими формулами:

$$c_{ij} = \sum_{k=1}^n a_{ik} \times b_{kj} = a_{i1}b_{1j} + a_{i2}b_{2j} + a_{i3}b_{3j} + \dots + a_{in}b_{nj} \quad (3.19)$$

$$(i=1,2,3,\dots,m; j=1,2,3,\dots,p).$$

Матрицы  $A \times B$  и  $B \times A$  не только не равны, но даже и разной структуры.

Пусть дана матрица  $A$

$$A = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}; \quad d(A) \neq 0$$

где  $d(A) \neq \Delta \neq 0$  *детерминант* (определитель) матрицы  $A$ .

*Обратная матрица*  $A^{-1}$  равна

$$A^{-1} = \frac{1}{d(A)} \begin{vmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{32} \\ A_{13} & A_{23} & A_{33} \end{vmatrix} \quad (3.20)$$

где  $A_{ij}$  - алгебраическое дополнение элементов  $a_{ij}$  в определителе  $d(A)$ .

Читателю рекомендуется обратить внимание на порядок индексов в матрице (3.20).

Матрица, построенная из алгебраических дополнений элементов не особой квадратной матрицы  $A$  в определителе  $d(A)$ , в которой алгебраические дополнения элементов строк расположены по столбцам и наоборот, называется *присоединенной матрицей матрицы  $A$*

и обозначается  $\tilde{A}$ .

$$\tilde{A} = \begin{vmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{32} \\ A_{13} & A_{23} & A_{33} \end{vmatrix} \quad (3.21)$$

Для двух квадратных матриц одного и того же порядка независимость их произведения от порядка сомножителей ( $A \times B = B \times A$ ) возможна лишь в исключительных случаях. Такие матрицы называют *коммутативными*.

### Примеры

$$1. \quad A = \begin{vmatrix} 1 & 2 \\ 3 & 1 \end{vmatrix}, \quad B = \begin{vmatrix} 2 & 1 \\ 0 & 3 \\ 1 & -1 \end{vmatrix}.$$

Произведение  $A \times B$  не имеет смысла. В то же время произведение  $B \times A$  можно найти

$$B \times A = \begin{vmatrix} 2 \times 1 + 1 \times 3 & 2 \times 2 + 1 \times 1 \\ 0 \times 1 + 3 \times 3 & 0 \times 2 + 3 \times 1 \\ 1 \times 1 - 1 \times 3 & 1 \times 2 - 1 \times 1 \end{vmatrix} = \begin{vmatrix} 5 & 5 \\ 9 & 3 \\ -2 & 1 \end{vmatrix}.$$

$$2. \quad A = \begin{vmatrix} 1 & 2 & 1 \\ 3 & 1 & 2 \end{vmatrix}, \quad B = \begin{vmatrix} 2 & -1 \\ 1 & 3 \\ 0 & 1 \end{vmatrix}.$$

Возможны оба произведения:

$$A \times B = \begin{vmatrix} 1 \times 2 + 2 \times 1 + 1 \times 0 & 1 \times (-1) + 2 \times 3 + 1 \times 1 \\ 3 \times 2 + 1 \times 1 + 2 \times 0 & 3 \times (-1) + 1 \times 3 + 2 \times 1 \end{vmatrix} = \begin{vmatrix} 4 & 6 \\ 7 & 2 \end{vmatrix}$$

$$B \times A = \begin{vmatrix} 2 \times 1 + (-1) \times 3 & 2 \times 2 + (-1) \times 1 & 2 \times 1 + (-1) \times 2 \\ 1 \times 1 + 3 \times 3 & 1 \times 2 + 3 \times 1 & 1 \times 1 + 3 \times 2 \\ 0 \times 1 + 1 \times 3 & 0 \times 2 + 1 \times 1 & 0 \times 1 + 1 \times 2 \end{vmatrix} = \begin{vmatrix} -1 & 3 & 0 \\ 10 & 5 & 7 \\ 3 & 1 & 2 \end{vmatrix}.$$

## 4 ОСНОВЫ РЕГРЕССИОННОГО АНАЛИЗА

### 4.1 Постановка задачи. Исходные предположения

В этой главе будет рассмотрен третий основной класс задач математической статистики, связанный с установлением формы и степени связи между несколькими - случайными переменными. С такими задачами приходится иметь дело при количественном исследовании многих физических явлений, когда целью эксперимента является построение математической модели явления. Большой класс моделей может быть представлен в форме аналитической зависимости между набором контролируемых переменных  $x_1, x_2, \dots, x_n$ , называемых факторами, и выходной переменной  $y$  (одной или несколькими), называемой откликом:

$$Y = \varphi(x_1, x_2, \dots, x_n) + \varepsilon \quad \text{или} \quad Y = \varphi(x) + \varepsilon, \quad (4.1)$$

где  $\eta = \varphi(x)$  - неизвестная функция, существование которой постулируется;

$\varepsilon$  - случайная величина, имеющая смысл погрешности модели; её природа - это ошибки измерения и действие неучтенных факторов.

Иногда причины, обуславливающие величину  $\varepsilon$ , действуют порознь.

Если заведомо известно, что между откликом и факторами существует детерминированная зависимость (в этом случае отклик  $Y = \eta$  - неслучайная величина), то величина  $\varepsilon$  порождена лишь погрешностями измерения.

Напротив, при отсутствии погрешностей измерения величина  $\varepsilon$  определяется лишь действием неучтенных факторов. В этом случае отклик является случайной величиной.

Как в том, так и в другом случаях величина  $\eta = \varphi(x) = M[Y/x]$ , т.е.  $\eta$  является условным средним значением  $Y$  при заданных значениях  $x$ .

Соотношение

$$Y = \varphi(x) \quad (4.2)$$

называется *уравнением регрессии переменной  $y$  по переменным  $x_1, x_2, \dots, x_n$* , а установление вида функции  $\varphi(x)$  в некоторой области  $Q \in X$  по результатам эксперимента - *регрессионным анализом* (здесь  $X$  -  $m$ -мерное пространство факторов).

В результате эксперимента в некоторых точках  $x_i \in Q$   $i=1, 2, 3, \dots, n$  измеряются значения  $\varphi(x_i)$ , на которые накладываются случайные ошибки  $\varepsilon_i$ , т.е. фиксируются величины

$$Y = \varphi(x_i) + \varepsilon_i \quad (4.3)$$

По этим данным, а также при некоторых априорных сведениях о функции  $\varphi(x)$  требуется, возможно, более точно, оценить неизвестную функцию  $\varphi(x)$ .

Обычно на основе физических соображений, а также предварительных исследований заранее известно, что функция  $\varphi(x)$  принадлежит к некоторому классу  $Z$  достаточно «хороших» (например, дифференцируемых и т.п.) функций.

Наиболее благоприятным является такой случай, когда  $Z$  - некоторое параметрическое семейство функций, заданных в области  $Q$ , т.е.

$$Z = \{ \varphi(x, \theta_1, \theta_2, \dots, \theta_j, \dots, \theta_r), x_i \in Q, \theta_j \in G_j, j=1, 2, 3, \dots, r \},$$

где  $\vec{\theta} = (\theta_1, \theta_2, \dots, \theta_j, \dots, \theta_r)^T$  - вектор неизвестных параметров ( $T$  - символ операции транспонирования, т.е. замены отроки столбцом).

Например, известно, что скорость протекания химической реакции между некоторыми компонентами пропорциональна объему исходного вещества. Тогда объем  $V(t)$  в момент времени  $t$  меняется по экспоненциальному закону

$$V(t) = Q_0 e^{-Q_1(t-t_0)} (t > t_0),$$

где  $Q_0, Q_1$  - неизвестные параметры модели, которые нужно оценить наилучшим образом,  $t_0$  - начальный момент времени.

Одним из наиболее общих методов оценки неизвестных параметров  $Q_0$  и  $Q_1$  является *метод наименьших квадратов*, который будет рассмотрен ниже.

К сожалению, рассмотренный выше случай является редким. Более реальна ситуация, когда о функции  $y^*$  известно лишь, что она является достаточно гладкой и может быть «с малой погрешностью» представлена отрезком разложения в ряд Тейлора по некоторым базисным функциям  $\{f_k(x)\}$ ,  $k=0, 1, 2, \dots$ , т.е.

$$\varphi(x) = \varphi^*(x) = \sum_{k=0}^l \beta_k f_k(x) \quad \text{или} \quad \varphi^*(x) = \beta^T f(x), \quad (4.4)$$

$\beta = (\beta_0, \beta_1, \dots, \beta_n)^T$  - вектор неизвестных параметров;



$f(x) = \{f_1(x), f_2(x), \dots, f_n(x)\}^T$  - вектор базисных функций (известных заранее);

$n$ - в общем случае неизвестный параметр, уточняемый в ходе построения модели.

$$Y = \beta_0 + \sum_{j=1}^n \beta_j x_j + \sum_{\substack{u,j \\ j \neq u}}^n \beta_{uj} x_u x_j + \sum_{j=1}^n \beta_{jj} x_{jj}^2 + \dots, \quad (4.5)$$

где  $\beta_j = \frac{\partial y}{\partial x_j}$ ;  $\beta_{uj} = \frac{\partial^2 y}{\partial x_j \partial x_u}$ ;  $\beta_{jj} = \frac{\partial^2 y}{\partial x_j^2}$ ;

Иными словами, исходный класс функций, к которому принадлежит искомая функция  $y^* = \varphi(x)$  заменяется некоторым классом более простых: по структуре функций, представляемых (4.5), и задача снова сводится к наилучшей оценке вектора неизвестных параметров  $\beta$ .

После обработки экспериментальных данных, поскольку модель и измерение имеют определенную погрешность  $\varepsilon$ , получают так называемые *выборочные коэффициенты регрессии*  $b_0, b_1, \dots, b_p, \dots, b_{uj}, \dots, b_{jj}$ , являющимися оценками теоретических коэффициентов  $\beta_0, \beta_j, \beta_{uj}, \beta_{jj}$ .

Уравнение регрессии, полученное на основании опыта, запишется следующим образом

$$Y^* = b_0 + \sum_{j=1}^n b_j x_j + \sum_{\substack{u,j \\ j \neq u}}^n b_{uj} x_u x_j + \sum_{j=1}^n b_{jj} x_{jj}^2 + \dots \quad (4.6)$$

где  $b_0$  - называется свободным членом уравнения регрессии;

$b_j$  - линейными эффектами;

$b_{uj}$  - эффектами взаимодействия;

$b_{jj}$  - квадратичными эффектами коэффициенты уравнения (4.6) определяются методом наименьших квадратов.

При такой постановке задачи исходная погрешность в классе  $Z^*$

$$\|y_i - \varphi^*(x_i)\| = \sqrt{\sum_{i=1}^n (y_i - \varphi^*(x_i))^2} = \sqrt{\sum_{i=1}^n \delta_i^2} = \Delta$$

может быть представлена в виде

$$\Delta^* = \|y_i - \varphi^*(x_i)\| = \|y_i - \varphi(x_i)\| + \|\varphi(x_i) - \varphi^*(x_i)\| \quad (4.7)$$

где  $\delta = \|\varphi^*(x_i) - \varphi(x_i)\| = \sqrt{\sum [\varphi^*(x_i, \beta) - \varphi(x_i)]^2}$  - определяет

систематическую погрешность модели  $\varphi^*(x)$  связанную с заменой исходного класса функций  $Z$  более узким  $Z^*$ , который будем называть аппроксимирующим;  $\Delta = \|y_i - \varphi^*(x_i)\| = \sum \varepsilon_i^2$  погрешность модели, обусловленная рассеянием отклика.

В силу (4.7) фразу о «малой погрешности» представления (4.4) нужно понимать в том смысле, что систематической погрешностью  $\delta$ , при замене  $Z$  на  $Z^*$ , можно пренебречь, по сравнению со случайной погрешностью  $\Delta$ , обусловленной случайными колебаниями наблюдений  $y_i$  в ходе эксперимента.

Именно на сопоставлении этих двух типов погрешностей и основаны правила *проверки адекватности* (качества) модели  $\eta = \varphi^*(x, \beta)$ , где вектор параметров  $\beta$  заманен вектором оценок  $b$ .

Одним из наиболее распространенных аппроксимирующих классов является класс полиномов, в котором в качестве базисных функций выбраны степени переменных  $x_1, x_2, \dots, x_n$ . Это объясняется простым видом полиномиальных моделей и удобством их интерпретации.

Простейшей полиномиальной моделью является модель 1-го порядка, линейная по всем переменным:

$$\eta = \sum_{k=0}^n \beta_k x_k$$

где  $x_0$  - фиктивная переменная.

С этой модели обычно начинают поиск искомой зависимости  $\eta = \varphi(x)$ .

Следующей по сложности является модель 1-го порядка со всеми взаимодействиями факторов, имеющая вид

$$\eta = \beta_0 + \sum_{j=1}^N \beta_j x_j + \sum_{\substack{u,j \\ j \neq u}}^N \beta_{uj} x_u x_j + \dots \quad (4.8)$$

Наконец, для поиска экстремальных условий эксперимента используется модель 2-го порядка, которая получается из предыдущей добавлением всех квадратов переменных:

$$\eta = \beta_0 + \sum_{j=1}^N \beta_j x_j + \sum_{\substack{u,j \\ j \neq u}}^N \beta_{uj} x_u x_j + \dots + \sum_{j=1}^N \beta_{jj} x_j^2 \dots$$

Более сложные полиномиальные модели 3-го порядка и выше используются сравнительно редко. Следует также отметить, что обычно имеются некоторые априорные сведения, позволяющие заранее исключить из модели общего вида некоторые члены, характеризующие малозначимые взаимодействия факторов и их квадраты. Это обстоятельство существенно упрощает поиск адекватной модели.

Следует подчеркнуть, что выражение (4.4) является самым общим видом линейной по параметрам модели и представляет не только полиномиальные модели. Например, в качестве базисных функций  $\{f_k(x)\}$  могут выступать тригонометрические  $\{\sin kx, \cos kx\}$ , показательные  $e^{kx}$  и т.д.

Если неизвестная функция  $\varphi(x)$  представлена в виде (4.4), то задача, тем самым, сведена к предыдущей ситуации, т.е. к оценке вектора неизвестных параметров  $\beta$ .

Исходные предположения. Результаты эксперимента, в общем случае, можно представить следующей матрицей П, в которой перечислены  $n$  различных точек  $x_1, x_2, \dots, x_n$ , факторного пространства X, в которых проводился эксперимент (первая строка); перечислены величины  $r_1, r_2, \dots, r_n$ ,  $\sum_{i=1}^n r_i = N$ , где  $r_i$  - число параллельных опытов в точке  $x_i$  (вторая строка) и указаны наблюдавшиеся в эксперименте значения отклика  $y$  в  $r_i$  при каждом значении  $x_i$  (третья строка):

$$P = \begin{bmatrix} x_1 & x_2 & \dots & x_n \\ r_1 & r_2 & \dots & r_n \\ \dots & \dots & \dots & \dots \\ (y_{11} \dots y_{1r}) & (y_{22} \dots y_{2r}) & \dots & (y_{n1} \dots y_{nr}) \end{bmatrix} \quad (4.9)$$

$$\sum_{i=1}^n r_i = N$$

Геометрическая интерпретация матрицы П представлена на рис.4.1 для  $m=2, n=4, r_i=3$ . Поверхность отклика определяется искомой функцией  $\varphi(x)$  и нам известна.

Крестиками отмечены значения  $y_{ij}$ .

Всего в результате эксперимента - проводится  $\sum_{i=1}^n r_i = N$  изме-

рений отклика  $y$  в  $n$  различных точках некоторой области  $Q \in X$  пространства факторов по  $r_i$  параллельных измерения в каждой точке.

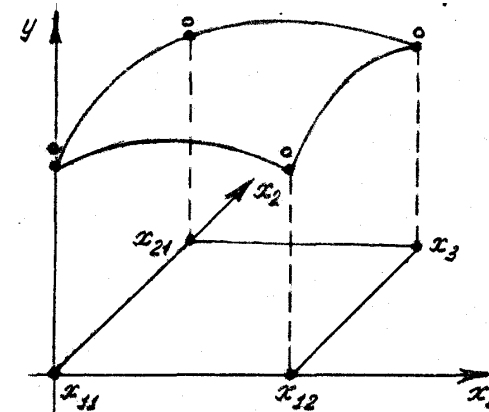


Рис. 4.1. Вид поверхности отклика

Будем пока считать, что все  $r_i=1$ , т.е. что параллельных опытов не проводится. Общий случай, как будет отмечено ниже, можно свести к рассматриваемому. Тогда с учетом допущения (4.4) результаты  $N$  измерений можно представить в виде системы  $N$  равенств

$$y_i = \sum_{k=0}^l \beta_k f(x_k) + \epsilon_i, \quad i=1, 2, \dots, N \quad (4.10)$$

или

$$y_i = \beta^T f(x_i) + \epsilon_i, \quad i=1, 2, \dots, N \quad (4.11)$$

Если на основе системы равенств (4.10), которая содержит в себе всю информацию, полученную в эксперименте, мы сумеем оценить наилучшим образом неизвестные параметры  $\beta_k$  т.е. найти значения  $b_k \approx \beta_k$ , то тем самым, будет найдена приближенная модель следующего вида:

$$y_i = \sum_{k=0}^l b_k f(x_k) \quad (4.12)$$

Эта модель будет наилучшей в классе  $Z$  выбранного базиса  $\{f_k(x)\}$  и числа базисных функций  $n$  (т.е. погрешность  $\Delta^*$  будет минимальной).

Однако остается ещё погрешность модели  $\delta$ , уменьшить которую, в принципе, можно, если удачно подобрать функции  $\{f_k(x)\}$  и их количество  $n$ . Таким образом, модель (4.12) требует в общем случае проверки на адекватность (на соответствие результатам эксперимента) и при необходимости уточнения, что будет рассмотрено ниже. Вернемся к системе равенств (4.10) и покажем, как можно найти наилучшие оценки  $b_k$  параметров  $\beta_k$ . С этой целью представим систему (4.10) в матричной форме, введя следующие матрицы:

матрицу значений отклика размера  $(N \times 1)$

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_N \end{bmatrix} (N \times 1)$$

Если рассматривается план эксперимента с повторными наблюдениями, то вместо  $Y$  рассматривают матрицу средних значений

отклика  $\bar{Y}$  размера  $(N \times 1)$ , где  $\bar{y}_i = \frac{1}{r_i} \sum_{j=1}^{r_i} y_{ij}$

$$Y = \begin{bmatrix} \bar{y}_1 \\ \bar{y}_2 \\ \dots \\ \bar{y}_N \end{bmatrix} (N \times 1)$$

матрицу независимых переменных размера  $(N \times m)$ , где  $m=n+1$

$$X = \begin{bmatrix} f_0(x_{11}) & f_0(x_{21}) & \dots & f_0(x_{m1}) \\ f_0(x_{12}) & f_0(x_{22}) & \dots & f_0(x_{m2}) \\ \dots & \dots & \dots & \dots \\ f_0(x_{1N}) & f_0(x_{2N}) & \dots & f_0(x_{mN}) \end{bmatrix} (4.13)$$

$(N \times m)$ , где  $m=n+1$

матрицу (вектор) - столбец независимых параметров размера  $(m \times 1)$

$$\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \dots \\ \beta_m \end{bmatrix} (m \times 1)$$

матрицу (вектор) столбец ошибок размера  $(N \times 1)$

$$\varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \dots \\ \varepsilon_N \end{bmatrix} (N \times 1)$$

Напомним, что произведение матрицы-строки  $f(x)=[f_0(x), f_1(x), \dots, f_m(x)]$  на матрицу-столбец  $\beta$ , есть сумма произведений соответствующих

элементов  $\sum_{k=0}^m f_k(x) \beta_k$ .

Следовательно, систему (4.10) можно представить в виде

$$Y = X \beta + \varepsilon \quad (4.14)$$

Уравнение (4.14) называется *линейной* по параметру  $\beta$  *регрессионной моделью*, и неизвестные параметры  $\beta_0, \beta_1, \dots, \beta_m$  - *коэффициентами регрессии*.

Смысл матричного уравнения (4.14) можно пояснить его развернутой формой записи

$$\begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_N \end{bmatrix} = \begin{bmatrix} \sum_k^m f_k(x_1) \times \beta_k \\ \sum_k^m f_k(x_2) \times \beta_k \\ \dots \\ \sum_k^m f_k(x_N) \times \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \dots \\ \varepsilon_N \end{bmatrix} \quad (4.15)$$

Для отыскания наилучшей оценки вектора  $\beta$  методом наименьших квадратов нужно сделать некоторые предположения, которые составляют основу регрессионного анализа.

1. *Предположения* об ошибках  $\varepsilon_i$  ( $i=1, 2, \dots, N$ ):

$M\epsilon_i = 0$  – т.е. систематическая ошибка отсутствует;  
 $M\epsilon_i\epsilon_j$  - ошибки некоррелированы (это ограничение можно снять, если матрица ковариаций ошибок известна);  
 $D\epsilon_i = \sigma^2$  - точках факторного пространства  $X$  измерения производятся с одинаковой точностью.

Последнее ограничение можно ослабить, потребовав лишь, чтобы  $D\epsilon_i = \sigma^2 h_i$ , т.е. чтобы при разной точности измерения в различных точках  $X$  коэффициенты пропорциональности  $h_i$  были известны.

2. *Предположение* о факторах  $(x_1, x_2, \dots, x_n) = X^T$ .

Значения переменных  $x_i$  в процессе эксперимента измеряются практически без ошибок (по сравнению с ошибкой измерения отклика  $y$ ). Это означает, что из соотношений (3.8) вытекают равенства

$$My_i = \sum_{k=0}^m f_k(x) \beta_k ; Dy_i = D\epsilon_i = \sigma^2,$$

ибо поскольку  $x_i$ - неслучайные величины, то неслучайной является

$$\sum_{k=0}^m f_k(x) \beta_k .$$

Дисперсия  $Dy_i = \sigma^2$  называется *дисперсией воспроизводимости отклика* и обозначается далее  $\sigma_y^2$ . Она доказывает, насколько велико рассеяние отклика в различных точках факторного пространства  $X$ .

*Подчеркнем*, что никаких предположений о законе распределения пока не делаем, так как для оценки вектора параметров  $\beta$  они не нужны. Однако при статистическом анализе модели (3.9) они потребуются.

**4.2 Оценка параметров линейной модели.**

**Метод наименьших квадратов.**

Матрицы  $F$  и  $Y$  в линейной модели (3.10) содержат всю информацию, полученную в результате эксперимента. По этим данным нам нужно оценить вектор неизвестных параметров  $\beta = \beta_0, \beta_1, \dots, \beta_m)^T$

Обозначим некоторую оценку  $\hat{\beta}$  через  $b$ .

*Напомним*, что:

- оценка  $b$  называется *несмещенной* для  $\beta$ , если  $Mb = \beta$
- оценка называется *наилучшей* (эффeктивной), если её дисперсионная матрица  $Db = \min$  среди всех других оценок (сравнение матриц производится по каждому элементу).

Метод наименьших квадратов (МНК) дает способ получения наилучшей несмещенной оценки в классе линейных оценок (т.е. зависящих от значений  $y_i$ ). Это утверждение составляет содержание *теоремы Гаусса-Маркова* [5,8,48], которую мы приведем без доказательства.

*Теорема 3.1.* Наилучшие линейные несмещенные оценки для параметра  $\beta^T = \beta_0, \beta_1, \dots, \beta_m$  в линейной модели (4.14) определяется согласно МНК матричным равенством

$$b = (F^T F)^{-1} F^T Y, \text{ или } b = M^{-1} Y \quad (4.16)$$

где  $M = F^T F$  - так называемая *информационная матрица Фишера*, которая предполагается неособенной (т.е. её определитель  $\det M \neq 0$ ), а, следовательно, обратимой;

$M^{-1}$  - обратная матрица к матрице  $M$ ;

$Y = F^T Y$  - преобразованная матрица наблюдений.

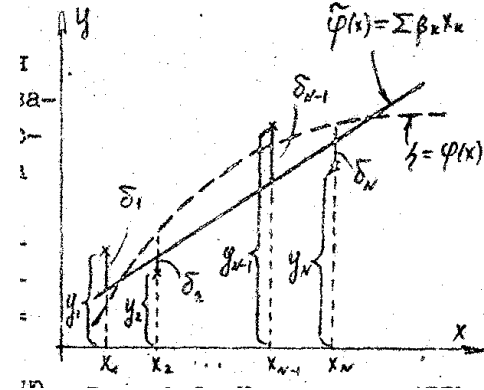


Рис. 4.2. Иллюстрация МНК при  $m=1$ .

Поясним идею метода наименьших квадратов (для случая  $m=1$ ) и происхождение формулы (3.11). Пусть  $m=1$  и  $\eta = \beta_0 + \beta_1 x$ .

На рис. 4.2 кривая, изображенная пунктиром - график неизвестной зависимости  $\eta = \varphi(x, \beta)$ , а крестиками обозначают значения отклика

$y$ , полученные в эксперименте (в точках  $x_1, x_2, \dots, x_n$ ). Неизвестную функцию  $\varphi(x, \beta)$  мы приближенно заменяем функцией

$$y^* = \varphi^*(x) = \sum_{k=0}^l \beta_k f(x_k) \text{ (класс таких функций } Z \text{) и хотим найти та-}$$

кую линейную функцию (т.е. подобрать параметры  $\beta_0, \beta_1, \dots, \beta_m$ ),



$$F = \begin{pmatrix} N & \sum_{i=1}^N x_{1i} & \sum_{i=1}^N x_{2i} & \sum_{i=1}^N x_{1i}x_{2i} \\ \sum_{i=1}^N x_{1i} & \sum_{i=1}^N (x_{1i})^2 & \sum_{i=1}^N x_{1i}x_{2i} & \sum_{i=1}^N (x_{1i})^2 x_{2i} \\ \sum_{i=1}^N x_{2i} & \sum_{i=1}^N x_{1i}x_{2i} & \sum_{i=1}^N (x_{2i})^2 & \sum_{i=1}^N x_{1i}(x_{2i})^2 \\ \sum_{i=1}^N x_{1i}x_{2i} & \sum_{i=1}^N (x_{1i})^2 x_{2i} & \sum_{i=1}^N x_{1i}(x_{2i})^2 & \sum_{i=1}^N (x_{1i}x_{2i})^2 \end{pmatrix} \quad (4.20)$$

Вид этой матрицы (4.5) полностью аналогичен виду определителя системы (4.18).

Информационная матрица  $M$  равняется

$$M = F^T \times F \quad (4.21),$$

где  $F^T$ -транспонированная матрица независимых переменных.

После того, как по выражению (4.21), с помощью изложенных выше правил, найдена информационная матрица  $M$ , нужно определить элементы  $s_{uo}^v$  и по ним найти  $S$ . Элементы  $s_{uo}^v$  находятся из обратной матрицы  $V$  по следующему правилу.

**Следствие 3.1.** Если матрица независимых переменных  $F$  квазиортогональна (т.е. скалярное произведение разных столбцов равно нулю), то матрица  $M = F^T F$  является диагональной и все коэффициенты  $b_0, b_1, \dots$  линейной модели оцениваются независимо по формуле, которая вытекает из формулы (4.16):

$$b_i = \sum_{i=1}^N \varphi_j(x_i) y_i / \sum_{i=1}^N \varphi_j^2(x_i) \quad \text{или} \quad b_i = \frac{1}{N} \sum_{i=1}^N \varphi_j(x_i) y_i \quad (4.22)$$

$$\varphi_j(x_i) = \pm 1, \quad j=1,2,\dots, m$$

Поясним, что матрицу  $F$  можно сделать квазиортогональной, а, следовательно, независимо оценить параметры  $b_0, b_1, \dots$  по формуле (4.22), если специальным образом выбрать точки  $x_i$  в пространстве факторов  $X$ . Такой способ выбора точек  $x_i$  называется *ортогональным планированием* и будет рассмотрен ниже.

Решить систему 4.18. в общем виде нельзя. Поэтому надо задать конкретным видом функции  $Y^*=f(x, b_0, b_1, \dots)$

**Пример.3.1.** Пусть число независимых параметров  $m=1, \eta = \beta_0 \beta_1 x$  (т.е число коэффициентов  $l=2, \frac{\partial \eta}{\partial \beta_0} = 1, \frac{\partial \eta}{\partial \beta_1} = x$ )

а) Пусть отклик  $y$  измерен в  $N$  произвольных точках  $x_1, x_2, \dots, x_N$ . Найдем оценки  $b_0$  и  $b_1$ , по формула (3.11), учитывая, что

$$Y = \begin{bmatrix} y_1 \\ \dots \\ y_N \end{bmatrix}, \quad b = \begin{bmatrix} b_0 \\ b_1 \end{bmatrix}, \quad F = \begin{bmatrix} 1 & x_1 \\ \dots & \dots \\ 1 & x_N \end{bmatrix}, \quad Y = F^T Y = \begin{bmatrix} \sum_{i=1}^N y_i \\ \sum_{i=1}^N x_i y_i \end{bmatrix}$$

Имеем:

$$M = F^T F = \begin{bmatrix} N & \sum_{i=1}^N x_i \\ \sum_{i=1}^N x_i & \sum_{i=1}^N x_i^2 \end{bmatrix}, \quad M^{-1} = (F^T F)^{-1} = \frac{1}{\det M} \begin{bmatrix} \sum_{i=1}^N x_i^2 & -\sum_{i=1}^N x_i \\ -\sum_{i=1}^N x_i & N \end{bmatrix},$$

$$\det M = N \sum_{i=1}^N x_i^2 - \left( \sum_{i=1}^N x_i \right)^2$$

По методу наименьших квадратов требуется определить коэффициенты линейного уравнения регрессии

$$Y^* = b_0 + b_1 x \quad (4.23)$$

Система нормальных уравнений для этого случая имеет вид

$$\begin{cases} \sum_{i=1}^N y_i - \sum_{i=1}^N (b_0 + b_1 x_i) = 0 \\ \sum_{i=1}^N y_i x_i - \sum_{i=1}^N (b_0 + b_1 x_i) x_i = 0 \end{cases}$$

Или

$$\begin{cases} b_0 N + b_1 \sum_{i=1}^N x_i = \sum_{i=1}^N y_i \\ b_0 \sum_{i=1}^N x_i + b_1 \sum_{i=1}^N x_i^2 = \sum_{i=1}^N y_i x_i \end{cases} \quad (4.24)$$

Коэффициенты и легко найти с помощью определителей

$$b_0 = \frac{\begin{vmatrix} \sum_{i=1}^N y_i & \sum_{i=1}^N x_i \\ \sum_{i=1}^N y_i x_i & \sum_{i=1}^N x_i^2 \end{vmatrix}}{\begin{vmatrix} N & \sum_{i=1}^N x_i \\ \sum_{i=1}^N x_i & \sum_{i=1}^N x_i^2 \end{vmatrix}} = \frac{\sum_{i=1}^N y_i \sum_{i=1}^N x_i^2 - \sum_{i=1}^N y_i x_i \sum_{i=1}^N x_i}{N \sum_{i=1}^N x_i^2 - (\sum_{i=1}^N x_i)^2} \quad (4.25)$$

$$b_1 = \frac{\begin{vmatrix} N & \sum_{i=1}^N y_i \\ \sum_{i=1}^N x_i & \sum_{i=1}^N y_i x_i \end{vmatrix}}{\begin{vmatrix} N & \sum_{i=1}^N x_i \\ \sum_{i=1}^N x_i & \sum_{i=1}^N x_i^2 \end{vmatrix}} = \frac{N \sum_{i=1}^N y_i x_i - \sum_{i=1}^N y_i \sum_{i=1}^N x_i}{N \sum_{i=1}^N x_i^2 - (\sum_{i=1}^N x_i)^2} \quad (4.26)$$

Или это же можно записать в матричной форме

$$b = \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} = M^{-1}Y = \frac{1}{\det M} \begin{bmatrix} \sum_{i=1}^N x_i^2 \sum_{i=1}^N y_i - \sum_{i=1}^N x_i (\sum_{i=1}^N x_i y_i) \\ - \sum_{i=1}^N x_i \sum_{i=1}^N y_i + N \sum_{i=1}^N x_i y_i \end{bmatrix},$$

откуда получаем те же формулы (4.25) и (4.26)

Кoeffициент  $b_0$  проще найти по известному  $b_1$  из системы (3.14)

$$b_0 = \bar{y} - b_1 \bar{x}, \quad (4.27)$$

где  $\bar{y}$  и  $\bar{x}$  средние значения  $y$  и  $x$ . Уравнение (4.27) показывает, что между коэффициентами  $b_0$  и  $b_1$  существует *корреляционная зависимость*. Для оценки силы линейной связи между коэффициентами вычисляется выборочный коэффициент корреляции  $r^*$

$$r^* = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{(N-1)S_x S_y} \quad (4.28)$$

где  $S_x$  и  $S_y$  выборочные среднеквадратичные отклонения.

б) Пусть теперь  $N=2n$  и значения  $y_1, y_2, \dots, y_n$  получены при  $x_1=1$ , а  $y_{n+1}, y_{n+2}, \dots, y_{2n}$  при  $x_2=1$ ; тогда матрица независимых переменных квазиортогональна:

$$F = \left\{ \begin{matrix} \begin{bmatrix} 1+1 \\ 1+1 \\ \dots \\ 1-1 \end{bmatrix} \\ n \end{matrix} \right\} \Rightarrow F^T F = M = \begin{bmatrix} N & 0 \\ 0 & N \end{bmatrix} \Rightarrow M^{-1} \Rightarrow \frac{1}{N} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

$$F^T Y = Y = \begin{bmatrix} \sum_{i=1}^N y_i \\ \sum_{i=1}^N x_i y_i \end{bmatrix}$$

Следовательно, приходим к формулам (4.22), которые дают отдельные

оценки:  $b_0 = \frac{1}{N} \sum_{i=1}^N y_i = \bar{y}$ ,  $b_1 = \frac{1}{N} \sum_{i=1}^N x_i y_i$

Зная оценки  $b_0$  и  $b_1$ , находим модель (4.12):  $Y^* = b_0 + b_1 x$ .

Определение точности оценок  $b_0, b_1, \dots$  и  $y^*(x)$

Точность оценки  $b_0, b_1, \dots$  характеризуется *ковариационной матрицей*

$$D[b] = M[(b-\beta)(b-\beta)^T] = \begin{bmatrix} Db_0 & k_{01} & \dots & k_{0l} \\ k_{i0} & Db_1 & & \\ \dots & \dots & \dots & \dots \\ k_{l0} & k_{l1} & \dots & Db_l \end{bmatrix}$$

$D$  - дисперсия оценки  $b_i$ ;  $k_{ij}$  - ковариация оценок  $b_i$  в  $b_i$

которая равна  $\rho_{ij} = \frac{k_{ij}}{\sqrt{Db_i} \sqrt{Db_j}}$ ;  $\rho_{ij}$  - коэффициент корреляции.

*Правило* определения матрицы  $D[b]$  при известной дисперсии воспроизводимости  $\sigma_y^2 = \sigma^2$  дает следующая теорема.

**Теорема 3.2.** Матрица ковариаций вектора МНК-оценок  $b^T = (b_0, b_1, \dots, b_l)$  линейной модели (4.14) имеет вид

$$D[b] = \sigma^2 M^{-1} \text{ или } D[b] = \sigma^2 C, \quad (4.29)$$

где  $\sigma^2$  - дисперсия воспроизводимости отклика (будет оценена позже);  $C = M^{-1} = (F^T F)^{-1}$  - матрица, обратная к информационной матрице  $M$ , называемая *матрицей Фишера*.

Зная матрицу ковариаций (4.29), можно найти дисперсии всех оценок  $b_i$  и их ковариаций  $k_{ij}$  или коэффициенты корреляции  $\rho_{ij}$

Именно, если  $C = [c_{ij}]$ , то  $Db_{ii} = \sigma^2 c_{ii}$ , а  $k_{ij} = \sigma^2 c_{ij}$ , или

$$\rho_{ij}^2 = c_{ij}^2 / c_{ii} c_{jj}$$

**Следствие 3.2.** Если  $F$  - квазиортогональная матрица и  $\varphi_k(x_i) = \pm 1$  (что всегда можно сделать преобразованием исходных переменных

$x_k$ ), то матрица  $C = \frac{1}{N} E$ , где  $E$  - единичная матрица ( $N \times N$ ) и, следовательно, из (4.30) вытекает

$$Db_i = \frac{1}{N} \sigma^2 \quad (4.30)$$

т.е. оценки  $b_0, b_1, \dots$  некоррелированы и их точность (измеряемая дисперсия) обратно пропорциональна числу измерений  $N$  (в различных точках  $X$ ).

Для определения точности оценки отклика  $y^*(x)$  используется дисперсия  $Dy^*(x)$ , определяемая согласно следующей теореме.

**Теорема 3.3.** Дисперсия оценки отклика  $y^*(x)$  в произвольной точке  $x \in X$  определяется по формуле

$$Dy^*(x) = \sigma^2 f^T(x) C f(x), \quad (4.31)$$

где  $f^T(x) = \{f_0(x), \dots, f_l(x)\}$  - значение вектора базисных функций в точке  $x \in X$ .

**Следствие 3.3.** Если  $F$  - квазиортогональная матрица и

$f_k(x_i) = \pm 1$ , то  $C = \frac{1}{N} E$ , и, следовательно,

$$Dy^*(x) = \frac{\sigma^2}{N} f^T(x) \times f(x) = \frac{\sigma^2}{N} \sum_{k=0}^l f_k^2(x) \quad (4.32)$$

из формулы (4.32) следует, что по мере удаления от центра области эксперимента (т.е. от точки  $x=0$ , где  $f_k(0)=0$   $k=1, 2, \dots, l$ ) дисперсия оценки отклика возрастает.

**Пример 3.2.** Начнем точность оценок  $b_0$  и  $b_1$ , модели отклика 1-го порядка  $Y^*(x)$

$$y^*(x) = b_0 + b_1 x$$

Используя соотношение (3.13), можно найти, что

$$Db_0 = \frac{\sigma^2}{N \sum_{i=1}^N (x_i - \bar{x})^2} \times \sum_{k=0}^N x_i^2, \quad Db_1 = \frac{\sigma^2}{N \sum_{i=1}^N (x_i - \bar{x})^2}, \quad k_{01} = \text{cov}(b_0 b_1) = \frac{\sigma^2 \bar{x}}{N \sum_{i=1}^N (x_i - \bar{x})^2}.$$

Используя соотношение (4.32), найдем, что

$$D_{y^*(x)} = \sigma^2 \left[ \frac{1}{N} + \frac{(x_i - \bar{x})^2}{\sum_{i=1}^N (x_i - \bar{x})^2} \right] \quad (4.33)$$

$$\text{где } \bar{x} = \frac{1}{N \sum_{i=1}^N x_i}.$$

Если дополнительно найдена оценка  $S_y^2$  дисперсии воспроизводимости  $\sigma^2$ , то, используя  $D_{y^*(x)}$ , можно найти приближенные доверительные границы уровня  $\gamma$  для отклика  $y^*(x) = b_0 + b_1 x$  в каждой точке  $X$  следующего вида:



$$\overline{y(x)} = y^*(x) + Z_{1-\frac{\alpha}{2}} \sqrt{\frac{1}{N} + \frac{(x_i - \bar{x})^2}{\sum_{i=1}^N (x_i - \bar{x})^2}},$$

$$\underline{y(x)} = y^*(x) - Z_{1-\frac{\alpha}{2}} \sqrt{\frac{1}{N} + \frac{(x_i - \bar{x})^2}{\sum_{i=1}^N (x_i - \bar{x})^2}}$$

изображенные на рис. 4.3 штриховыми линиями ( $Z_{1-\alpha/2}$  - квантиль нормального закона уровня  $1 - \alpha/2$ ,  $\alpha-\gamma$ ). Как видим, точность прогноза на основе регрессионной модели  $y^*(x) = b_0 + b_1x$  падает по мере удаления от области проведения эксперимента  $\Omega$ . Этот вывод остается в силе для любого числа факторов  $m$ .

*Замечание 3.1. Оценивание параметров модели при повторных наблюдениях.* Мы предполагали, что все  $N$  измерений отклика проводятся в различных точках пространства факторов  $X$  (т.е. число повторных или, как говорят, параллельных опытов в каждой точке  $r_i = 1$ ). Все расчетные формулы

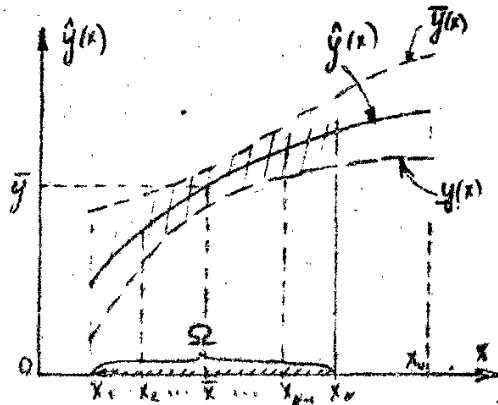


Рис. 4.3. Доверительные границы  $y(x)$  и  $\underline{y(x)}$  для отклика (4.16), (4.22) - (4.31), (4.32), - приведены для этого случая. Что изменится, если число различных точек будет  $n=N/r$ , где  $r$  - кратность повторных наблюдений.

Можно показать [1], что все расчетные формулы сохраняются, если вместо матрицы значений отклика  $y$ , везде использовать матрицу средних значений отклика  $\overline{y}$  размера  $(n-1)$ , где  $i$ -элемент равен  $y_i = \frac{1}{r} \sum_{j=1}^r y_{ij}$  - среднему значению из  $r$  повторных наблюдений в  $i$ -й точке  $x_i=1, \dots, n$ . При этом везде вместо  $N$  будет стоять  $n$ , а вместо  $\sigma^2$  -  $\sigma^2/r$ , так как дисперсия среднего арифметического  $\overline{y_i}$   $r$  наблюдений в  $r$  раз меньше исходной дисперсии  $y$ . Ясно также, что размерность матрицы  $F$  будет не  $(N \times l)$ , а  $(n \times l)$ .

### 4.3 Статистический анализ регрессионной модели

*Статистический анализ модели* (4.12) состоит из следующих этапов:

- оценка дисперсии воспроизводимости  $\sigma^2 = \sigma_y^2$ ;
- проверка адекватности модели;
- оценка значимости коэффициентов модели.

Для решения этих задач необходимо сделать дополнительные, к уже сделанным двум, еще одно предположение о законе распределения случайной величины  $\epsilon_i$ . Будем считать, что

$$\epsilon_i \sim N(0, \sigma), i=1, \dots, n, \tag{4.34}$$

т.е. в каждой точке  $x_i \in X$  ошибка  $\epsilon_i$  имеет нормальный закон распределения с параметрами  $0$  и  $\sigma$ . Отсюда следует в силу равенства (4.3) и предположений регрессионного анализа, что измерение в  $i$ -й точке  $x_i \in X$  значение отклика  $y_i$  имеет нормальный закон распределения с параметрами  $M y_i = \varphi(x)$  и  $D y_i = \sigma^2$ , т.е.

$$y_i \sim N[\varphi(x), \sigma^2]. \tag{4.35}$$

Это допущение позволяет использовать для решения перечисленных задач аппарат теории проверки статистических гипотез.

#### Оценка дисперсии воспроизводимости модели

При оценке точности модели по формулам (4.23), (4.24) требуется знание дисперсии воспроизводимости  $\sigma_y^2$ , причем при её оценке нужно одновременно проверять первое допущение о постоянстве  $\sigma_y^2$  в различных точках  $X$ .

Чтобы оценить дисперсию воспроизводимости  $\sigma^2 = \sigma_y^2$ , необходимо иметь параллельные наблюдения в каждой точке  $x_i \in X$ ,  $i = 1, 2, \dots, n$ , т.е. матрица результатов эксперимента должна иметь вид (4.9).

Если  $y_{i1}, y_{i2}, \dots, y_{ir_i}$  - значения отклика в точке  $x_i$  то наилучшая оценка дисперсии  $y_i$  в этой точке  $\sigma_i^2 = Dy_i$  имеет вид (при  $r_i = r$ )

$$\sigma_i^2 \cong S_i^2 = \frac{1}{r-1} \sum_{j=1}^r (y_{ij} - \bar{y})^2 \quad (4.25)$$

Для проверки справедливости допущения о постоянстве  $\sigma_y^2$  проверяем гипотезу  $H_0: \sigma_i^2 = \sigma^2$  по критерию Кохрена (на некотором уровне значимости  $\alpha$ ). Если  $H_0$  принимается, то общую оценку дисперсии  $\sigma_y^2$  можно получить усреднением оценок  $S_i^2$ , т.е.

$$\sigma^2 = S^2 = \frac{1}{n(r-1)} \sum_{i=1}^n \sum_{j=1}^r (y_{ij} - \bar{y})^2 \quad (4.27)$$

При отклонении  $H_0$  нужно провести анализ причин и принять меры для обеспечения однородности  $Dy_i = \sigma_i^2$  (например, сужая область проведения эксперимента  $\Omega \in X$ ).

Заметим, что возможность проверки гипотезы  $H_0$  обеспечена предположением о нормальности распределения ошибок (4.26).

Замечание 3.2. Если  $r_i \neq r$ , то для проверки однородности нужно использовать *критерий Бартлета*, а усреднение оценок  $S_i^2$  производить с учетом числа повторных наблюдений:

$$S^2 = \sum_{i=1}^n \alpha_i S_i^2, \quad \alpha_i = (r_i - 1) / \sum_{i=1}^n (r_i - 1). \quad (4.28)$$

#### Проверка адекватности модели

Наша модель имеет вид (4.12)

$$y^*(x) = \sum_{k=0}^l b_k \varphi_k(x)$$

и является наилучшей в аппроксимирующем пространстве  $Z$ , зависящем от вида и числа базисных функций  $\{\varphi_k(x)\}$  в том смысле, что погрешность

$$\delta = \|y_i - y^*_i\| = \sum_{k=1}^n [\bar{y} - y^*(x_i)]^2 = \min$$

Однако это не означает, что величина  $\delta$  нас удовлетворяет и модель (4.12) может быть признана адекватной (т.е. хорошо согласующейся с наблюдениями  $y_{ij}$ ). Все зависит от того, насколько погрешность  $\delta$  велика на фоне «шума», связанного с погрешностями эксперимента  $\varepsilon_i$ . Общую погрешность модели (4.12), называемую остаточной суммой квадратов - отклонений  $Q_{ост}$ , можно представить в виде

$$\sum_{i=1}^n \sum_{j=1}^{r_i} (y_{ij} - y_i^*)^2 = \sum_{i=1}^n r_i (\bar{y}_i - y_i^*)^2 + \sum_{i=1}^n \sum_{j=1}^{r_i} (y_{ij} - \bar{y}_i)^2 \quad (4.29)$$

где  $Q_1 = \sum_{i=1}^n r_i (\bar{y}_i - y_i^*)^2 = \delta$  - сумма квадратов отклонений, обусловленная неадекватностью модели;

$Q_2 = \sum_{i=1}^n \sum_{j=1}^{r_i} (y_{ij} - \bar{y}_i)^2$  - сумма квадратов отклонений отклика, обусловленная погрешностями эксперимента.

Если  $Q_1$  невелика по сравнению с  $Q_2$ , то это и означает, что расширять пространство  $Z$  за счет, например, увеличения числа членов разложения  $l$  в модели (3.9) нецелесообразно и модель следует признать адекватной. Эти качественные рассуждения на языке математической статистики означают проверку гипотезы адекватности модели  $H_0$ , состоящей в том, что  $My = F\beta$  против альтернативы  $H_1: My \neq F\beta$ , где  $F$  - выбранная нами матрица независимых переменных, определяющая вид модели (4.14).

Можно показать [5,20], что независимо от того, истинна гипотеза  $H_0$  или нет, величина  $V_i = Q_2 / \sigma^2$  имеет  $\chi^2$ -распределение с

$\sum_{i=1}^n (r_i - 1)$  степенями свободы. Величина

$$S_y^2 = Q_2 / \sum_{i=1}^n (r_i - 1) \quad (4.30)$$

совпадает до виду с оценкой (4.28) и является несмещенной оценкой дисперсии воспроизводимости  $\sigma_i^2 = \sigma^2$ .

Аналогично можно показать [18, 20] что если  $H_0$  истинна, то  $V_t = Q_1 / \sigma^2$  имеет  $\chi^2$ -распределение с  $n-l-1$  степенями свободы, где  $l' = l+1$  - число неизвестных параметров модели (4.14). При этом величина  $S_{ad} = Q_1 / (n-l-1)$  также является несмещенной оценкой  $\sigma_y^2 = \sigma^2$  и называется *дисперсией, связанной с неадекватностью модели*. Таким образом, если гипотеза  $H_0$  истинная, то отношение

$$T = S_{ad} / S_y^2 \sim F(v_1, v_2), \quad v_1 = n-l-1; \quad v_2 = \sum_{i=1}^n (r_i - 1)$$

т.е. имеет F-распределение Фишера с  $v_1$  и  $v_2$  степенями свободы и, следовательно, проверка  $H_0$  осуществляется стандартным образом по критерию Фишера. А именно, если найдено табличное значение критерия Фишера

$F_T(v_1, v_2, \alpha)$ , отвечающее уровню значимости  $\alpha$ , то  $H_0$  принимается при

$$S_{ad} / S_y^2 \leq F_T(v_1, v_2, \alpha)$$

и отклоняется в противном случае. Отклонив  $H_0$ , мы должны построить более сложную модель, увеличив, например, число базисных функций или, быть может, выбрав другой их набор  $\{\varphi_k(x)^*\}$

#### Оценка значимости коэффициентов модели

Некоторые из базисных функций  $\varphi_k(x)$  могли быть включены ошибочно в модель (4.12), т.е. на самом деле отклик  $y$  от  $\varphi_k(x)$  не зависит, и потому соответствующий коэффициент  $\beta_k$ , должен быть равен нулю. Но при этом оценка  $b_k \neq 0$ , хотя и близка к нему. Проверка значимости коэффициента означает проверку гипотезы  $H_0: \beta_k = 0$  против альтернативы  $H_1: \beta_k \neq 0$ . Если  $b_k \cong 0$  и гипотеза  $H_0$  принимается, то коэффициент  $\beta_k$  считается незначимым и соответствующий член из модели (4.12) выбрасывается.

Проверка  $H_0$  основана на том, что  $b_k$  имеет нормальный закон распределения с  $Mb_k = 0$  и  $Db_k = \sigma_k^2$  (где  $\sigma_k^2$  есть  $k$ -й элемент на диагонали матрицы ковариации  $Db_k$ ), так как оценка  $b_k$  линейно зависит от наблюдений  $y_i$  (или  $\bar{y}_i$ ), которые по предположению распределены нормально. Следовательно, можно использовать критерий Стьюдента

$$T_k = \frac{b_k - Mb_k}{\sqrt{Db_k}} = \frac{b_k \sqrt{N}}{S_y} \sim S(v=N) \quad (4.31)$$

где  $N = \sum_{i=1}^n r_i$  - общее число опытов.

Если найдено табличное значение  $T(N, \alpha)$  критерия Стьюдента, отвечающее  $N = \sum_{i=1}^n r_i$  степеням свободы  $N = ng$  при  $g = r_i$  и уровню значимости  $\alpha$ , то  $H_0$  принимается при  $|b_k| \sqrt{N} / S_y \leq T_T(N, \alpha)$  и отклоняется в противном случае.

Регрессионная модель, прошедшая проверку на адекватность и значимость коэффициентов, может быть использована для решения различных практических задач. Основными из них являются:

- нахождение экстремальных условий протекания процесса, модель которого построена;
- определение значения отклика в той части факторного пространства, где эксперимент не проводился, т.е. либо интерполяция, либо экстраполяция (прогнозирование) отклика.

Отметим в заключение кратко вопросы, касающиеся построения математических моделей по результатам уже проведенного эксперимента, которые остались вне нашего рассмотрения.

#### 1. Построение моделей нелинейных по параметрам.

С такими моделями приходится встречаться довольно часто. Их исследование значительно сложнее. Для отыскания МНК - оценок параметров таких моделей приходится использовать численные методы оптимизации [2. 8, 20].

#### 2. Построение моделей с векторным откликом.

Если вектор отклика  $y^T = (y_1, y_2, \dots, y_m)$  состоит из некоррелированных компонентов, то все сводится к построению отдельных моделей для каждого компонента. Сложность возникает, когда  $y_1, y_2, \dots, y_m$  связаны между собой. Этот случай разобран в [22].

#### 3. Плохая обусловленность информационной матрицы $M = F^T F$ .

Это означает, что определитель  $\det M \approx 0$  и обращение матрицы для отыскания МНК - оценок вектора параметров  $\beta$  либо вообще невозможно, либо сильно затруднено. В таком случае нужно переходить к моделям *неполного ранга* [20], либо использовать *метод регуляризации* [8]. Заметим, что эта трудность исчезнет, если эксперимент спланирован заранее.

Пример 3.3. Производительность химического реактора  $y$  зависит в основном (как показали предварительные исследования) от двух факторов: температуры  $T$  и времени протекания реакции  $t$ . В окрестности точки

$T=220^\circ\text{C}$  и  $t=4\text{ч}$  была построена линейная модель 1-го порядка, показавшая, что отклик растет пропорционально росту  $T$  и  $t$  примерно с одинаковой скоростью. Для поиска максимального значения  $y$  было решено исследовать полиномиальную модель 2-го порядка в окрестности точек ( $T=260^\circ\text{C}$ ,  $t=8\text{ч}$ ), т.е. модель вида

$$\eta = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{12} x_1 x_2 \quad (4.32)$$

где  $x_1 = t - 8$ ,  $x_2 = (T - 260)/10$  - кодированные переменные, при которых исходная точка ( $t=8$ ,  $T=260$ ) переходит в точку ( $x_1=0$ ,  $x_2=0$ ), а интервалы варьирования  $\Delta t = 1\text{ч}$  и  $\Delta T = 10^\circ\text{C}$  переходят в  $\Delta x_1 = 1$  и  $\Delta x_2 = 1$ .

Было решено провести эксперимент в  $n=9$  различных точках, продублировав наблюдения в каждой точке  $x_i$  (т.е.  $r_i=r=2$ ). Таким образом, число неизвестных параметров  $l'=l+1=6$ ; число степеней свобода  $v_1=n-l'=3$ ; число степеней свободы  $v_2=n(r-1)=9$

Результаты эксперимента и значения отклика, которые дает модель  $\eta=y^*(x)$ , приведенная ниже, представлены в табл.

Таблица 4.1

$i$	$x_{1i}$	$x_{2i}$	$y_{1i}$	$y_{2i}$	$\bar{y}_i$	$y^*(x_i)$
1	0	0	88,0	88,2	88,1	88,97
2	1	0	89,2	89,1	89,15	89,2
3	0	1	88,2	82,1	82,15	82,3
4	-1	0	83,7	83,8	83,75	83,8
5	0	-1	87,3	87,4	87,35	87,3
6	1	1	82,6	82,7	82,65	82,6
7	1	-1	89,6	89,6	89,6	89,6
8	-1	-1	82,2	82,1	82,15	82,2
9	-1	1	79,3	79,1	79,2	79,1

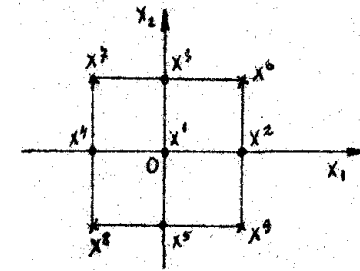


Рис. 4.4. Точки плана эксперимента

На основании табл. 4.1. находим вектор МНК - оценок  $b$  в соответствии с формулами раздела 4.2 данной главы, определив матрицу  $F$ ,  $M=F^T F$ ,  $F^T Y$  и  $C=M^{-1}$

$$F = \begin{bmatrix} x_0 & x_1 & x_2 & x_0 x_1 & x_0 x_2 & x_1 x_2 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & -1 & 0 & 1 & 0 & 0 \\ 1 & 0 & -1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & 1 & 1 & -1 \\ 1 & -1 & -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 & 1 & -1 \end{bmatrix}$$

$$b = (F^T F)^{-1} F^T Y = [87,97; 2,72; -2,52; -1,45; -3,15; -1]^T$$

Тем самым модель (4.32) в кодированных переменных  $x_1$  и  $x_2$  найдена.

$$F^T F = \begin{bmatrix} 9 & 0 & 0 & 6 & 6 & 0 \\ 0 & 6 & 0 & 0 & 0 & 0 \\ 0 & 0 & 6 & 0 & 0 & 0 \\ 6 & 0 & 0 & 6 & 0 & 0 \\ 6 & 0 & 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4 \end{bmatrix}, (F^T F)^{-1} = \begin{bmatrix} 5/9 & 0 & 0 & -1/3 & -1/3 & 0 \\ 0 & 1/6 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/6 & 0 & 0 & 0 \\ 6 & 0 & 0 & 1/2 & 0 & 0 \\ 6 & 0 & 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/4 \end{bmatrix},$$

$$CF^T Y = \begin{bmatrix} 87,97 \\ 2,72 \\ -2,52 \\ -1,45 \\ -3,15 \\ -1,0 \end{bmatrix}$$

Статистический анализ модели (4.32)

Находим вначале

$$Q_1 = \sum_{i=1}^9 r(\bar{y}_i - y_i^*)^2 = 0,121;$$

$$S_{ад}^2 = \frac{1}{v_1} Q_1 = \frac{1}{3} 0,121 = 0,0403$$

$$Q_2 = \sum_{i=1}^9 \sum_{j=1}^2 (y_{ij} - \bar{y}_i)^2 = 0,07$$

$$S_y^2 = \frac{1}{v_2} Q_2 = \frac{1}{9} 0,07 = 0,0078$$

$$S_y^2 = \frac{1}{r_2} S_y^2 = \frac{1}{2} 0,0078 = 0,0039$$

Проверка адекватности модели (при  $\alpha=0,01$ ):

$$F = \frac{S_{ад}^2}{S_y^2} = \frac{0,0403}{0,0078} = 5,2 < F(\alpha, v_1, v_2) = 6,99, \text{ следовательно, модель можно}$$

считать адекватной на уровне значимости  $\alpha=0,001$ ,  $t_T(\alpha, f)=2,9$ ):

$$S_0 = \sqrt{S_y^2 C_{00}} = \sqrt{0,008 \left(\frac{5}{9}\right)} = 0,07 \Rightarrow t_0 = \frac{87,9}{0,07} \gg t_T$$

$$S_1 = \sqrt{S_y^2 C_{11}} = \sqrt{0,008 \left(\frac{1}{6}\right)} = 0,036 \Rightarrow t_1 = \frac{2,72}{0,036} \gg t_T$$

$$S_2 = \sqrt{S_y^2 C_{22}} = \sqrt{0,008 \left(\frac{1}{6}\right)} = 0,036 \Rightarrow t_2 = \frac{2,52}{0,036} \gg t_T$$

$$S_3 = \sqrt{S_y^2 C_{33}} = \sqrt{0,008 \left(\frac{1}{2}\right)} = 0,063 \Rightarrow t_3 = \frac{1,45}{0,063} \gg t_T$$

$$S_4 = \sqrt{S_y^2 C_{44}} = \sqrt{0,008 \left(\frac{1}{2}\right)} = 0,063 \Rightarrow t_4 = \frac{3,15}{0,063} \gg t_T$$

$$S_5 = \sqrt{S_y^2 C_{55}} = \sqrt{0,008 \left(\frac{1}{4}\right)} = 0,063 \Rightarrow t_5 = \frac{1}{0,045} \gg t_T.$$

*Вывод:* все коэффициенты значимы. Доверительный интервал  $\gamma=0,9$  ( $Z_\gamma=1,65$ ) для  $\eta$  имеет вид  $y^*(x) \pm Z_\gamma \sqrt{Dy^*(x)}$ , где

$$Dy^*(x) = S_y^2 \varphi^T(x) C \varphi(x)$$

Для отыскания точки  $x^*$ , в которой  $y^*(x^*) = \max$ , решаем систему уравнений (линейных):

$$\left. \begin{array}{l} \frac{\partial y^*(x)}{\partial x_1} = b_1 + 2b_3x_1 + b_5x_2 = 0 \\ \frac{\partial y^*(x)}{\partial x_2} = b_2 + 2b_3x_2 + b_5x_2 = 0 \end{array} \right\} \begin{array}{l} 2,9x_1 + x_2 = 2,72 \\ -x_1 + 6,3x_2 = -2,5 \end{array} \Rightarrow \begin{array}{l} x_1^* = 1,14 \\ x_2^* = -0,58 \end{array}$$

Возвращаясь к исходным переменным  $t$  и  $T$  находим точку экстремума

$$t^* = x_1^* + 8 = 9,14; \quad T^* = 10x_2^* + 260 = 254,2^\circ\text{C}$$

## 5 ЭЛЕМЕНТЫ ТЕОРИИ ПЛАНИРОВАНИЯ ЭКСПЕРИМЕНТА

Из всего множества задач, которые приходится решать экспериментатору при исследовании интересующего его объекта или процесса, мы выделим следующие две встречающиеся на практике, пожалуй, наиболее часто:

- построение математической модели объекта, представляющей собой аналитическую зависимость между выходной переменной (откликом) и набором входных переменных (факторов);
- поиск оптимальных условий поведения объекта (протекания процесса), т.е. поиск таких значений факторов, при которых отклик (или некоторый функционал от него) достигает экстремума.

До сих пор рассматривались вопросы построения модели по результатам пассивного эксперимента. При этом вторая задача в этих условиях даже не рассматривалась, так как возможность её решения связана именно с целенаправленным поиском точек проведения эксперимента в пространстве факторов. В части II этого пособия рассмотрены обе отмеченные задачи в предположении, что экспериментатор имеет возможность целенаправленно влиять на условия проведения эксперимента (нужным образом устанавливать значения контролируемых переменных и число опытов) или, иными словами, может планировать эксперимент.

### 5.1 Сопоставление возможностей пассивного и активного экспериментов

Рассмотрим построение математической модели исследуемого процесса или явления в виде аналитической зависимости вида

$$y = f(x_1, x_2, \dots, x_m) + \varepsilon, \quad (5.1)$$

где  $y$  - наблюдаемое значение интересующего параметра (отклика);  $x_1, x_2, \dots, x_m$  - исследуемые факторы;  $\varepsilon$  - погрешность модели, связанная как с действием неучтенных факторов, так и с ошибками эксперимента;

$\eta = f(x_1, x_2, \dots, x_m)$  - среднее значение отклика в точке  $X = (x_1, x_2, \dots, x_m)$ . Приближенное определение функции  $f(x_1, x_2, \dots, x_m)$  по результатам пассивного эксперимента является задачей регрессионного анализа. Её решение было рассмотрено выше. При использовании результатов пассивного эксперимента мы предполагаем, что задана матрица плана

$$\Pi = \begin{bmatrix} x_1 & x_2 & \dots & x_i & \dots & x_n \\ r_1 & r_2 & \dots & r_1 & \dots & r_n \end{bmatrix} \quad (5.2)$$

где  $x_i$  - точка факторного пространства  $X$ , где производятся измерения отклика ( $i=1, 2, \dots, n$ );

$r_i$  - число параллельных опытов в точке  $x_i$ , так что  $\sum_{i=1}^n r_i = N$  - общее

число проведённых опытов.

Зная матрицу  $\Pi$  и результаты измерений отклика в виде матрицы  $Y = [\bar{y}_1, \bar{y}_2, \dots, \bar{y}_n]^T$ , в регрессионном анализе рассматривают линейную модель для оценки  $f$  следующего вида:

$$Y = F(\Pi)\beta + \varepsilon \quad (5.3)$$

где матрица независимых переменных

$$F(\Pi) = \begin{bmatrix} f_0(x_{11}) & f_0(x_{21}) & \dots & f_0(x_{m1}) \\ f_0(x_{12}) & f_0(x_{22}) & \dots & f_0(x_{m2}) \\ \dots & \dots & \dots & \dots \\ f_0(x_{1N}) & f_0(x_{2N}) & \dots & f_0(x_{mN}) \end{bmatrix}$$

зависит от выбранного плана эксперимента  $\Pi$ .

Используя МНК, мы нашли наилучшие (для данной матрицы  $\Pi$ ) оценки  $b_k$  неизвестных параметров модели (1.3) и построили модель вида

$$y^*(x) = \sum_{k=0}^l b_k f_k(x) \quad (5.4)$$

На этом достижения пассивного эксперимента заканчиваются.

Спрашивается, а нельзя ли добиться большего путем специального выбора матрицы плана  $\Pi$ , т.е. путем планирования эксперимента? Например, мы знаем, что матрица ковариаций оценок  $D(b) = \sigma_y^2 C(\Pi)$ ,  $C(\Pi) = (F^T F)^{-1}$ , (5.5)

зависит от  $\Pi$  и, следовательно, можно попытаться найти такую матрицу  $\Pi^*$ , при которой  $(D(b, \Pi^*))$  была бы наименьшей в определенном смысле.

Таким образом, целенаправленный выбор матрицы плана П, т.е. проведение активного эксперимента может дать по-видимому (и действительно, дает!), существенный выигрыш в точности определения модели.

Критерий наилучшего выбора матрицы плана П может быть выбран по-разному - в зависимости от целей эксперимента; отсюда и требования к *наилучшей* матрице П могут быть различными. Различные, критерии оптимальности будут рассмотрены ниже.

Но сначала убедимся в том, что выбор матрицы плана П действительно влияет на точность оценок  $b_k$ .

### 5.2 Однофакторный и многофакторный эксперимент, преимущества последнего

Традиционный способ проведения эксперимента для установления зависимости вида (5.3) до появления теории планирования эксперимента (ТПЭ) являлся однофакторным. Это означает, что контролируемые переменные (факторы) варьировались по очереди при фиксировании в данном опыте всех остальных факторов. Оказывается, это не наилучший способ проведения эксперимента!

Если проводить *многофакторный* эксперимент, т.е. менять одновременно уровни нескольких факторов, то можно добиться более точной оценки модели (при том же объеме наблюдений N!), Происходит это за счет более рационального использования пространства факторов X, что можно проиллюстрировать на следующем примере.

Задача о взвешивании. Требуется определить массы А, В и С трёх предметов, используя для этой цели пружинные весы.

Традиционную схему взвешивания (отвечающую схеме однофакторного эксперимента) можно представить в виде табл. 5.1 (где + 1 означает, что предмет находился на весах, а -1, что он не взвешивался).

Таблица 5.1

Номер опыта	А	В	С	Результат взвешивания
1	-1	-1	-1	$y_0$
2	+1	-1	-1	$y_1$
3	-1	+1	-1	$y_2$
4	-1	-1	+1	$y_3$

Здесь вначале делается «холостое» взвешивание для определения нулевой точки весов, а затем по очереди взвешивается каждый из предметов. Модель взвешивания можно записать в виде

$$y = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3, \quad (5.6)$$

где  $x_i$

$$= \begin{cases} 1, & \text{если } i\text{-й из 3 предметов взвешивался,} \\ 0, & \text{в противном случае.} \end{cases}$$

Ясно, что при всех  $x_i = 0$  имеем  $y_0 = \alpha_0$ ; при  $x_1=1, x_2=x_3=0$  имеем

$A = \alpha_1 = y_1 - y_0$  и т.д. - каждая масса есть разности результатов 2 взвешиваний независимо от числа предметов (факторов).

Дисперсия результатов взвешивания, следовательно, определяется в виде (для 1-го предмета)

$$D(A) = D(y_1 - y_0) = 2Dy,$$

где  $Dy$  - ошибка взвешивания.

Аналогично  $D(B) = D(C) = 2Dy$ , и не зависит от числа взвешиваемых предметов.

Запишем модель (5.6) в кодированных переменных (что нужно для сравнения этого способа взвешивания с другим - многофакторным):

$$x_i = \frac{x_i - 1/2}{1/2} = \begin{cases} +1 & \text{при взвешивании } i\text{-го предмета} \\ -1 & \text{в противном случае} \end{cases}.$$

Имеем

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3, \quad (5.7)$$

причем легко видеть, что  $\alpha_i = 2\beta_i$ ,  $\alpha_0 = \beta_0 - 2\sum\beta_i$

Обратим внимание на вид матрицы независимых переменных в рассматриваемом случае:

$$F_1 = \begin{bmatrix} 1 & -1 & -1 & -1 \\ 1 & +1 & -1 & -1 \\ 1 & -1 & +1 & -1 \\ 1 & -1 & -1 & +1 \end{bmatrix} \quad (5.8.)$$

Первый её столбец не ортогонален остальным, т.е.  $F_1^T F_1 \neq D$ , где  $D$  - диагональная матрица.



Рассмотрим другую схему взвешивания, учитывающую влияние нескольких факторов одновременно (схема *многофакторного* эксперимента), и представленную табл.5.2.

Таблица 5.2

Номер опыта	A	B	C	Результат взвешивания
1	+1	-1	-1	$y_0$
2	-1	+1	-1	$y_1$
3	-1	-1	+1	$y_2$
4	+1	+1	+1	$y_3$

Матрица независимых переменных для модели (6.7) в этом случае имеет вид

$$F_2 = \begin{bmatrix} 1 & +1 & -1 & -1 \\ 1 & -1 & +1 & -1 \\ 1 & -1 & -1 & +1 \\ 1 & +1 & +1 & +1 \end{bmatrix} \quad (5.9)$$

т.е.  $F_2^T F_2 = D$ , так как все её столбцы ортогональны.

Оценки масс A, B и C и их точность найдем по общим формулам, полученным ранее для ортогональных матриц F:

$$b_1 = \frac{\sum_{i=1}^N f_1(x_i) y_i}{N} = \frac{1}{4}(y_1 - y_2 - y_3 + y_4) \quad \text{или} \quad \alpha_1 = \frac{1}{2}(y_1 - y_2 - y_3 + y_4)$$

$$b_2 = \frac{\sum_{i=1}^N f_2(x_i) y_i}{N} = \frac{1}{4}(-y_1 + y_2 - y_3 + y_4) \quad \text{или} \quad \alpha_2 = \frac{1}{2}(-y_1 + y_2 - y_3 + y_4)$$

$$b_3 = \frac{\sum_{i=1}^N f_3(x_i) y_i}{N} = \frac{1}{4}(-y_1 - y_2 + y_3 + y_4) \quad \text{или} \quad \alpha_3 = \frac{1}{2}(-y_1 - y_2 + y_3 + y_4)$$

Учитывая независимость измерений  $y_i$ , имеем

$D\alpha_1 = DA = \frac{4}{4} Dy = Dy$ ; аналогично  $D\alpha_2 = Db = Dy$ ,  $D\alpha_3 = DC = Dy$ , т.е. дисперсия, связанная с ошибкой взвешивания, подучается вдвое меньше, чем при традиционном способе взвешивания, хотя в обоих случаях выполнялось по 4 опыта.

За счет чего получен выигрыш в точности? В первом случае эксперимент был поставлен так, что каждая масса была определена по результатам 2 опытов, во втором случае она вычислялась уже по результатам всех 4 опытов. Отсюда и удвоение точности.

Можно также сказать, что более высокая эффективность многофакторного эксперимента связана с более рациональным размещением точек эксперимента  $x_i$  в пространстве факторов. Это можно пояснить следующим образом. Столбцы матрицы  $F_1$  и  $F_2$  (кроме первого) задают координаты точек в трехмерном пространстве факторов  $X = R_3$ , которые расположены в вершинах куба с гранями, проходящими через точки  $\pm 1$ . Легко видеть, что матрица  $F_1$  задает 4 смежные вершины куба (на рис. 5.1 они отмечены крестиками), которые заполняют лишь небольшую часть всего куба, т.е. возможной области изменения факторов  $x_1, x_2, x_3$ .

Матрица  $F_2$ , напротив, задает 4 вершины куба, максимально разнесенные в пространстве  $X = R_3$ . Они образуют вершины правильного тетраэдра, вписанного в куб (на рис. 5.1 эти точки отмечены кружками).

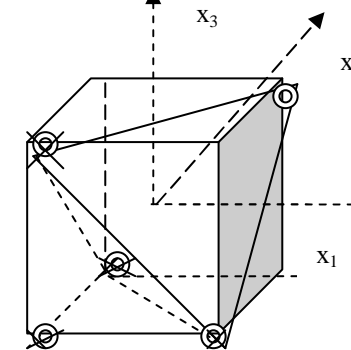


Рис. 5.1. Расположение точек в пространстве факторов при однофакторном эксперименте (отмечены крестиками) и при многофакторном (отмечены кружками)

Матрица  $F_2$ , напротив, задает 4 вершины куба, максимально разнесенные в пространстве  $X = R_3$ . Они образуют вершины правильного тетраэдра, вписанного в куб (на рис. 5.1 эти точки отмечены кружками).

Таким образом, радиус обследуемой сферы факторного пространства при многофакторном эксперименте существенно больше (как нетрудно подсчитать,  $r = \sqrt{3}$ ), это обстоятельство играет *решающее значение* для моделей типа (5.7), линейных по факторам. Чем дальше мы разносим точки  $x_i$  в пространстве факторов X, тем точнее определяем вектор параметров модели  $\beta$ .

Пусть, например, все координаты точки  $x_i$  увеличены в  $k$  ( $k > 1$ ) раз, т.е. наблюдения проводятся в точках  $x_i^* = kx_i$ . Тогда матрица ковариаций  $D(b)$  вектора оценок преобразуется следующим образом:

$$D(b) = \sigma^2 [F^T(kx_i)F(kx_i)]^{-1} = \frac{\sigma^2}{k^2} (F^T F)^{-1} \rightarrow 0 \text{ при } k \rightarrow \infty$$

Таким образом, модель вида (5.7) будет тем точнее, чем дальше разнесены наблюдения  $x_i$ .

Для одного фактора этот вывод иллюстрируется рис. 5.2.

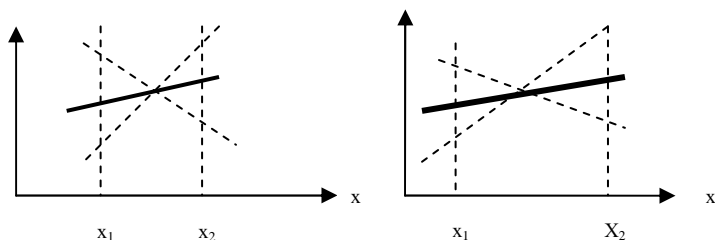


Рис. 5.2. Зависимость ошибки в оценке коэффициента  $b_1$  в модели  $f(x) = \beta_0 + \beta_1 x$  от расстояния между точками (черные кружки - истинные значения, светлые кружки - значения, измеренные с ошибками)

Однако область изменения факторов Q в любом реальном эксперименте ограничена. Поэтому мы не можем сколь угодно увеличивать радиус обследуемой сферы пространства факторов. Тем не менее, оказывается, что, используя многофакторные планы, мы увеличиваем радиус обследуемой сферы просто за счет свойств многомерного пространства, не увеличивая при этом интервалов варьирования по каждой переменной. Если каждая из переменных  $x_1, x_2, x_3$  варьируется на двух уровнях (т.е. принимает значения  $\pm 1$ ), то объем обследуемого пространства ограничен кубом, координаты вершин которого задаются перестановкой чисел  $(\pm 1, \pm 1, \pm 1)$ .

Очевидно, радиус сферы, в которую вписан трехмерный куб,  $r = \sqrt{3}$ . Переходя к пространству 4 факторов, нетрудно убедиться, что радиус сферы, в которую вписан четырехмерный куб с вершинами  $(\pm 1, \pm 1, \pm 1, \pm 1)$ ,  $r = \sqrt{4}$ ; в пространстве 5 факторов радиус аналогичной сферы  $\sqrt{5}$  и т.д.

Мы видим, что с ростом числа независимых переменных растет радиус обследуемой сферы при неизменных интервалах варьирования до каждой переменной!

### 5.3 Критерии оптимальности планов

Мы уже отмечали выше, что наилучший выбор матрицы плана  $\Pi$  при проведении активного эксперимента можно осуществить различным образом в зависимости от целей эксперимента. Иными словами, критерии оптимальности плана  $\Pi$  могут быть различными.

Различают три группы критериев оптимальности плана  $\Pi$ :

- 1) критерии, связанные со свойствами оценок коэффициентов модели  $b_k$  (которые определяются матрицей ковариаций  $D(b_k)$ );
- 2) критерии, связанные с предсказательными свойствами модели, т.е. с дисперсией отклика  $Dy^*(x) = d(x)$ , которая зависит от матрицы  $C = (F^T F)^{-1}$ ;
- 3) критерии, формулируемые без использования матрицы плана  $\Pi$  дисперсионной матрицы  $C = (F^T F)^{-1}$ .

Прежде чем формулировать критерии оптимальности, поясним геометрический смысл матрицы ковариаций  $D(b_k)$  и дисперсии отклика  $Dy^*(x)$ .

*Геометрический смысл  $D(b_k)$  (для 2 неизвестных параметров).* В пространстве параметров  $(\beta_1, \beta_2)$  вектор их оценок  $b = (b_1, b_2)$  является двумерной нормально распределенной случайной величиной с матрицей ковариаций  $D(b_k)$  и средним значением  $\beta = (\beta_1, \beta_2)$ . Возможные значения отклонений  $b_1 - \beta_1$  и  $b_2 - \beta_2$  характеризуются эллипсом рассеяния оценок (рис. 5.3). Площадь этого эллипса  $S = 2\pi \sqrt{D(b_1)} \sqrt{D(b_2)}$  пропорциональна  $\det D(b)$  - определителю матрицы  $D(b)$ , который называют *обобщенной дисперсией оценки*  $b$ .

В пространстве параметров с числом измерений  $m > 2$  вместо эллипса рассматривают эллипсоид рассеяния, добавив еще, что длины

осей эллипсоида рассеяния определяются средними квадратическими отклонениями  $\sqrt{Db_k}$  оценок  $b_k$  соответствующих параметров модели.

Таким образом, определитель матрицы  $Db$  (обобщенная дисперсия) является обобщенной мерой точности вектора оценок  $b$  и геометрически пропорционален объему эллипсоида рассеяния этого вектора.

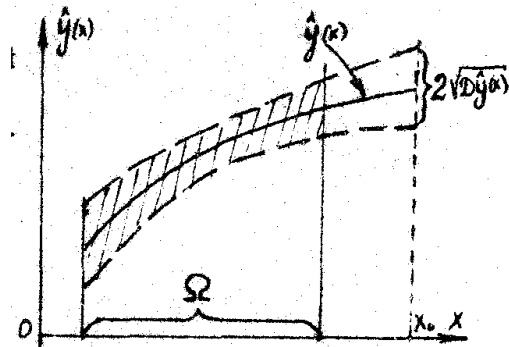


Рис. 5.3. Эллипс рассеяния вектора оценок

*Геометрическая интерпретация дисперсии отклика.* В пространстве факторов  $X$  величина  $Dy^*(x)$  характеризует точность определения отклика  $y(x)$  в произвольной точке  $x \in X$ .

Случай одного фактора иллюстрирует рис. 5.4.

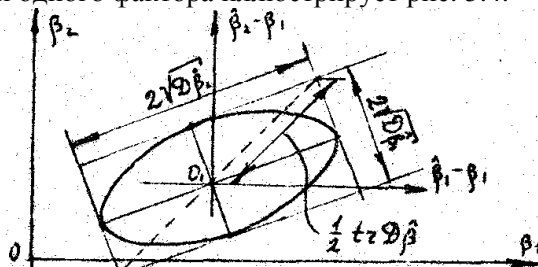


Рис. 5.4. График модели  $y^*(x)=f(x,b)$  (сплошная кривая) и коридор возможных значений  $y^*(x)$  (пунктирные линии). Заштрихованная часть относится к области  $\Omega$  значений  $x$ , где проводился эксперимент.

Таблица 5.3

	Критерий	Алгебраический смысл	Статистический смысл и геометрическая интерпретация
1-я группа	D-оптимальность	$\min \det Db_k = \min \det C(\Pi)$ $\Pi \quad \Pi$	Минимум обобщенной дисперсии. Минимум объема эллипсоида рассеяния оценок параметров
	A-оптимальность	$\min \max \text{tr } C(\Pi)$ $\Pi$ $\text{tr } C(\Pi)$ - сумма диагональных элементов матрицы $C(\Pi)$	Минимальная средняя дисперсия оценок коэффициентов. Минимум суммы квадратов осей эллипсоида рассеяния или минимума длины диагонали прямоугольника, описанного около эллипсоида
	E-оптимальность	$\min \max \sqrt{Db_k}$ $\Pi \quad k$	Отдельные оценки параметров не обладают слишком большими дисперсиями. Минимум максимальной оси эллипсоида рассеяния
	Ортогональность	$C(\Pi) = \frac{1}{N} E$ (E-единичная матрица)	Оценки параметров независимы, т.е. все направления осей эллипсоида рассеяния параллельны координатным осям пространства параметров
2-я группа	G-оптимальность	$\min \max Dy^*(x_i)$ $\Pi \quad i$	Минимум максимального значения дисперсии оценок отклика
	Q-оптимальность	$\min \sum_{i=1}^n Dy^*(x_i)$ $\Pi$	Минимум средней дисперсии отклика
	Ротатабельность	$Dy^*(x_i) = f(\rho)$ $\rho = \sqrt{\sum_{i=1}^n (x_i)^2}$ Расстояние от точки $x = (x_1, x_2, \dots, x_m)$ до центра эксперимента	Постоянство дисперсии предсказания на равных расстояниях от центра эксперимента
3-я группа	Насыщенность	_____	Степень близости общего числа опытов к числу неизвестных параметров модели
	Композиционность	_____	Свойство плана $\Pi$ , означающее возможность разделить эксперимент на части и реализовать каждую часть последовательно

Таким образом, *двухсигмовый интервал* возможных значений  $y^*(x_0)$  в точке  $x_0$  имеет вид  $y^*(x_0) \pm \sqrt{Dy^*(x)}$ .

Геометрическая интерпретация величин  $Dy$  и  $Dy^*(x)$  позволяет дать наглядный смысл каждому из критериев оптимальности плана. Сводка критериев оптимальности и их интерпретация приведены в табл. 5.3.

Список возможных критериев далеко не исчерпывается табл. 5.3. Другие критерии отмечены в [3]. Но даже если ограничиться приведенными критериями, то возникает вопрос, как поступать «простому экспериментатору», какой из критериев выбирать?

К сожалению, общей теории выбора критериев оптимальности пока не существует [1, 3]. Поэтому при решении конкретных задач приходится руководствоваться ограничениями и возможностями конкретной ситуации. Очень часто удобно использовать ортогональные планы, позволяющие значительно упростить числовые расчеты, особенно при поиске адекватной модели. Наиболее общий, казалось бы, критерий  $D$  - оптимальности требует, как правило, слишком большого числа опытов и довольно громоздких расчетов, связанных с применением ЦВМ. Поэтому его использование ограничивается, обычно, теми случаями, когда возможно взять готовый план из каталога планов [4].

Положение, однако, существенно упрощается для важного на практике класса моделей - полиномиальных линейных моделей 1-го порядка. Общий вид такой модели для  $m$  факторов дает выражение

$$y(x_1, x_2, \dots, x_m) = \beta_0 + \sum_{k=1}^m \beta_k x_k \quad (5.10)$$

где  $x_1, x_2, \dots, x_m$  - независимые переменные (кодированные).

Оказывается, что для модели (1.10) ортогональный план (при котором  $F^T(\Pi) F(\Pi) = NE$ , где  $E$  - единичная матрица) является оптимальным и по всем другим критериям [1,5].

Таким образом, планирование эксперимента при использовании линейных моделей 1-го порядка сводится к изучению ортогональных планов. Этот вопрос, а также вопросы планирования эксперимента при построении более общих моделей, чем (5.10), будет рассмотрен ниже.

## 5.4 ПЛАНИРОВАНИЕ ЭКСПЕРИМЕНТА ПРИ ПОСТРОЕНИИ МОДЕЛИ

Данная глава будет посвящена вопросам планирования эксперимента применительно к полиномиальным моделям 1-го и 2-го порядков.

### 5.4.1 Линейные модели 1-го порядка.

#### Полный Факторный эксперимент типа $2^m$

При обсуждении критериев оптимальности, которыми руководствуются при планировании эксперимента (гл.5.2), отмечалось, что ортогональное планирование играет особую роль для моделей 1-го порядка, так как для них ортогональные планы являются одновременно оптимальными по всем остальным критериям. Поэтому в данном параграфе будут подробно рассмотрены вопросы построения ортогональных планов и их возможности применительно к моделям 1-го порядка.

В основе всех ортогональных планов лежит так называемый *полный факторный эксперимент (ПФЭ) типа  $2^m$* , при котором каждый из факторов в эксперименте может принимать одно из двух возможных значений или, как говорят, устанавливаться на одном из двух возможных уровней. В ПФЭ уровни каждого фактора комбинируются со всеми уровнями других факторов. Число экспериментов и при этом растет как  $2^m$ , т.е. достаточно быстро. Поэтому при больших значениях  $m$  ( $m > 3$ ) важное практическое значение имеют неполные (или, как говорят, дробные) факторные эксперименты, сохраняющие свойство ортогональности.

Прежде, чем переходить к рассмотрению ПФЭ типа  $2^m$ , перечислим те вопросы, которые, непосредственно предшествуют выбору матрицы плана и составляют содержание так называемого этапа «предпланирования эксперимента».

*Этап предпланирования эксперимента.* На этом этапе, непосредственно предшествующем выбору матрицы плана и проведению эксперимента, решаются следующие основные вопросы:

1. Выбор факторов  $z_k$  ( $k=1, 2, \dots, m$ ), подлежащих изучению, и выходной переменной (отклика)  $y$ .
2. Выбор области проведения эксперимента  $\Omega$  в факторном пространстве  $Z$ .
3. Выбор базовой точки  $z_0 = (z_{01}, z_{02}, \dots, z_{0m}) \in \Omega$  - центра эксперимента.

4. Выбор интервалов варьирования  $\Delta z$  ( $k=1, 2, \dots, m$ ) по каждому фактору  $z_k$  (предполагается, что каждый фактор  $z_k$  будет принимать два уровня  $z_{k0} + \Delta z_k$  и  $z_{k0} - \Delta z_k$ ).

5. Кодирование переменных.

Рассмотрим подробнее каждый перечисленных пунктов.

*Выбор факторов  $z$*  ( $k=1, 2, \dots, m$ ), которые включаются в модель, производится с учетом следующих обстоятельств.

*Во-первых*, факторы должны быть управляемыми, так как в противном случае невозможно производить эксперимент в соответствии с заданным планом. Например, такой фактор, как доля примеси в сырье при производстве некоторого продукта не является управляемым, его следует отнести к шумовому фону.

*Во-вторых*, факторы должны быть функционально независимыми. Например, рассматривая такие переменные, как ток, напряжение и активную мощность источника питания, в качестве факторов можно взять любые два из трех (например, напряжение и ток), так как все три переменные связаны функциональной зависимостью.

*В-третьих*, факторы должны быть совместимыми, т.е. комбинации их уровней, предусмотренных планом, должны быть безопасными и не давать бракованной продукции.

*Наконец* предполагается, что в рассмотрение включены все существенные факторы. Решение этого вопроса находится в компетенции специалиста, располагающего априорной информацией о свойствах исследуемого объекта. Существуют специальные методы выделения существенных факторов и отсеивания несущественных [2].

Выбор выходной переменной (или переменных) определяется смыслом и целевой установкой решаемой задачи. При этом следует иметь в виду, что выходная переменная должна отражать свойства и качества процесса в целом и должна измеряться достаточно точно.

*Область проведения эксперимента  $\Omega$*  выбирается путем компромисса между желанием охватить возможно более широкую часть пространства факторов  $Z$  и стремлением получить простую и адекватную модель объекта, что возможно, как правило, в небольшой части области  $Z$ . При этом, разумеется, должны учитываться те ограничения на факторы  $z_k$ , которые определяются реальными условиями и смыслом задачи (например, время  $t \geq 0$ , температура  $T$  не выше некоторого значения  $T_{\max}$  и т.д.).

*Базовая точка  $z_0 = (z_{01}, z_{02}, \dots, z_{0m})$*  выбирается либо в центре области ограничений  $\Omega$ , если модель объекта строится сразу для всей области  $\Omega$ , либо в центре каждой части  $\Omega_0$ , на которые  $\Omega$  разбивается с целью поиска в каждой части наиболее простой модели объекта. В этом случае, когда решается задача поиска оптимальных условий, в качестве базовой точки выбирается такая, в которой целевая функция принимает наилучшее значение по сравнению со всеми другими точками, изучаемыми в процессе предварительных исследований.

*Интервалы варьирования  $z_k$ ,  $k=1, 2, \dots, m$*  должны отвечать следующим условиям:

-  $\Delta z_k$  должна существенно превышать, погрешность измерения по соответствующему фактору  $x_k$ ;

-  $\Delta z_k$  не должна превышать размеров области допустимых значений  $\Omega$ , и не должна быть настолько большой, чтобы математическое описание поверхности отклика было слишком грубым;

- приращение  $\Delta \eta$  выходной переменной  $\eta$ , вызванное изменением управляемого фактора  $z_k$  на  $\Delta z_k$  должно существенно превышать погрешность измерения выхода  $\Delta \eta$  (чтобы было возможным выделение  $\Delta \eta$  на фоне шумов при небольшом числе параллельных опытов).

Ясно, что эти условия могут служить лишь некоторым ориентиром и не дает формализованного правила выбора  $\Delta z x_k$ .

После выбора базовой точки  $z_0 = (z_{01}, z_{02}, \dots, z_{0m})$  и интервалов варьирования  $\Delta z_k$  ( $k=1, 2, \dots, m$ ) производится кодирование переменных  $z$  и переход к переменным  $x_k$  по формулам кодирования

$$x_k = \frac{z_k - z_{0k}}{\Delta z_k}, \quad k=1, 2, \dots, m \quad (5.11)$$

Поскольку каждая переменная  $z_k$  принимает в каждом опыте лишь два уровня нижний  $z_{k0} - \Delta z_k$  верхний  $z_{k0} + \Delta z_k$ , то кодированная переменная  $x_k$

( $k=1, 2, \dots, m$ ) в каждом опыте может принимать значения либо +1 (верхний уровень), либо -1 (нижний уровень).

Преобразование переменных (5.11) есть линейное преобразование, которое переводит начало координат исходного пространства  $Z$  в точку  $z_0$ , а прямоугольную область  $\Omega_x = \{|x_{k0} - x_k| < \Delta x_k, (k=1, 2, \dots, m)\}$  в  $m$ -мерный гиперкуб  $\Omega_x$ , грани которого отсекают на осях отрезки  $\pm 1$ , т.е.

$$\Omega_x = \{-1 \leq x_k \leq +1, (k=1, 2, \dots, m)\}.$$

Оказывается, что если точки эксперимента  $x_k=(x_{k1}, x_{k2}, \dots, x_{km})$  выбирать в вершинах гиперкуба  $\Omega_x$  определённым образом (как именно - указано ниже), то план  $\Pi$  будет ортогональным, т.е. информационная матрица  $M=(F^T F)$  будет диагональной:  $M = F^T F = NE$ , где

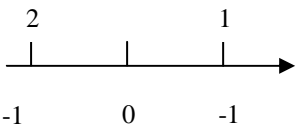
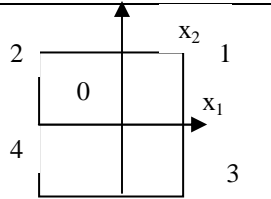
$$E = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \text{единичная матрица.}$$

*Полный факторный эксперимент типа  $2^m$ .* Один из способов построения ортогонального плана  $\Pi$  - это выбор в качестве точки эксперимента  $x_k$  в факторном пространстве всех вершин  $m$ -мерного куба  $\Omega_x$ , т.е. построение ПФЭ  $2^m$ .

*Определение.* Полным факторным экспериментом типа  $2^m$  называется множество всех точек в  $m$ -мерном пространстве кодированных переменных, координаты которых равны либо +1, либо -1.

Число точек в этом плане  $n=2^m$ , и они заполняют все вершины  $m$ -мерного куба. Следовательно, и минимальное число опытов при проведении ПФЭ  $2^m$  равно  $n=2^m$ . Их больше  $n$ , если в некоторых точках проводятся параллельные опыты.

*Примеры ПФЭ  $2^m$  и общее правило построения плана*

Размерность пространства	Точки эксперимента	План эксперимента
1) $m=1$		$\Pi_1 = \begin{bmatrix} +1 \\ -1 \end{bmatrix} = \begin{bmatrix} + \\ - \end{bmatrix}$
2) $m=1$		$\Pi_2 = \begin{bmatrix} +1 & +1 \\ -1 & +1 \\ +1 & -1 \\ -1 & -1 \end{bmatrix} = \begin{bmatrix} + & + \\ - & + \\ + & - \\ - & - \end{bmatrix}$

Нумерация точек плана может осуществляться по-разному. Мы будем следовать правилу «наращивания», которое применительно к плану  $\Pi_2$  и состоит в следующем:

$$\Pi_2 = \begin{bmatrix} \Pi_1 & l_1 \\ \Pi_2 & l_2 \end{bmatrix}, \text{ где } l_j = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} + \\ + \end{bmatrix}$$

$$3). m=3 \quad \Pi_3 = \begin{bmatrix} \Pi_2 & l_2 \\ \Pi_2 & -l_2 \end{bmatrix}, \text{ где } l_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Вообще, если построен уже план  $\Pi_m$ , то план  $\Pi_{m+1}$  (определяется по рекуррентному правилу (правилу «наращивания»):

$$\Pi_{m+1} = \begin{bmatrix} \Pi_m & l_m \\ \Pi_m & -l_m \end{bmatrix}, \text{ где } l_m = \begin{bmatrix} 1 \\ \dots \\ \dots \\ 1 \end{bmatrix}$$

Убедимся в том, что ПФЭ типа  $2^m$ , задаваемый матрицей плана  $\Pi_m$ , является ортогональным.

Пусть вначале  $m = 3$ , и пусть функция отклика имеет вид

$$\eta = \sum_{k=0}^3 \beta_k x_k + \sum_{k < j}^3 \beta_{kj} x_k x_j + \beta_{123} x_1 x_2 x_3 \quad (5.12)$$

т.е. является полиномиальной моделью 1-го порядка со всеми взаимодействиями: парными  $x_j x_k$   $k \neq j$  (или взаимодействиями 1-го порядка) и тройным  $x_1 x_2 x_3$  (взаимодействием 2-го порядка). Попутно мы покажем, что для такой функции отклика могут быть получены МНК-оценки всех неизвестных коэффициентов  $\{\beta_k\}$ ,  $\{\beta_{kj}\}$ ,  $\beta_{123}$ .

Обозначим через  $x_{ku}$  значение переменной  $x_k$  ( $k=1,2,3$ ) в  $u$ -м опыте

( $u = 1, 2, \dots, 2^m = 8$ ) (все  $x_{ku} = \pm 1$ ) и выпишем матрицу независимых переменных, учитывая, что базисные функции в рассматриваемом случае равны:

$$f_0(x) = 1, f_1(x) = x_1, f_2(x) = x_2, f_3(x) = x_3, f_4(x) = x_1 x_2 = x_4, f_5(x) = x_1 x_3 = x_5,$$

$$f_6(x) = x_2 x_3 = x_6,$$

$f_7(x) = x_1 x_2 x_3 = x_7$ , где переменные  $x_4, \dots, x_7$  - введены для удобства обозначений.

Матрица независимых переменных  $F=[f_k(x_u)]=[x_{ku}]$ ,  $k=0,1,\dots,7$ ;  $u=1,2,\dots,8$ . Или в развернутом виде

$$F = \begin{bmatrix} x_0 & x_1 & x_2 & x_3 & x_4 = x_1x_2 & x_5 = x_1x_3 & x_6 = x_2x_3 & x_7 = x_1x_2x_3 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 & -1 & -1 & 1 & -1 \\ 1 & 1 & -1 & 1 & -1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & -1 & 1 & -1 & -1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 \end{bmatrix}$$

Из последнего представления  $F$  видно, что 2-й, 3-й и 4-й столбцы образуют матрицу плана ПФЭ  $2^3$ , а последующие столбцы матрицы (с 5-го по 8-й) для парных и тройных взаимодействий получаются перемножением соответствующих столбцов матрицы плана  $2^3$ .

Для матрицы независимых переменных  $\Gamma$  выполняются следующие соотношения:

$$\sum_{u=1}^n x_{ku} = 0 \quad (k=1,2,\dots,7) \quad (5.13)$$

$$\sum_{u=1}^n (x_{ku})^2 = n \quad (k=0,1,2,\dots,7) \quad (5.14)$$

$$\sum_{u=1}^n x_{kj}x_{ku} = 0 \quad (k=1,2,\dots,7, k \neq j) \quad (5.15)$$

Соотношение (5.13) является условием *симметричности* плана, оно означает, что точки плана расположены симметрично относительно начала координат (в пространстве кодированных переменных). Соотношение (5.14) называют условием *нормировки*, оно означает, что в ПФЭ предусмотрены лишь 2 уровня +1 и -1. Соотношение (5.15) означает, что столбцы матрицы  $F$  попарно *ортогональны* и, следовательно, матрица  $(F^T F) = M$  называемая *информационной матрицей Фишера*, является диагональной:  $M = (F^T F) = nE_n$ ,

$$\text{где } E_n = \begin{bmatrix} 1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & 0 \\ 0 & 0 & \dots & 1 \end{bmatrix} \quad [n \times n] \text{- единичная матрица.}$$

Тем самым мы убедились в том, что план  $P_3$  для ПФЭ  $2^m$  ( $m = 3$ ) является ортогональным. Ясно, что свойство сохраняется, если функция отклика (5.12) не содержит некоторые или даже все взаимодействия.

Число неизвестных коэффициентов в модели (5.12) равно  $l+1=1+m+C_m^2+C_m^3=8$  при  $m=3$ , а число опытов  $n = 2^m=8$  тоже, т.е. все коэффициенты  $\beta_0, \beta_1, \beta_2, \beta_3, \beta_{12} = \beta_4, \beta_{13} = \beta_5, \beta_{23} = \beta_6, \beta_{123} = \beta_7$ , могут быть оценены независимо. В соответствии с общими формулами для ортогонального плана (см. ч.1, гл.,3) имеем следующие МНК-оценки коэффициентов  $\beta_k$  и их дисперсии:

$$b_k = \frac{1}{N} \sum_{u=1}^n \varphi_k(x_u) y_u, \text{ или } b_k = \frac{1}{N} \sum_{u=1}^n x_{ku} y_u \quad (5.16)$$

$$Db_k = \frac{\sigma^2}{n} \quad (2.7) \quad k=0,1,2,\dots,7$$

т.е. оценки  $\{b_k\}$ ,  $k=0,1,2,\dots,7$  некоррелированы.

Если заранее известно, что некоторые (или все) взаимодействия в модели (5.12) отсутствуют, то нетрудно видеть, что формулы для вычисления МНК-оценок остальных коэффициентов при этом сохраняются прежними.

Ясно, что приведенные рассуждения и выводы остаются справедливыми для ПФЭ  $2^m$  при любом  $m > 3$ ), если только функция отклика представляет собой полиномиальную модель 1-го порядка, включающую, быть может, все взаимодействия переменных  $x_1, x_2, \dots, x_m$ . Действительно, пусть функция отклика имеет вид

$$\eta = \beta_0 + \sum_{k=1}^m \beta_k x_k + \sum_{k < j} \beta_{kj} x_k x_j + \sum_{k < i < j} \beta_{kij} x_k x_i x_j \quad (5.18)$$

где произведение  $x_1 x_2 \dots x_m$  называется взаимодействием  $(m-1)$ -го порядка факторов  $x_1, x_2, \dots, x_m$ .

Тогда общее число неизвестных коэффициентов в модели (5.15)

$$l+1=1+m+C_m^2+C_m^3+\dots+C_m^m=\sum_{k=0}^m C_m^k=2^m$$

Но число опытов в ПФЭ  $2^m$  равно  $n=2^m$  тоже. Таким образом, ПФЭ  $2^m$  позволяет независимо оценить все коэффициенты в модели (5.18).

Нетрудно также проверить, что матрица независимых переменных состоит из попарно ортогональных столбцов, а, следовательно, матрица

$\{F^T F\} = M$  является диагональной. Это значит, что план  $\Pi_m$  в ПФЭ  $2^m$  является ортогональным и, следовательно, МНК-оценки и их дисперсии для коэффициентов модели (5.15)

$$\beta_0, \beta_1, \beta_2, \dots, \beta_m, \beta_{12}=\beta_{m+1}, \dots, \beta_{1\dots m} = \beta_l = 2^m - 1$$

определяются по формулам, аналогичным (4.22) и (4.30) из раздела 4:

$$b_n = \frac{1}{n} \sum_{u=1}^n x_{uk} x_{uk} \quad k=0, 1, 2, \dots, l \quad (5.19)$$

$$D(b_k) = \frac{\sigma^2}{m}, \quad k=0, 1, \dots, l, \quad \text{cov}(b_i b_j), i \neq j \quad (5.20)$$

где, напомним,  $x_{ku}$  - значение переменной  $x_k$  ( $k=0, 1, \dots, l$ ) в  $u$ -м опыте ( $u=0, 1, \dots, m$ ), причем  $x_0 = 1$ ,

$$x_{m+1} = x_1 x_2, \quad x_{m+2} = x_1 x_3, \quad x_l = x_1 x_2 \dots x_m$$

Заметим, что выражение вида  $b_k x_k$  называют часто линейным эффектом переменной а  $b_{1\dots d} x_1 x_2 \dots x_d$  ( $d=2, 3, \dots, m$ )-эффектом взаимодействия факторов  $x_1 x_2 \dots x_d$

### 5.4.2 Линейные модели 2-го порядка

Если изучение отклика у производится с целью отыскания его экстремальных значений, модели 1-го порядка перестают работать - нужны модели 2-го порядка вида (4.9), где присутствуют лишь парные взаимодействия. К таким моделям приходится переходить также в том случае, когда проверка адекватности показала плохое согласие с моделью 1-го порядка:

$$\eta = \beta_0 + \sum_{k=1}^m \beta_k x_k + \sum_{i=m+1}^{2m} \beta_{ii} x_i^2 + \beta_{2m+1} x_1 x_2 + \beta_l x_{m+1} x_m \quad (5.21)$$

Общее число коэффициентов модели  $l^* = 1 + m + m + C_{m+2}^2$ .

Для их оценки ПФЭ типа  $2^m$  - на двух уровнях - уже не годится, так как значений у на границе области проведения эксперимента  $\Omega$  недостаточно для определения характера поверхности отклика. Поэтому для модели 2-го порядка нужно проводить эксперимент на 3 уровнях. Но ПФЭ типа  $3^m$  требует слишком много наблюдений (по сравнению с числом неизвестных параметров) и, кроме того, не является ортогональным. Степень избыточности такого плана иллюстрирует табл. 5.4.

Таблица 5.4.

m	1	2	3	4	5
$C_{m+2}^2 = l^*$	3	6	10	15	21
$3^m$	3	9	27	81	243

Поэтому для квадратичных моделей распределения применяют так называемые *композиционные планы*, предложенные Боксом. Хотя они и не являются D - оптимальными, но не слишком сильно проигрывая им в эффективности (измеряемой обобщенной дисперсией  $\det D[b_k]$  они существенно более экономичны в числе наблюдений.

*Композиционные планы Бокса.* Композиционный план, как это следует из названия, скомпонован из «ядра», которым является план (полный либо дробный факторный) модели 1-го порядка, и дополнительных точек: центральной  $(0, 0, 0, \dots, 0)$ , в которой проводится  $n_0$  измерений, и  $2m$  так называемых «звездных» точек с координатами  $(\pm\alpha, 0, \dots, 0)$ ,  $(0, \pm\alpha, 0, \dots, 0)$ ,  $(0, 0, \dots, \pm\alpha)$ , расположенных на расстоянии  $\alpha$  от середины ребер гиперкуба.

Выбором длины «плеча»  $\alpha$  и числа  $n$  композиционному плану могут быть приданы различные свойства, например, ортогональность или рототабальность.

Рассмотрим ортогональные центральные композиционные планы (ОЦКП).

Для обеспечения ортогональности достаточно положить  $n_0=1$  и преобразовать модель 2-го порядка (5.21) следующим образом:

$$\eta = b_0 + \sum_{k=1}^m b_k x_k + \sum_{i=m+1}^{2m} b_{ii} (x_i^2 - \theta) + b_{2m+1} x_1 x_2 + b_l x_{m+1} x_m \quad (5.22)$$

$$\text{где} \quad \theta = \frac{1}{N} \sum_{i=1}^N (x_{ij})^2 = \frac{2^{n-p} + \alpha^2}{N} \quad (5.23)$$



В равенстве (5.23) N- общее число точек в плане;  
 $2^{n-p}$ - число точек «ядра» композиционного плана, являющегося  
 ДФП типа  $2^{n-p}$ . Ясно, что (с учетом  $n_0 = 1$ )  $N = 2^{n-p} + 2m + 1$ .

От модели (5.23) легко перейти к модели (5.22), учитывая, что

$$\beta_0 = b_0 - \theta \sum_{i=m+1}^{2m} b_i \quad (5.24)$$

Таблица 5.5

	Но- мер опы- та	$x_0$	Матрица F									Мат рица Y
			Матрица плана			$x_1$	$x_2$	$x_3$	$x_1 x_2$	$x_1 x_3$	$x_2 x_3$	
			$X_1$	$X_2$	$X_3$	$\theta$	$\theta$	$\theta$	$\theta$	$\theta$	$\theta$	
«Яд- ро» плана $2^m$	1	1	+1	+1	+1	$I-\theta$	$I-\theta$	$I-\theta$	+1	+1	+1	$y_1$
	2	1	-1	+1	+1	$I-\theta$	$I-\theta$	$I-\theta$	-1	-1	+1	$y_2$
	3	1	+1	-1	+1	$I-\theta$	$I-\theta$	$I-\theta$	-1	+1	-1	$y_3$
	4	1	-1	-1	+1	$I-\theta$	$I-\theta$	$I-\theta$	+1	-1	-1	$y_4$
	5	1	+1	+1	-1	$I-\theta$	$I-\theta$	$I-\theta$	+1	-1	-1	$y_5$
	6	1	-1	+1	-1	$I-\theta$	$I-\theta$	$I-\theta$	-1	+1	-1	$y_6$
	7	1	+1	-1	-1	$I-\theta$	$I-\theta$	$I-\theta$	-1	-1	+1	$y_7$
	8	1	-1	-1	-1	$I-\theta$	$I-\theta$	$I-\theta$	+1	+1	+1	$y_8$
«Звёзд- ные» точки	9	1	$\alpha$	0	0	$\alpha^2-\theta$	$-\theta$	$-\theta$	0	0	0	$y_9$
	10	1	$-\alpha$	0	0	$\alpha^2-\theta$	$-\theta$	$-\theta$	0	0	0	$y_{10}$
	11	1	0	$\alpha$	0	$-\theta$	$\alpha^2-\theta$	$-\theta$	0	0	0	$y_{11}$
	12	1	0	$-\alpha$	0	$-\theta$	$\alpha^2-\theta$	$-\theta$	0	0	0	$y_{12}$
	13	1	0	0	$\alpha$	$-\theta$	$-\theta$	$\alpha^2-\theta$	0	0	0	$y_{13}$
	14	1	0	0	$-\alpha$	$-\theta$	$-\theta$	$\alpha^2-\theta$	0	0	0	$y_{14}$
Центр	15	1	0	0	0	$-\theta$	$-\theta$	$-\theta$	0	0	0	$y_{15}$

Для оценки коэффициентов модели (5.22) достаточно знать матрицу независимых переменных F и матрицу отклика Y, представленные в табл. 5.5 для 3 факторов ( $m = 3$ ), где матрица ПФЭ  $2^m$  обведена рамкой.

Можно проверить, что для обеспечения ортогональности столбцов матрицы F плечо  $\alpha$  нужно определять из выражения

$$\alpha = \sqrt{2^{\frac{m-p}{2}} (\sqrt{N} - 2^{\frac{m-p}{2}})}$$

Значения  $\alpha$  для различных  $m$  приведены в табл. 5.5. В силу ортогональности столбцов матрицы F информационная матрица плана  $M = F^T F$  будет диагональной следующего вида:

$$M = \begin{bmatrix} m_0 & 0 & 0 & 0 \\ 0 & m_1 I_m & 0 & 0 \\ 0 & 0 & m_2 I_m & 0 \\ 0 & 0 & 0 & m_3 I_m \end{bmatrix},$$

где  $m_0 = N = 2^{(m-p)} + 2m + 1$ ;

$m_1 = 2^{(m-p)} + 2\alpha^2$ ;

$m_2 = 2^{(m-p)}(1-\theta)^2 + 2(\alpha^2 - \theta^2) + (2n-1)\theta^2$ ;

$m_3 = 2^{(m-p)}$ ;

I - единичная матрица ( $m \times m$ );

дисперсионная матрица  $C = M^{-1}$  получается из M обращением диагональных элементов  $m_i$ , т.е.

$$C = \begin{bmatrix} c_0 & 0 & 0 & 0 \\ 0 & c_1 I_m & 0 & 0 \\ 0 & 0 & c_2 I_m & 0 \\ 0 & 0 & 0 & c_3 I_m \end{bmatrix}, \text{ где } c_i = \frac{1}{m_i}, i=0, 1, 2, 3.$$

Используя общие формулы (4.16) и (4.22), а также учитывая (2.14), для коэффициентов исходной модели (5.21) можно записать следующие выражения:

$$b_0 = \frac{1}{N} \sum_{j=1}^N y_j - \theta \sum_{i=m+1}^{2m} b_i \quad (5.25).$$

$$b_i = \begin{cases} \frac{1}{m} \sum_{j=1}^N x_{ij} y_j, & \text{при } i = 1, 2, \dots, m \\ \frac{1}{m} \sum_{j=1}^N [(x_{ij})^2 - \theta] y_j, & \text{при } i = m+1, \dots, 2m \\ \frac{1}{m_3} \sum_{j=1}^N x_{rj} x_{kj} y_j, & \text{при } \dots r \neq k \quad i = 2m+1, \dots, l \end{cases}$$

$$S_i^2 = \begin{cases} S^2 c_0 & \text{при } i = 0 \\ S^2 c_1 & \text{при } i = 1, 2, \dots, m \\ S^2 c_2 & \text{при } i = m + 1, \dots, 2m \\ S^2 c_3 & \text{при } i = 2m + 1, \dots, l \end{cases} \quad S_{b_0}^2 = S^2(c_0 + n\theta^2 c_2),$$

$S^2$ -оценка дисперсия воспроизводимости.

Для удобства пользователей сведём значения  $\theta$ ,  $\alpha$ ,  $c_i$  в табл. 5.6.

Таблица 5.6

Размерность	Ядро плана	N	$\alpha$	$\theta$	Элементы матрицы			
					$c_0$	$c_1$	$c_2$	$c_3$
2	$2^2$	9	1	0,66	0,11	0,16	0,5	0,25
3	$2^3$	15	1,215	7	1	7	0,23	0,125
4	$2^4$	25	1,414	0,73	0,06	0,09	0,12	0,062
				0,8	7	0,05	5	
					0,04			
5	$2^{5-1}$	27	1,547	0,77	0,03	0,04	0,08	0,062
6	$2^{6-1}$	45	1,722	0,84	7	8	7	0,031
7	$2^{7-1}$	79	1,885	3	0,02	0,02	0,05	0,016
				0,9	2	6	6	
					0,01	0,01	0,03	
					3	4	9	

Заметим, что для числа факторов  $m \leq 4$  в качестве «ядра» ортогонального композиционного плана может быть использован только ПФЭ. Для  $m=5, 6, 7$  можно использовать (в качестве «ядра») полуреплики  $2^{m-1}$ , и лишь для  $m=8$  можно использовать четвертьреплику  $2^{m-2}$ . Это отражено в табл. 5.6.

**Пример 2.1.** Изучался процесс нормирования лакосажевых резисторов. В качестве выходной переменной  $y$  рассматривался процент годных изделий. Варьируемыми параметрами являлись: вязкость суспензии  $z_1$ , температура при нанесении суспензии  $z_2$ , скорость нанесения суспензии на подложку  $z_3$ , температура при термической обработке  $z_4$ . Для поиска оптимальных условий в окрестности режима,

принятого по технологии, был проведен эксперимент на основе ОЦКП 2-го порядка для получения модели в виде полинома 2-го порядка. Уровни варьирования факторов выбраны были следующими: по фактору

$$z_1: 7,0 \pm 0,5, \text{ т.е. } x_1 = \frac{z_1 - 7}{0,5};$$

$$z_2: 25,0 \pm 5, \text{ т.е. } x_2 = \frac{z_2 - 25}{5};$$

$$z_3: 4 \pm 2, \text{ т.е. } x_3 = \frac{z_3 - 4}{2};$$

$$z_4: 130 \pm 57, \text{ т.е. } x_4 = \frac{z_4 - 130}{57}.$$

В соответствии с ОЦКП эксперимент был проведен в  $N = 2^4 + 2 \times 4 + 1 = 25$  точках и дал следующие значения  $y$  в соответствующих точках (вид плана аналогичен табл. 5.5);

- 1) 40,7; 6) 43,04; 11) 50,07; 16) 50,05; 21) 44,84;
- 2) 30,6; 7) 49,9; 12) 49,53; 17) 49,9; 22) 50,1;
- 3) 44,6; 8) 47,3; 13) 50,4; 18) 43,9; 23) 43,74;
- 4) 37,6; 9) 49,5; 14) 50,15; 19) 49,48; 24) 50,41;
- 5) 49,2; 10) 46,8; 15) 50,33; 20) 50,12; 25) 49,97.

Из табл. 5.7 находим  $\alpha=1,414$  и  $\theta=0,8$ ,  $c_0=0,04$ ,  $c_1=0,05$ ,  $c_2=0,125$ ,  $c_3=0,062$ . На основе формул (5.21) и (5.22) получаем модель процесса в виде

$$y^*(x) = 49,6 - 3,9x_1 + 1,98x_2 + 4,83x_3 + 6,3x_4 - 2,6x_1^2 + 0,3x_2^2 - 2,03x_3^2 - 2,43x_4^2 + 1,09x_1x_2 + 1,36x_1x_3 + 2,75x_1x_4 - 1,17x_2x_3 - 1,61x_2x_4 - 3,8x_3x_4$$

Проверим адекватность модели. Оценка дисперсии воспроизводимости отклика  $S^{2*} = 1,19$  с числом степеней свобода  $\nu_2 = 3$  получена путем параллельных наблюдений, проведенных в некоторой точке вне

рамок плана. Остаточная сумма квадратов  $\sum_{i=1}^N (y_i - y_i^*)^2 = 52,2$  при

числе степеней свободы  $\nu_1$ , равном разности числа наблюдений и числа  $l$  неизвестных параметров модели, т.е.  $\nu_1 = N - l = 25 - 15 = 10$ .

Следовательно, дисперсия неадекватности в данном случае  $S_{ao}^2 = 5,22$  и дисперсионное отношение  $F = 5,22/1,19 = 4,38$ . табличное значение  $F$  - критерия при  $\alpha = 0,05$   $F_T(\alpha, v_1, v_2) = 8,79$ .

Так как  $F = 4,38 \leq F_T(\alpha, v_1, v_2) = 8,79$ , то модель считается *адекватной*.

Проверка значимости коэффициентов (при  $\alpha = 0,05$ ) показывает, что незначимы коэффициенты при  $x_1^2$ ,  $x_1x_2$ ,  $x_2x_3$ , которые следует исключить из модели, причем в силу ортогональности плана все остальные коэффициенты не изменяются, кроме  $b_0$ , который в соответствии с (5.25) будет равен 49,7.

*Рототабельные центральные композиционные планы (РЦКП)* - это планы, которые обеспечивают постоянство дисперсии отклика  $Dy^*(x)$  на равных расстояниях от центра эксперимента (независимо от направления). При этом свойство ортогональности теряется, хотя информационная матрица таких планов  $F^T F = M$  близка к диагональной, что позволяет получить удобные для ручного счета формулы оценок коэффициентов модели и их дисперсий [2].

*РЦКП* строится аналогично ортогональному плану, т.е. к «ядру» плана, в качестве которого может быть использован полный ( $2^m$ ) либо дробный ( $2^{m-p}$ ) факторный план, добавляются  $2m$  звездных точек с «плечом»  $\alpha = 2^{(m-p)/4}$  и  $n_0 > 1$  опытов в центре плана. Величина  $n_0$  выбирается из довольно искусственных соображений о постоянстве  $Dy^*(x)$  внутри сферы единичного радиуса с центром в начале координат (кодированных переменных). Её значения для разных  $m$  приведены в [2].

Поскольку рототабельные планы не имеют явных преимуществ перед ОЦКП, проигрывая последним по точности, измеряемой обобщенной дисперсией (что будет отмечено ниже), мы не будем подробней их здесь рассматривать, отсылая интересующихся к [2].

*Планами типа  $V_m$*  называются композиционные планы с «ядром» в виде полного ( $2^m$ ) факторного плана, дополненного  $2m$  звездными точками с «плечом»  $\alpha = 1$ . Иными словами, план типа  $V_m$  состоит из  $2^m$  вершин  $m$ -мерного гиперкуба с координатами  $\pm 1$  и  $2m$  центров  $(m-1)$ -мерных граней. Общее число точек плана  $N = 2^m + 2m$ . Эти планы изучались и сравнивались с непрерывными  $D$ -оптимальными планами в [23,24.25], где показано их существенное превосходство над рассмотренными выше ОЦКП и РЦКП.

Сопоставление этих планов приведено в табл. 5.7, заимствованной из [24] (где  $d_{max} = Dy^*(x)$ )

Таблица 5.7

Тип плана	m=4		m=5		m=6	
	N	$d_{max}$	N	$d_{max}$	N	$d_{max}$
Непрерывный D-оптимальный $V_m$	-	15	-	21	-	28
Ортогональный Рототабельный	24	18,5	42	34	76	66
	25	-	43	149	77	312
	31	267	52	692	91	1726

Из табл. 5.7 видно, что по такому важному критерию, как максимальное значение дисперсии предсказания  $d_{max}$ , планы типа  $V_m$ , несомненно уступают непрерывным  $D$ -оптимальным планам и существенно превосходят ОЦКП и РЦКП Бокса (при меньшем числе опытов  $N$ ). Дополнительное преимущество планов типа  $V_m$ , состоит в том, что их звездные точки находятся в пределах заданной области эксперимента  $\Omega$ , а не выходят за неё, что часто бывает невозможно осуществить.

В работе [26] получены формулы для оценок коэффициентов квадратичной модели и их дисперсий, имеющие вид:

$$b_0 = \frac{1}{2(m-1)} \left( \sum_{j=N_0+1}^N \bar{y}_j - \frac{1}{2^{m-1}} \sum_{j=1}^{N_0} \bar{y}_j \right), N_0 = 2^m \quad (5.27)$$

$$b_{jj} = \frac{1}{2} \sum_{j=N_0+1}^N (x_{ij})^2 \bar{y}_j - b_0, \quad b_j = \frac{1}{2 + 2^m} \sum_{j=1}^{N_0} x_{ij} \bar{y}_j; \quad (5.28)$$

$$b_{jk} = \frac{1}{2^m} \sum_{j=1}^{N_0} x_{ij} x_{kj} \bar{y}_i; \quad (5.29)$$

$$Db_0 = \frac{c_1 \sigma^2}{r}; \quad Db_{ii} = \frac{c_2 \sigma^2}{r}; \quad Db_i = \frac{\sigma^2}{r(2 + 2^m)}; \quad Db_0 = \frac{\sigma^2}{r 2^m};$$

где  $c_1 = (2 + m 2^m) / (m-1) 2^{m+1}$ ,  $c_2 = \{ [m(m-3) + 3] 2^m + 2 \} / (m-1) 2^{m+1}$ ;  $N_0 = 2^m$  число точек «ядра» плана;  $N = N_0 + 2m$ ;

$\bar{y}_i$  - среднее значение отклика в  $j$ -й точке плана, определяемое по  $r$  параллельным опытам;  $\sigma^2$  - дисперсия воспроизводимости.

Если некоторые из оценок параметров вида  $b_{ij}$  окажутся незначимыми, то необходимо произвести уточнение оценок остальных параметров  $b_{ij}$  и параметра  $b_0$  по формулам:

$$b_0 = \frac{1}{A} \{ (n - v - 1) [2^{m-1} \sum_{j=N_0+1}^N \bar{y}_j - \sum_{j=1}^{N_0} \bar{y}_j] + (1 + 2^{m-1}) \sum_{i=1}^v \sum_{j=N_0+1}^N (x_{ij})^2 \bar{y}_j \};$$

$$b_{ij} = \frac{1}{A} \sum_{j=N_0+1}^N \bar{y}_j (x_{ij})^2 - \frac{1}{A} \{ (m-1) [2^{m-1} \sum_{j=N_0+1}^{N_0} \bar{y}_j - \sum_{j=1}^{N_0} \bar{y}_j] - [(m-2)2^{m-1} - 1] \sum_{i=1}^v \sum_{j=N_0+1}^N (x_{ij})^2 \bar{y}_j \}$$

(5.30)

где  $A = 2^m [(m-1)^2 - v(n-2)] + 2v$ .

В случае незначимости некоторых из оценок параметров вида  $b_i$  или  $b_{ik}$  уточнение оценок остальных параметров не производится.

Дисперсия оценки отклика определяется в виде

$$Dy^*(x) = \frac{\sigma^2}{r} (c_1 + c_4 \sum_{j=1}^m x_j^2 + c_2 \sum_{j=1}^m x_j^4 + c_5 \sum_{j < k} x_j^2 x_k^2) \quad (5.31)$$

где  $c_4 = \frac{1}{2 + 2^m} - c_3$ ,  $c_5 = [(m-1)^2 - (m-2)2^m + 1] / (m-1)^2 2^m$ .

Приведем пример использования плана типа  $B_m$  [27].

*Пример 2.2.* Изучался процесс электролиза для нанесения медного покрытия на стальные детали различной формы. В качестве отклика у рассматривалась доля анодного выхода металла, в качестве факторов были выбраны следующие переменные: концентрация главного компонента  $z_1$  значение рН раствора  $z_2$ , катодная плотность тока  $z_3$ .

Область действия указанных факторов и кодированные переменные имеют вид (без указания единиц измерения  $z_k$ ):

$$0 \leq z_1 \leq 120 \Rightarrow x_1 = \frac{(z_1 - 60)}{60};$$

$$8,2 \leq z_2 \leq 9 \Rightarrow x_1 = \frac{(z_2 - 8,6)}{0,4};$$

$$0,5 \leq z_3 \leq 4 \Rightarrow x_1 = \frac{(z_3 - 2,25)}{1,75}.$$

В каждой точке плана  $B_3$ , состоящем из  $N_0 = 2^3 = 8$  точек «ядро» плана, а также в  $2m=6$  точках, расположенных в центрах граней 3-мерного куба и в центре плана (всего  $N=15$ ) было проведено  $r=3$  параллельных опыта, на основании которых вычислены следующие усредненные значения отклика:

- 1)  $\bar{y}_1=80,0$     4)  $\bar{y}_4=87,4$ ;    7)  $\bar{y}_7=62,0$ ;
- 2)  $\bar{y}_2=119,0$ ;    5)  $\bar{y}_5=62,5$ ;    8)  $\bar{y}_8=102,0$ ;
- 3)  $\bar{y}_3=75,0$ ;    6)  $\bar{y}_6=230,0$ ;    9)  $\bar{y}_9=81,3$ ;
- 10)  $\bar{y}_{10}=893$ ;    12)  $\bar{y}_{12}=82,0$ ;    14)  $\bar{y}_{14}=107,0$ ;
- 11)  $\bar{y}_{11}=91,0$ ;    13)  $\bar{y}_{13}=68,2$ ;    15)  $\bar{y}_{15}=84,4$ .

По формулам (5.27) - (5.30) были вычислены оценки параметров квадратичной модели и их дисперсии, в результате чего получена модель следующего вида (где все коэффициенты оказались значимыми):

$$y^*(x) = 77,5 - 29,7x_1 + 17,3x_2 - 10,4x_3 + 9,1x_1^2 - 8,0x_2^2 + 6,8x_3^2 - 19,3x_1x_2 + 19,4x_1x_3 - 11,4x_2x_3 \quad (5.32)$$

Оценка дисперсии воспроизводимости отклика (по  $r=3$  параллельным опытам), усредненная по всем  $N=15$  точкам, равна 5,93.

Оценка дисперсии адекватности модели  $\sigma_{ad}^2$  равна 219,6. При уровне значимости  $\alpha=0,05$  гипотеза об адекватности модели по критерию Фишера не подтвердилась. Это не означает, что модель (5.32) не дает полезной информации об изучаемом процессе. Её можно использовать для приближенного отыскания оптимальных условий протекания процесса. Последующее уточнение может быть найдено преобразованием этой модели с помощью перехода к новым факторам, специальным образом связанными с исходными. Пример такого перехода приведен в [28].

*Некоторые сведения о D-оптимальных планах.* Напомним, что D-оптимальный план  $\Pi$  позволяет получить оценки параметров регрессионной модели, имеющие минимально возможное рассеяние относительно центра распределения, так как он обеспечивает минимальный объем эллипсоида рассеяния оценок.

Теория D - оптимальных планов [8] развита применительно к так называемым *непрерывным планам*, под которыми понимаются планы вида

$$\Pi = \begin{bmatrix} x_1 & x_2 & \dots & x_n \\ v_1 & v_2 & \dots & v_n \end{bmatrix},$$

где  $\sum_{i=1}^n v_i = 1$ , а величина  $v_i$  может принимать любые значения между 0 и 1;

$v_i$  имеет смысл доли или частоты наблюдений в точке  $x_i$ .

Если  $v_i = \frac{r_i}{N}$ , т.е. существует такое целое число N, что число

$Nv_i = r_i$ , тоже является целым, то непрерывный план называется точным, т.е. его можно в точности реализовать при проведении эксперимента. Однако не для всякого N можно найти такой точный план, что соотношение частот наблюдений в n точках плана было одинаковым с непрерывным планом. Это можно сделать, очевидно, лишь тогда, когда все произведения  $Nv_i$  -

целые числа. Например, если в точках  $x_1, x_2, x_3$  частоты наблюдений  $v_1=0,2; v_2=0,5; v_3=0,3$ , то такой план будет точным, ибо при  $N=10$  имеем  $r_1 = Nv_1=2; r_2 = Nv_2=5; r_3 = Nv_3=3$ .

Но если частоты равны  $v_1=0,5; v_2=0,5 \cdot \frac{1}{\sqrt{2}}; 1/\sqrt{2}; v_3 = \frac{1}{\sqrt{2}}$ , то точного плана не существует, однако можно подобрать всегда приближенный дискретный план.

Приведём без доказательств ряд теорем о непрерывных D-оптимальных планах, характеризующих их «хорошие» свойства [2, 8, 25].

**Теорема 2.1.** Всегда существует непрерывный D -оптимальный план, сосредоточенный в конечном числе точек n, причем

$$l+1 \leq n \leq \frac{(l+1)(l+2)}{2} = C_{l+2}^2$$

где  $l+1=l^*$ - число неизвестных коэффициентов регрессионной модели.

Эта теорема позволяет, экспериментатору из бесконечного множества непрерывных D -оптимальных планов для выбранной модели выделить те, которые сосредоточены в малом числе точек.

**Теорема 2.2.** Непрерывный план  $\Pi$  тогда и только тогда D -оптimalен, когда

$$\max_{x \in \Omega} d(x) = l+1$$

где  $\Omega$  область планирования эксперимента.

Эта теорема даёт способ проверки всякого непрерывного плана на близость к D - оптимальному путем вычисления показателя

$$\delta = (\max_{x \in \Omega} d(x)/l^*) - 1$$

Чем ближе  $\delta$  к нулю, тем ближе рассматриваемый непрерывный план и D -оптимальному.

**Теорема 2.3.** D -оптимальный план не меняется при линейном преобразовании области эксперимента  $\Omega$ .

Это позволяет рассматривать D -оптимальные планы для некоторых стандартных областей, например, для гиперкуба с координатами вершин в точках  $\pm 1$ , для шара и т.д. Изменение масштаба при переходе к исходным единицам измерения не повлияет на расположение и частоты точек D - оптимального плана. На основе этого свойства составлены каталоги D - оптимальных планов для типовых полиномиальных моделей [4].

**Теорема 2.4.** Непрерывный D -оптимальный план является одновременно и G-оптимальным (т.е. минимизирует  $\max d(x)$ ), а если, кроме того, информационная матрица плана является диагональной, то он является и A - оптимальным.

В частности, как уже отмечалось выше, планы полного и дробного факторного эксперимента для линейных моделей 1-го порядка являются D, - G- и A -оптимальными при планировании на гиперкубе с координатами вершин в точках  $\pm 1$ , если матрица F плана не содержит одинаковых столбцов.

*Непрерывные D -оптимальные планы на отрезке  $\Omega = (-1, 1)$ .* В случае одного фактора непрерывные D - оптимальные планы найдены и в аналитическом виде для полиномиальной модели произвольного порядка

$$\eta = \sum_{k=0}^d \beta_k(x_k)$$

Эти планы сосредоточены в  $d+1$  точке, причем частота всех точек одинаковы, т.е.  $\nu_1 = \nu_2 = \dots = \nu_{d+1} = \frac{1}{d+1}$ , а координаты точек плана  $x_i$  являются корнями полинома  $G(x)$  степени  $d+1$ , представляющего собой коэффициент при  $z^{d+1}$  в разложении функции

$$g(x) = \sqrt{1 + 2xz + z^2} \text{ в ряд Тейлора по степеням } z \text{ в точке } z=0.$$

*Пример 2.3.* Определить координаты точек непрерывного  $D$ -оптимального плана для квадратичной модели  $\max \eta(x) = \beta_0 + \beta_1 x + \beta_2 x^2$ . Найдем коэффициент разложения  $g(x)$  по степеням  $z$  при  $z^3$  (полагая  $z=0$ );

$$\frac{1}{3} \frac{\partial^3 g(z)}{\partial z^3} \Big|_{z=0} = \frac{1}{3} \frac{\partial^3}{\partial z^3} \sqrt{1 - 2xz + z^2} \Big|_{z=0} = \frac{1}{2} x(1 - x^2)$$

Следовательно, искомые точки  $D$ -оптимального непрерывного плана есть корни полинома  $G(x) = \frac{1}{2} x(1 - x^2)$  т.е.  $x_1 = 0, x_2 = \pm 1, x_3 = -1$ ,

а частоты равны  $\nu_1 = \nu_2 = \nu_3 = \frac{1}{3}$ . Отсюда легко получить точный  $D$ -

оптимальный план, если положить  $N = r(d+1)$ , где  $r$  - число повторений опыта в каждой точке задаваемое произвольно.

В [2] приведены значения координат точек непрерывных  $D$ -оптимальных планов для степени полинома  $d = 2, 3, \dots, 8$ , а также дисперсионные матрицы этих планов. •

*Непрерывные  $D$ -оптимальные планы Кифера и Коно.* В случае квадратичной модели (2.26) и планирования на гиперкубе  $\Omega = \{|x_i| \leq 1, i=1, 2, \dots, m\}$  в работах Кифера [29] и Коно [30] непрерывные  $D$ -оптимальные планы построены аналитически. Заметим, что в общем случае (т.е. для более сложных моделей и областей  $\Omega$ ) непрерывные  $D$ -оптимальные планы можно построить лишь численными методами.

Планы Кифера сосредоточены в следующих точках гиперкуба:  
- в вершинах гиперкуба (множество точек плана  $E_0$ ) с одинаковой частотой  $\nu_0$ ;

- в серединах ребер (множество  $E_1$ ) с одинаковой частотой  $\nu_1$ ;

- в центрах двумерных граней (множество  $E_2$ ) с одинаковой частотой  $\nu_2$

Планы Коно отличаются от планов Кифера лишь тем, что вместо множества точек  $E_n$  берется центр гиперкуба  $E_n$ , с частотой  $\nu_n$ .

В табл. 5.9 приведены значения частот для этих планов (знаменатель соответствует плану Коно) и общее количество точек планов. Отметим, что планы Кифера существуют лишь для числа факторов  $m \leq 5$ . В последнем столбце табл. 5.9 для сравнения приведены максимальные числа  $n_{\max}$  точек непрерывного  $D$ -оптимального плана.

Из табл. 5.9 видно, что планы Коно начиная  $m \geq 3$  являются более экономичными, чем планы Кифера, хотя и они при  $m = 6, 7$  превышают максимально необходимое число точек  $n_{\max}$ .

Практическое использование этих планов состоит в приближенной замене непрерывного  $D$ -оптимального плана дискретным планом. Например, для  $m = 4$  план Коно содержит 49 точек с частотами 0,45 для вершин куба, 0,5 - для середины ребер и 0,047 - для центра плана.

Таблица 5.9

m	Вершина куба $E_0$		Середина ребер $E_1$		Центры граней $E_2$	Центр куба $E_n$		Общее число точек плана		Число коэф-фициентов	$n_{\max}$	
	$\nu_0$	$n_0 = 2^m$	$\nu_1$	$n_1$	$\nu_2$	$n_2$	$\nu_n$	$n_n$	Ки-фера		Ко-но	$l+1$
1	0,666	2	0,333	1	0	0	-	-	3	3	3	5
2	0,583	4	0,321	4	0,096	1	0,096	1	9	9	6	15
3	0,576/0,61	8	0,227/0,42	12	0,197	6	0,065	1	26	21	10	35
4	0,593/0,45	16	0,123/0,502	32	0,284,	24	0,047	1	72	49	15	70
5	0,617/0,4	32	0,025/0,56	80	0,358	80	0,037	1	192	113	21	126
6	-/0,36	64	-/0,61	192	-	-	0,027	1	-	257	28	210
7	-/0,33	128	-/0,65	448	-	-	0,022	1	-	577	36	330

Если выделено  $N = 100$  наблюдений, то их нужно приблизительно расположить так: по  $r=3$  наблюдения в вершинах куба ( $3 \times 16 = 48, N\nu_1=45$ ),

по  $r=2$  наблюдения в серединах ребер ( $2 \times 32=64, N\nu_2=50$ ) и

$N\nu_3=4,7 \approx 5$  наблюдений в центре - всего 119 опытов вместо 100.

В тех случаях, когда область эксперимента не имеет правильной геометрической формы, а модель представляется неполным полиномом второй степени или полиномом более высокой степени, для построения D - оптимальных непрерывных планов приходится применять численные методы с использованием процедуры последовательного планирования эксперимента [2,8].

## 6. ПЛАНИРОВАНИЕ ЭКСПЕРИМЕНТА ПРИ ПОИСКЕ ОПТИМАЛЬНЫХ УСЛОВИЙ

Как уже отмечалось, одной из основных задач, решаемых экспериментатором, является поиск оптимальных условий поведения объекта или, другими словами, поиск значений факторов  $X=X^*$ , при которых отклик (или некоторый функционал от него) достигает экстремума.

Для решения этой задачи существуют два различных подхода.

*Первый*, традиционный, основан на предварительной построении аналитической модели объекта и на применении различных методов оптимизации, отмеченных, например, в [7, 22].

*Второй* подход, появившийся вместе с теорией планирования эксперимента, не предполагает известным математическое описание объекта и требует лишь возможности проводить на объекте эксперимент. Это означает, что экспериментатор может получить значение отклика у при любом, вообще говоря, интересующем его значении вектора – факторов  $x$ .

В настоящее время известны несколько вариантов реализации второго подхода:

- метод крутого восхождения, или метод Бокса-Уилсона (по имени авторов, опубликовавших его в 1951 г.);
- симплексный метод [2, 3, 7];
- метод эволюционного планирования [2].

Здесь мы ограничимся рассмотрением первого метода.

### 6.1 Метод крутого восхождения

Ранее описаны методы построения экспериментально-статистических моделей в виде уравнений регрессии. Рассмотрим использование этих моделей для оптимизации процессов или свойств многокомпонентных систем. Качество процесса характеризуется несколькими функциями отклика, но найти такое сочетание значений влияющих факторов, при котором одновременно достигаются экстремумы всех интересующих экспериментатора функций отклика трудно. Отметим, что и влияющие факторы, и функции отклика изменяются только в определенных пределах. Так, концентрации реагентов не могут быть отрицательными, температура и давление в аппарате не должно превышать безопасных пределов и т. п.

Следовательно, оптимизацию процессов, как правило, осуществляют в условиях ограничений на влияющие факторы и функции отклика.

Величина, характеризующая уровень оптимизации процессов, называется критерием оптимальности [64, 7].

В частном случае критерием оптимальности может быть одна из функций отклика, характеризующих процесс. Оптимизация процесса представляет собой целенаправленный поиск значений влияющих факторов, при которых достигается экстремум критерия оптимальности с учетом ограничений, наложенных на все влияющие факторы и функции отклика.

Г. Бокс и К. Уилсон [64] предложили использовать для оптимизации результаты полного факторного эксперимента или эксперимента по методу дробных реплик. Сущность такой оптимизации состоит в следующем. Пусть, например, критерием оптимальности служит функция отклика  $y$ , представленная в виде уравнения (4.5). По теореме о разложении аналитической функции в ряд Тейлора частные производные функции отклика по факторам равны по величине и знаку соответствующим коэффициентам регрессии.

Следовательно, при изменении факторов пропорционально их коэффициентам регрессии в ту сторону, которую указывает знак коэффициента, движение осуществляется по градиенту. Эффективность градиента существенно зависит от характера поверхности отклика. Можно доказать, что движение из некоторой точки поверхности в направлении градиента — кратчайший путь к оптимуму. Это утверждение следует, например, из теории поля, согласно кота: градиент скалярной функции задается соотношением

$$\text{Grad } F = \left( \frac{\partial F}{\partial x_1} \vec{i} + \frac{\partial F}{\partial x_2} \vec{j} + \dots + \frac{\partial F}{\partial x_m} \vec{m} \right) \quad (6.1)$$

где  $\text{grad } F$  — обозначение градиента;  $\frac{\partial F}{\partial x_i}$  — частная производная

функции отклика по  $i$ -му фактору  $\vec{i}, \vec{j}, \dots, \vec{m}$  — единичные векторы в направлении координатных осей факторного пространства.

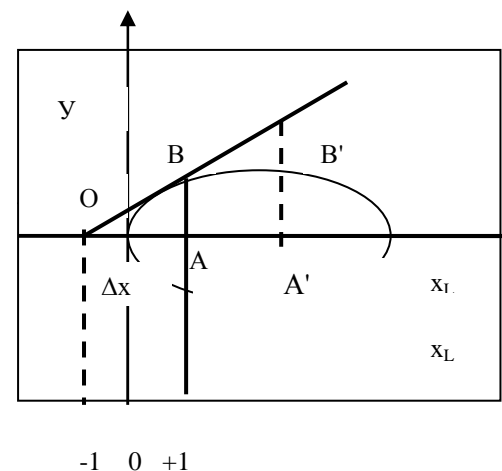


Рис.6.1. Демонстрация метода крутого восхождения

Вычисление значений функций, лежащих на градиенте, удобно рассматривать на простейшем примере случае одного фактора (рис. 6.1). Если значение коэффициента регрессии, равно тангенсу угла между линейной функцией и осью данного фактора, умножить на интервал варьирования, который является прилежащим катетом в прямоугольном треугольнике  $OAB$ , то противолежащий катет— $AB$ , дает координаты точки, расположенной на градиенте. Обобщение на случай  $k$  фактов делается механически — все эффекты независимы друг от друга. Важно, что существенно только соотношение произведений коэффициентов на соответствующие интервалы. Их абсолютные величины могут все одновременно умножаться или делиться на любое положительное число. Таким образом, находим точки, лежащие на градиенте, только с другим шагом. Последовательно алгебраически прибавляя строки из величин, пропорциональных составляющим градиента, к нулевому уровню, получают шаги.

Примем один из влияющих факторов за базовый, и для него вычислим произведение соответствующего коэффициента регрессии на шаг варьирования. Например, для первого фактора это произведение имеет вид  $b_1 \Delta x_1$ .

Затем для базового фактора выберем шаг движения  $\Delta x_1^*$ , с которым будет осуществляться оптимизация. Обычно  $\Delta x_1^* \leq \Delta x_1$ . После этого вычисляем



$$\gamma = \frac{\Delta x_i^*}{b_i \Delta x_i} \quad (6.2)$$

Для всех остальных факторов шаги движения к оптимальным значениям рассчитываем по формуле

$$\Delta x_i^* = \gamma b_i \Delta x_i \quad (6.3)$$

Движение к оптимуму начинаем из центра плана, который пользовался для получения математического описания функции отклика. Значения факторов на каждом новом шаге определяем путем прибавления  $\Delta x_i^*$  к соответствующим предыдущим значениям.

Так осуществляется оптимизация при методе движения по градиенту—по методу *крутого восхождения*.

При отыскании минимума функции отклика новые значения факторов находим из предыдущих путем вычитания  $\Delta x_i^*$ . Такой способ оптимизации называется методом *наискорейшего спуска*.

Движение к оптимуму прекращают в следующих случаях.

1. Значения одного или нескольких факторов или функций отклика вышли на границы допустимых значений.
2. Экстремум критерия оптимальности у достигнут.

В первом случае на этом оптимизация заканчивается, а во втором — в области экстремума функции у ищут ее новое математическое описание, используя полный факторный эксперимент или метод дробных реплик. Получив адекватное описание этой функции в виде (4.5), продолжают оптимизацию методом крутого восхождения. Если же в области оптимума не удастся получить адекватного уравнения регрессии вида (4.5), переходят к планированию эксперимента по математическому описанию функции у в виде многочлена более высокой степени.

*Пример 4.* Пусть в результате полного факторного эксперимента (пример 2) получено адекватное уравнение регрессии

$$y_1 = 35,6 + 1,95x_1 - 1,35x_2 \quad (6.4)$$

Здесь  $y_1$  — выход продукта реакции,  $x_1$  — температура,  $x_2$  — концентрация реагента.

Температура Т, при построении уравнения регрессии, варьировалась в пределах  $45^{\circ}\text{C} \div 55^{\circ}\text{C}$ , а концентрация n в пределах  $24\% \div 26\%$ .

Введем также в рассмотрение функцию отклика  $y_2$ , характеризующую скорость химической реакции ( $\text{кмоль} \cdot \text{м}^3 \cdot \text{ч}^{-1}$ ). Пусть требуется, чтобы выполнялось условие  $y_2 \geq 2,5$ .

Допустим, что ограничения на влияющие факторы имеют вид  $30^{\circ} \leq T = z_1 \leq 120^{\circ}$ ;  $10\% \leq n = z_2 \leq 70\%$ .

Будем оптимизировать выход продукта реакции методом крутого восхождения.

*Решение.* Найдем центр плана, используя в размерной (реальной) системе координат

$$z_{01} = \frac{z_{1\max} + z_{1\min}}{2} = \frac{55 + 45}{2} = 50$$

$$z_{02} = \frac{z_{2\max} + z_{2\min}}{2} = \frac{26 + 24}{2} = 25$$

Найдём интервалы варьирования

$$\Delta z_1 = \frac{z_{1\max} - z_{1\min}}{2} = \frac{55 - 45}{2} = 5$$

$$\Delta z_2 = \frac{z_{2\max} - z_{2\min}}{2} = \frac{26 - 24}{2} = 1$$

В качестве базового фактора возьмем температуру. Поскольку шаг движения при крутом восхождении должен подчиняться неравенству  $\Delta z_i^* \leq \Delta z_i$ , т.е.  $\Delta z_1^* \leq 5^{\circ}$ , выбираем шаг движения на крутом восхождении равным  $4^{\circ}$ , тогда

$$\gamma = \frac{\Delta z_1^*}{b_1 \Delta z_1} = \frac{4}{1,95 \times 5} = 0,41$$

Шаг по концентрации на крутом восхождении

$$\Delta z_2^* = \gamma b_2 \Delta z_2 = 0,41(-1,35) \times 1 = -0,55\%$$

Для удобства ведения эксперимента шаги движения, рассчитанные по формуле (26.1), можно округлять. В данном случае удобно принять

$$\Delta z_2^* = -0,5\%$$

В безразмерной (кодированной) системе координат, к которой переходят по формулам

$$x_k = \frac{z_{0k} - z_k}{\Delta z_k},$$

шаги восхождения будут равны

$$\Delta x_1^* = \frac{4}{5} = 0,8, \Delta x_2^* = \frac{24,5 - 25}{1} = -0,5$$

Обозначим через  $y_1^p$  - рассчитанное по выражению (6.4) значение функции отклика  $y_1$ , а через  $y_1^э$  - экспериментальное значение той же функции;  $y_2^э$  - экспериментальное значение функции  $y_2$ .

Результаты, полученные по методу крутого восхождения внесём в таблицу 6.1

Таблица 6.1

Результаты опытов по методу крутого восхождения

Характеристика и номер опыта	Z <sub>1</sub>	Z <sub>2</sub>	X <sub>1</sub>	X <sub>2</sub>	y <sub>1</sub> <sup>p</sup>	y <sub>1</sub> <sup>э</sup>	y <sub>2</sub> <sup>э</sup>
Центр плана	50	25	0	0	35,6	35,1	2,9
Интервал варьирования	5	1	1	1	-	-	-
Шаг движения	4	-0,5	0,8	-0,5	-	-	-
№ опыта 1	54	24,5	0,8	-0,5	36,5	36,9	3,2
2	58	24,0	1,6	-1,0	37,4	37,2	3,7
3	62	23,5	2,4	-1,5	38,2	38,5	2,8
4	66	23,0	3,2	-2,0	39,1	40,7	2,3
5	70	22,5	4,0	-2,5	40,0	38,1	1,9
6	74	22,0	4,8	-3,0	40,9	37,2	1,6

Как видно из результатов опытов, выполненных по методу крутого восхождения (табл. 6.1), в опыте 4 достигнут максимальный выход продукта реакции, однако скорость процесса в этом случае меньше допустимого значения. По-видимому, оптимальным режимом процесса следует считать условия опыта 3. Ограничения на  $z_1$  и  $z_2$  в ходе оптимизации не нарушены.

## 6.2 Последовательное планирование эксперимента

Очень часто обработку в планирование эксперимента приходится вести по мере поступления данных, т.е. последовательно. Например, если построенная по исходным данным регрессионная модель оказалась неадекватной, то дополнительный эксперимент позволяет уточнить её, для чего нужно, вообще говоря, пересчитать оценки неизвестных параметров (исключение составляет ортогональное планирование). Можно, конечно, каждый раз при получении новой порции результатов эксперимента решать задачу как бы заново, используя всю накопленную информацию. Однако такой путь весьма нерационален. Гораздо удобнее с вычислительной точки зрения при последовательном способе построения модели использовать рекуррентные алгоритмы оценивания. Такой путь особенно удобен при использовании ЦВМ.

Приведем здесь краткое изложение такого подхода, отсылая интересующихся к литературе [2, 9].

Будем искать регрессионную модель уже отмеченного ранее вида (см. раздел 4. 1):

$$\eta(x) = \sum_{k=0}^l \beta_k f_k(x); \text{ или } \eta = \beta^T f(x)$$

Пусть на основе N наблюдений найдены оценка  $b_N$  вектора коэффициентов  $\beta$ , а также дисперсионная матрица  $C_N$  и остаточная сумма квадратов  $S_N$ , имеющих известные уже нам выражения, снабженные лишь индексом N

$$b_N = C_N F_N^T Y_N, \quad C_N = (F_N^T F_N)^{-1} \quad (6.5)$$

$$S_N = (Y_N - F_N b_N) (Y_N - F_N b_N)^T = Y_N^T Y_N - (b_N)^T F_N^T Y_N \quad (6.6)$$

Пусть теперь к N имевшимся наблюдениям добавлены ещё M наблюдений в точках  $x_{N+1}, \dots, x_{N+M}$  определяющих матрицу независимых переменных  $f_M = [f(x_{N+1}), \dots, f(x_{N+M})]^T$  и матрицу  $Y_M = [Y_{N+1}, \dots, Y_{N+M}]^T$ .

Предполагается, что дисперсия ошибки наблюдений в новых точках та же, что и в предыдущих. Требуется построить алгоритм вычисления:

- 1) оценок  $b_{N+M}$  через оценку  $b_N$  матрицы  $f_M$  и  $Y_M$ ;
- 2) дисперсионной матрицы  $C_{N+M}$  через матрицы  $C_N$  и  $f_M$ ;
- 3) остаточной суммы квадратов  $S_{N+M}$  через  $S_N$ ,  $b_N$ ,  $f_M$  и  $Y_M$ .

В [2] приведен вывод следующих соотношений, дающих ответ на поставленные вопросы:

$$C_{n+M} = C_n - C_n F_M^T [F_M C_n F_M^T + I_M]^{-1} F_M C_n \quad (6.7)$$

где  $I_M$  - единичная матрица размера  $(M \times M)$ ;

$$b_{N+M} = b_N + C_n F_M^T [F_M C_n F_M^T + I_M]^{-1} (Y_M - F_M b_N); \quad (6.8)$$

$$S_{n+M} = S_n (Y_M - F_M b_N)^T (I_M + F_M C_n F_M^T)^{-1} (Y_M - F_M b_N) \quad (6.9)$$

Приведенные соотношения принимают особенно простой вид при  $M = 1$ , т.е. когда пересчет оценок коэффициентов ведется после каждого дополнительного опыта:

$$C_{n+1} = C_n - \frac{C_n f(x_{N+1}) f^T(x_{N+1}) C_n}{1 + f^T(x_{N+1}) C_n f(x_{N+1})} \quad (6.10)$$

$$b_{N+1} = b_N + \frac{[Y_{N+1} - f^T(x_{N+1}) b_N]}{1 + f^T(x_{N+1}) C_n f(x_{N+1})} \quad (6.11)$$

$$S_{n+1} = S_n + \frac{C_n f(x_{N+1}) [Y_{N+1} - f^T(x_{N+1}) b_N]}{1 + f^T(x_{N+1}) C_n f(x_{N+1})} \quad (6.12)$$

Зная остаточную сумму квадратов, легко вычислить остаточную дисперсию (дисперсию неадекватности):

$$(S_{ад}^2)_{N+M} = S_{N+M} / (N+M+1)$$

Проиллюстрируем применение приведенных формул.

*Пример 3.2.* Пусть зависимость между откликом  $y$  и тремя факторами  $x_1, x_2, x_3$  ищется в виде

$$\eta(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$$

где  $x_i$  - кодированные переменные  $x_1 = z_1 - 4$ ;  $x_2 = (z_2 - 1,4)/0,1$ ;  $x_3 = (z_3 - 3)/0,5$

Сначала был реализован ДФЭ  $2^{3-1}$  с генерирующим соотношением

ищем

$x_3 = x_1 x_2$ . Статистический анализ результатов показал, что  $b_2 \cong 0$ . Причина такого результата могла быть объяснена чрезмерно большой дисперсией ошибок наблюдений, в связи с чем было предложено продолжить эксперимент для уточнения модели. После добавления двух опытов взятых для плана

$2^{3-1}$  с генерирующим соотношением  $x_3 = x_1 x_2$  были получены значимые оценки всех коэффициентов и опыты были прекращены.

Результаты расчетов приведены в табл. 6.2. Оценка дисперсии воспроизводимости, определенная по ряду параллельных опытов, равна 10,9 при числе степеней свободы  $\nu_2 = 4$ .

Таблица 6.2

i	План			$y_i$	Оценки параметров				$D b_{ki}$	$t_{кр}(S_k^2)_i$	$S_i$	$(S_k^2)_i$	$F = \frac{(S_k^2)_i}{S}$
	$x_{1i}$	$x_{2i}$	$x_{3i}$		$b_{0i}$	$b_{1i}$	$b_{2i}$	$b_{3i}$					
1	-1	-1	-1	45	-	-	-	-	-	-	-	-	-
2	+1	-1	+1	84,6	-	-	-	-	-	-	-	-	-
3	-1	+1	+1	59,1	-	-	-	-	-	-	-	-	-
4	+1	+1	-1	82,4	67,77	15,72	2,97	4,07	2,73	3,53	0	0	-
5	-1	-1	+1	58,2	68,4	15,09	3,34	4,70	2,38	3,28	12,7	12,75	1,17
6	+1	-1	-1	57,9	67,3	14,02	3,41	5,77	2,04	3,04	49,3	24,68	2,26

Используя формулы (6.10), (6.11), найдем

$$C_5 = \frac{1}{32} \begin{bmatrix} 7 & 1 & 1 & -1 \\ 1 & 7 & -1 & 1 \\ 1 & -1 & 7 & 1 \\ -1 & 1 & 1 & 7 \end{bmatrix}$$

$$b_5 = \begin{bmatrix} 67,8 \\ 15,7 \\ 2,97 \\ 4,07 \end{bmatrix} + \frac{1}{8} \begin{bmatrix} 1 \\ -1 \\ 1 \\ 1 \end{bmatrix} (58,2 - 53,15) = \begin{bmatrix} 68,4 \\ 15,09 \\ 2,34 \\ 4,7 \end{bmatrix}, S_5 = \frac{1}{2} (58,2 - 53,15)^2 = 12,75$$

$$\text{Дисперсии оценок коэффициентов равны } (S_k^2)_5 = \frac{7}{32 S^2} = 2,38$$

из табл. 6.2 видно, что после 5-го опыта коэффициент при  $x_2$  остается незначимым.

Вычисления после 6-го опыта дают такие результаты:

$$b_5 = \begin{bmatrix} 68,4 \\ 15,09 \\ 2,34 \\ 4,7 \end{bmatrix} + \frac{1}{64} \begin{bmatrix} 7 & 1 & 1 & -1 \\ 1 & 7 & -1 & 1 \\ 1 & -1 & 7 & 1 \\ -1 & 1 & 1 & 7 \end{bmatrix} \times \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} (67,9 - 76,45) = \begin{bmatrix} 67,33 \\ 14,03 \\ 3,41 \\ 5,77 \end{bmatrix}$$

$$S_6 = 12,75 + \frac{(67,9 - 76,45)^2}{2} = 24,68;$$

$$C_6 = \frac{1}{32} \begin{bmatrix} 6 & 0 & 2 & 0 \\ 0 & 6 & 0 & 2 \\ 2 & 0 & 6 & 2 \\ 0 & 2 & 0 & 6 \end{bmatrix} \Rightarrow (S_k^2) = \frac{6}{32} S^2 = 2,04$$

Из табл. 6.2 видно, что после 6-го опыта все коэффициенты значимо отличаются от нуля. Кроме того, сопоставляя полученное, значение  $f = 2,26$  с  $f_{кр} = 10,64$  при уровне значимости  $\alpha = 0,05$ ,  $\nu_1 = 2$ ,  $\nu_2 = 4$ , убеждаемся в адекватности линейной модели

#### ЛИТЕРАТУРА

1. Адлер Ю.П., Маркова Е.В., Грановский Ю.В. Планирование эксперимента по поиску оптимальных условий. - М.: Наука, 1988.
2. Планирование эксперимента в исследовании технологических процессов. Под ред. Левкого Э.К.. - М.: Мир, 1987.
3. Налимов В.В, Голикова Т.К. Логические основания методов планирования эксперимента. - М.: Металлургия, 1991.
4. Голикова Т.Н., Панченко О.А., Фридман М.З. Каталог планов второго порядка. Ч. I, II. - М.: Изд-во МГУ, 2000.
5. Асатурян В.Н. Теория планирования эксперимента. Ч. I. - М.: Изд. МИЭМ, 1994.
6. Адлер Ю.П. Предпланирования эксперимента. - М.: Знание, 1980
7. Мойсюк Б.Н. Элементы теории оптимального эксперимента. Ч II. - М.: изд. МЭИ, 1996.
8. Федоров В.В. Теория оптимального эксперимента. - М.: Изд. Наука, 2001.
9. Налимов В.З. Теория эксперимента. - М.: Наука, 1981.
10. Налимов В.В., Чернова Н.А. Статистические методы планирования экстремальных экспериментов. - М.: Наука, 1989
11. Гумбель К. Теория экстремальных распределений. - М.: Наука, 1999.
12. Дейвиц Г. Порядковые статистики. - М.: Наука, 1999.
13. Смирнов Н.Б., Дунин - Барковский И.В. Курс теории вероятностей и математической статистики для технических приложений. М.: Знание, 1990
14. Смирнов Г.В., Смирнов Д.Г. Статистическая связь между заболеваниями щитовидной железы, сахарным диабетом, желчекаменной и мочекаменной болезнями // Материалы Международной научно- практ. Конференции: «Электронные средства и системы управления». Томск, изд-во института оптики атмосферы СОРАН 2004г., в трех частях, часть 2 с. 40-45.
15. Смирнов Г.В., Смирнов Д.Г. Статистический анализ связи между минеральным составом воды и заболеваниями человека зобом, уролитиазом, сахарной и желчекаменной болезнями. // Материалы Международной научно- практ. Конференции: «Электронные средства и системы управления». Томск, изд-во института оптики атмосферы СОРАН 2004г., в трех частях, часть 2 с. 45-47.

16. Смирнов Г.В., Смирнов Д.Г., Зотова Н.В. Статистика некоторых заболеваний по районам Томской области// Проблемы информационной безопасности государства, общества и личности: Материалы Седьмой Всероссийской научно- практической конференции. Томск, 16-18 февраля 2005 г. – Томск: Издательство ИОА СО РАН, 2005.- с.127-130

17. Смирнов Г.В., Смирнов Д.Г., Зотова Н.В. Корреляционные связи между некоторыми заболеваниями/ Материалы I Международной научно- практической конференции по лесопользованию, экологии и охране лесов: фундаментальные и прикладные аспекты. 21-22 марта 2005 г. – Томск. С 292-293.

18. Смирнов Г.В., Смирнов Д.Г., Зотова Н.В. Корреляционные связи между некоторыми заболеваниями// Материалы I Международной научно- практической конференции по лесопользованию, экологии и охране лесов: фундаментальные и прикладные аспекты. 21-22 марта 2005 г. – Томск. С 292-293.

1	ОЦЕНКА НЕИЗВЕСТНЫХ ПАРАМЕТРОВ	5
1.1	Основные понятия и определения	5
1.2	Основные выборочные характеристики	7
1.3	Точечные оценки параметров	10
1.4	Интервальные оценки параметров	13
1.5	Точные доверительные интервалы для параметров нормального закона	14
1.6	Обработка результатов измерений (две основные задачи теория ошибок)	16
2	ПРОВЕРКА СТАТИСТИЧЕСКИХ ГИПОТЕЗ	17
2.1	Понятия критерия и правило проверки гипотез	17
2.2	Правило наилучшего выбора критической области	19
2.3	Проверку стандартных гипотез	21
3	ЭЛЕМЕНТЫ ТЕОРИИ ОПРЕДЕЛИТЕЛЕЙ	24
4	ОСНОВЫ РЕГРЕССИОННОГО АНАЛИЗА	32
4.1	Постановка задачи. Исходные предположения	32
4.2	Оценка параметров линейной модели Метод наименьших квадратов	40
4.3	Статистический анализ регрессионной модели	51
5	ЭЛЕМЕНТЫ ТЕОРИИ ПЛАНИРОВАНИЯ ЭКСПЕРИМЕНТА	61
5.1	Сопоставление возможностей пассивного и активного экспериментов	61
5.2	Однофакторный и многофакторный эксперимент, преимущества последнего	63

5.3 Критерии оптимальности планов	68
5.4 ПЛАНИРОВАНИЕ ЭКСПЕРИМЕНТА ПРИ ПОСТРОЕНИИ МОДЕЛИ	72
5.4.1 Линейные модели 1-го порядка	
Полный Факторный эксперимент типа $2^m$	72
5.4.2 Линейные модели 2-го порядка	79
6. ПЛАНИРОВАНИЕ ЭКСПЕРИМЕНТА ПРИ ПОИСКЕ ОПТИМАЛЬНЫХ УСЛОВИЙ	95
6.1 Метод крутого восхождения	95
6.2 Последовательное планирование эксперимента	101