ФЕДЕРАЛЬНОЕ АГЕНТСТВО ПО ОБРАЗОВАНИЮ
Государственное образовательное учреждение высшего профессионального образования
«ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

СТАТИСТИЧЕСКИЕ МЕТОДЫ КОНТРОЛЯ И УПРАВЛЕНИЯ

Рекомендовано в качестве учебно-методического пособия
Редакционно-издательским советом
Томского политехнического университета

Издательство
Томского политехнического университета
2008

УДК 519.2
ББК 22.17

Дядик В.Ф.

Статистические методы контроля и управления: учебно-методическое пособие / Дядик В.Ф., Байдали Т.А., Байдали С.А. – Томск: Изд-во Томского политехнического университета, 2008. – 73 с.

В пособии системно изложены теоретические положения, алгоритмы расчета и последовательность выполнения лабораторных практикумов по экспериментальному определению оценок статистических характеристик случайных величин, систем случайных величин, случайных процессов, которые широко применяются при разработке систем автоматического контроля и управления технологическими процессами. В пособии также представлены необходимые учебные материалы для изучения статистических методов планирования экспериментов и обработки их результатов, широко используемых для описания реальных технологических процессов.

УДК 519.2
ББК 22.17

Рецензенты

Доктор химических наук, ведущий научный сотрудник Института химии нефти СО РАН
Т.А. Сагаченко

Зам. главного прибориста Сублиматного завода Сибирского химического комбината
О.П. Савитский

© Дядик В.Ф., Байдали Т.А., Байдали С.А., 2008
© Томский политехнический университет, 2008
© Оформление. Издательство Томского политехнического университета, 2008

СОДЕРЖАНИЕ

Введение

В инженерной практике для анализа и синтеза систем автоматического контроля и управления широко и продуктивно применяются статистические методы. Они позволяют строить динамические модели промышленных объектов на основании обработки их входных и выходных сигналов в режиме нормальной эксплуатации, решать задачи определения величин и показателей контролируемого объекта, включая фильтрацию измеряемых величин от помех. Статистические методы планирования экспериментов весьма эффективны при разработке статических моделей многомерных технологических объектов. Поэтому для специальности 140306 «Электроника и автоматика физических установок» курса «Статистические методы контроля и управления» является обязательной специально дисциплиной.

Пособие состоит из следующих разделов:

− Оценка законов распределения случайных величин;

− Эмпирическая плотность распределения (гистограмма выборки) случайной величины;

− Оценивание характеристик систем случайных величин;

− Оценка автокорреляционной функции стационарного случайного процесса. Экспоненциальное сглаживание измеряемых сигналов;

− Планирование экспериментов. Планы первого порядка;

− Планирование экспериментов. Планы второго порядка.

В первых четырех разделах системно изложены теоретические положения, алгоритмы расчета и последовательность выполнения лабораторных практикумов по экспериментальному определению оценок статистических характеристик случайных величин, систем случайных величин, случайных процессов, которые широко применяются при разработке систем автоматического контроля и управления технологическими процессами. В пособии также представлены необходимые учебные материалы для изучения статистических методов планирования экспериментов и обработки из результатов, широко используемых для описания реальных технологических процессов. Учебно-методическое обеспечение дисциплины представлено библиографией из 18 источников.

1. Оценка законов распределения случайных величин

1.1. Теория

Случайной величиной называют такую величину, значения которой изменяются при повторении опытов некоторым, заранее не предсказуемым образом.

В отличие от неслучайных, детерминированных величин для случайной величины нельзя заранее точно сказать, какое конкретное значение она примет в определенных условиях, а можно только указать закон ее распределения.

Полной характеристикой случайной величины X с вероятностной точки зрения является ее закон распределения.

Закон распределения считается заданным, если

- указано множество возможных значений случайной величины;
- указан способ количественного определения вероятности попадания случайной величины в любую область множества возможных значений.

Естественной формой закона распределения случайной величины X с конечным числом значений служит таблица распределения, в которой приведены все ее возможные значения x_1, x_2; . . . x_n , и соответствующие им вероятности.

Универсальной формой закона распределения служит функция распределения – это такая функция, значения которой в точке $\boldsymbol{x}$ равно вероятности того, что при проведении испытания случайная величина окажется меньше, чем $\boldsymbol{x}$:

$$F(x) = P(X < x) \tag{1.1}$$

Основные определения

1. Выборка.
2. Вариационный ряд.
3. Статистический ряд.
4. Полигон частот выборки.
5. Диаграмма накопления частот (эмпирическая функция распределения).

Выборкой объёма “n” называется набор значений x_1, x_2; . . . x_n – случайной величины X, полученных в результате “n” опытов.

Вариационным рядом называется выборка, расположенная в порядке возрастания ее элементов.

Если выборка объёма “n” содержит r различных элементов z_1; z_2; . . . z_r, причем элемент z_i встречается m_i раз, то число m_i называют частотой элемента z_i. Сумма частот равна объему выборки $\sum_{i=1}^{r} m_i = n$,

Отношение частот m_i к объему выборки $w_i = \frac{m_i}{n}$ – называется относительной частотой элемента z_i. Очевидно, что сумма относительных частот равна единице.

Статистическим рядом называется последовательность пар (z_i, m_i); (z_i, w_i). Обычно статистический ряд записывается в виде таблицы.

Полигоном частот выборки называется графическое изображение статистического ряда. Полигоном частот выборки является ломанная с вершинами в точках: (z_i, m_i); (z_i, w_i).

Эмпирической функцией распределения называют функцию $\hat{F}_n(x) = \frac{k}{n}$. То есть, $\hat{F}_n(x)$ есть отношение **k** – числа элементов выборки, не превосходящих **x,** к **n** – объему выборки.

Функция $\hat{F}_n(x)$ служит оценкой неизвестной функции распределения F(x), т.е. $\hat{F}_n(x) \approx F(x)$

Функция $\hat{F}_n(x)$ – неубывающая, ее значения принадлежат отрезку [0 – 1].

$$\hat{F}_n(x) = \begin{cases} 0 & при\ x \le x_1 \\ F_1(x) & при\ x_1 < x \le x_2 \\ F_2(x) & при\ x_2 < x \le x_3 \\ \ldots & \\ F_{m-1}(x) & при\ x_{r-2} < x \le x_{r-1} \\ F_m(x) & при\ x > x_r \end{cases} \qquad (1.2)$$

График эмпирической функции представляет собой ломаную линию. В промежутках между соседними членами вариационного ряда функция $\hat{F}_n(x)$ сохраняет постоянное значение. При переходе через точки оси x, равные членам выборки, $\hat{F}_n(x)$ претерпевает разрыв, скачком возрастая на величину $\frac{1}{n}$, если какие-либо из x_i совпадают λ

раз, то на величину $\frac{\lambda}{n}$. Таким образом график имеет ступенчатый вид (скачки в точках выборки).

1.2. Алгоритм расчета

Оценку законов распределения случайных величин выполнить с помощью электронной таблицы ***Excel***.

1. Записывается выборка в виде вариационного ряда.
2. Записывается статистический ряд в виде таблицы:

Элемент выборки	z_i						
Частота элемента	m_i						
Относительная частота элемента	w_i						

где m_i - частота элемента z_i,

$n = \sum_{i=1}^{r} m_i$ - объём выборки,

$w_i = \dfrac{m_i}{n}$ - относительная частота элемента z_i.

$i = 1,2,\ldots r$.

2.1. Строиться полигон частот выборки (графическое изображение статистического ряда):

ломанная с вершинами в точках (z_i, m_i) – полигон частот;

ломанная с вершинами в точках (z_i, w_i).– полигон относительных частот;

3. Записывается эмпирическая функция распределения в виде таблицы:

x_i						
k						
$\hat{F}_n(x) = \dfrac{k}{n}$						

где $x_i = z_i$, n – объем выборки, k – число элементов выборки, не превосходящих x_i, $i = 1,2,\ldots r$.

3.1. Строиться диаграмма накопления частот (эмпирическая функция распределения) $-\hat{F}_n(x)$, которая служит оценкой неизвестной функции распределения F(x), т. е. $\hat{F}_n(x) \approx F(x)$.

1.3. Последовательность выполнения задания

1. Выписать в рабочую тетрадь все основные определения и формулы для оценки законов распределения случайной величины.
2. Выписать из таблицы 1 согласно номера варианта выборку случайной величины X.
3. Выполнить задание.

1.4. Задание

1. Записать выборку в виде вариационного ряда.
2. Образовать статистический ряд в виде таблицы. Построить полигон частот выборки (графическое изображение статистического ряда).
3. Записать эмпирическую функцию распределения в виде таблицы и в аналитическом виде. Построить диаграмму накопления частот – эмпирическую функцию распределения.

2. Эмпирическая плотность распределения (гистограмма выборки) случайной величины

2.1. Теория

Закон распределения, выраженный в форме функции распределения или плотности распределения, дает исчерпывающую характеристику случайной величины X с вероятностной точки зрения.

Для описания непрерывной случайной величины X наиболее часто употребляется плотность распределения случайной величины f(x):

$$f(x) = F^{'}(x) = \lim_{\Delta x \to 0} \frac{P[x < X \le x + \Delta x]}{\Delta x} \qquad (2.1)$$

Эмпирическим аналогом плотности распределения f(x) является гистограмма $\hat{f}\left(x_n\right)$.

Расчет статистического распределения выборки

Для оценки f(x) по дискретной выборке (полученной в результате квантования непрерывной величины) (x_1; x_2; . . ., x_n) строят гистограмму выборки.

Гистограмма строится по группированным данным. Находят предварительное количество интервалов группировки на которое должна быть разбита область значений X. Это количество k определяют с помощью оценочной формулы (формула Стэрджесса):

$$\mathrm{k} = 1 + 3{,}2 \lg \mathrm{n},$$

где n – объём выборки.

Найденное значение k округляется до ближайшего целого нечетного числа.

Длина интервала определяется по формуле:

$$h = \Delta x = \frac{x_{\max} - x_{\min}}{k}.$$

Если h < 1 , то h = 1, если h > 1, то округляют число до большего ближайшего значения первой значащей цифры после запятой.

Например: h = 0,890 (1.0), h = 1,529 (1,6).

Середину области изменения выборки (центр гистограммы) находят по формуле: $C = \frac{x_{\max} + x_{\min}}{2}$.

За счет округления шага до большего ближайшего значения границы области значений X изменяются. В связи с этим находят:

– начало гистограммы по формуле: $a = C - \frac{k}{2} * h$.

– граничные значения каждого i – ого интервала: $[x_i \div x_{i+1} = x_i + h], \quad i = 1,2,...k,$ принимая начало первого интервала $x_1 = a$.

Подсчитывают число ν_i (ν_1; ν_2; . . .; ν_k) элементов выборки, попавших в интервал h_i.

ν_i равно числу элементов выборки вариационного ряда, для которых справедливо неравенство:

$$x_i < z \le x_i + \Delta x$$

т.е. граничные значения x_{i-1} и x_i относят к i (к правому интервалу "i").

Оценка плотности распределения случайной величины x (плотности вероятности) в точке x_i – центре интервала h_i вычисляется по формуле:

$$\hat{f}(x_i) \approx \frac{\nu_i}{n \cdot h_i} \quad (i = 1; 2; \ldots; k)$$

Проверка гипотезы о нормальном распределении генеральной совокупности X по критерию Пирсона на основании заданного эмпирического распределения

Во многих практических задачах точный закон распределения исследуемого признака *X* генеральной совокупности неизвестен. В этом случае необходимо проверить гипотезу о предполагаемом законе распределения. Выдвигаются нулевая гипотеза H_0 и ей конкурирующая H_1.

H_0: признак X имеет нормальный закон распределения.

H_1: признак X имеет закон распределения, отличный от нормального.

Нулевая гипотеза проверяется с помощью ***критерия согласия***.

Критерий χ^2 ("хи-квадрат") ***Пирсона*** – наиболее часто употребляемый критерий, может применяться для проверки гипотезы о любом законе распределения. Независимо от того, какое распределение имеет *X*, распределение случайной величины χ^2:

$$\chi^2 = \sum_{i=1}^{k} \frac{(\nu_i - \nu_i')^2}{\nu_i'} ,$$

где ν_i – эмпирические частоты, ν_i' – теоретические частоты; при $n \to \infty$ стремится к– χ^2 распределению с k степенями свободы.

Теоретические частоты определяются, исходя из предположения о законе распределения генеральной совокупности, в данном случае о нормальном законе. Так как $p_i = \frac{v_i}{n}$, где p_i – теоретическая вероятность, то $v_i' = n \cdot p_i$.

Для дискретного ряда:

$$p_i = \frac{h}{\sigma^*_{\text{в}}} \cdot \varphi(u_i),$$

где $u_i = \frac{x_i - m_x^*}{\sigma_{\text{в}}}$, $\varphi(u_i) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{u_i^2}{2}}$ – дифференциальная функция нормированного нормального распределения,

$h = x_i - x_{i-1}$ – шаг,

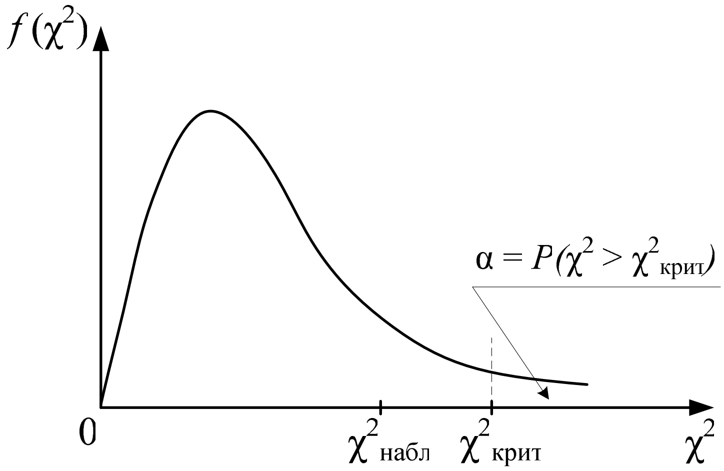
m^* – выборочное математическое ожидание вычисляемое по формуле:

$$m^* = \frac{\sum_{i=1}^{n} x_i}{n},$$

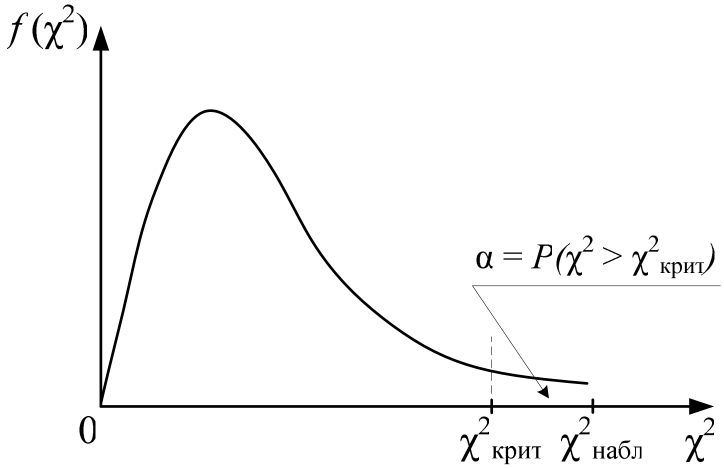
$\sigma^*_{\text{в}}$ – выборочное среднее - квадратичное отклонение, вычисляемое по формуле:

$$\sigma_B^* = \sqrt{\frac{\sum_{i=1}^{n} (x_i - m_x^*)^2}{n-1}}.$$

Рассчитав теоретические частоты, находят $\chi^2_{\text{набл}}$. Из таблицы *критических точек* распределения χ^2 (Приложение А. Таблица 2) по заданному уровню значимости α=0,05 и числу степеней свободы l находят $\chi^2_{\text{крит}}(\alpha, l)$ – границу правосторонней критической области (см. рис. 2.1). Здесь l =k –r–1 , где k – число различных дискретных значений x_i дискретного или число интервалов (x_{i-1} – x_i), r – число параметров предполагаемого закона распределения. Нормальное распределение определяется двумя параметрами: математическим ожиданием и средним квадратичным отклонением σ. Так как оба параметра оцениваются по выборке, то r = 2, отсюда l = k – 2 – 1. Затем сравнивают $\chi^2_{\text{набл}}$ и $\chi^2_{\text{крит}}(\alpha, l)$ и делают вывод.



а)



б)

Рис. 2.1

При формулировке вывода руководствуются следующим правилом:

- если наблюдаемое значение критерия $\chi^2_{набл}$ попало в область принятия гипотезы ($\chi^2_{набл} < \chi^2_{крит}(\alpha, l)$) как показано на рис. 2.1 а), то нет оснований отвергать нулевую гипотезу, по данным наблюдения признак X имеет нормальный закон распределения, расхождение между эмпирическими и теоретическими частотами случайное;
- если наблюдаемое значение критерия $\chi^2_{набл}$ попало в критическую область ($\chi^2_{набл} > \chi^2_{крит}(\alpha, l)$) как показано на рис. 2.1 б), то нулевая гипотеза отвергается, справедлива конкурирующая гипотеза, то есть признак X имеет закон распределения, отличный от нормального, расхождение между эмпирическими и теоретическими частотами значимо.

Построение гистограммы выборки и теоретической нормальной кривой

Для построения гистограммы в прямоугольной системе координат на оси абсцисс откладываем интервалы, а на каждом из интервалов h_i, как на основании, строим прямоугольник с основанием h_i и высотой $\hat{f}(x_i) \approx \frac{\nu_i}{n \cdot h_i}$. Полученная таким образом фигура и называется гистограммой выборки или эмпирической плотностью распределения.

На этом же графике строим гипотетическую (предположительную) теоретическую нормальную кривую.

Для этого берем середины интервалов x_i и <u>*выравнивающие частоты* ν_i</u> наносим точки с координатами $\left(x_i, \frac{\nu'_i}{nh}\right)$, соединяя их плавной линией, получаем искомую кривую.

2.2. Алгоритм расчета

Если X – непрерывная случайная величина с неизвестной плотностью распределения (плотностью вероятности) - f(x), то для оценки f(x) по дискретной выборке, полученной в результате квантования непрерывной величины (x_1; x_2; . . ., x_n) разбивают область значений X на интервалы h_i (i = 1; 2; . . .; k), где k–количество интервалов.

1. Расчет статистического распределения выборки.

1.1. Подсчитывается количество интервалов по формуле:
$k = 1 + 3{,}2 \lg n$, где n – объём выборки.

Полученное значение k – округляют до ближайшего целого нечетного числа.

1.2. Подсчитывается длина каждого интервала:
$h = \Delta x = \frac{x_{\max} - x_{\min}}{k}$ (с учетом округления)

1.3. Находится центр гистограммы: C=$\frac{x_{\max} + x_{\min}}{2}$.

1.4. Находится начало гистограммы $a = C - \frac{k}{2} * h$.

1.5. Определяются граничные значения каждого i-ого интервала: $[x_i \div x_{i+1} = x_i + h], \quad i = 1,2,...k,$ принимая начало первого интервала $x_1 = a$.

1.6. Подсчитывается число элементов выборки ν_i (ν_1; ν_2; . . .; ν_k), попавших в интервалы h_i, при этом граничные значения x_{i-1} и x_i относят к i (к правому интервалу "i").

1.7. По формуле $\hat{f}\left(x_i^*\right) \approx \frac{\nu_i}{n \cdot h}$ (i = 1; 2; . . .; k) вычисляют ординаты оценок плотности вероятности в точках x_i^* – центрах интервалов h_i.

1.8. Результат представляется в виде таблицы:

№	Интервал	Элементы	Число элементов ν_i	$\hat{f}\left(x_i^*\right) \approx \frac{\nu_i}{n \cdot h}$
1			ν_1	$f(x_1^*)$
2			ν_2	$f(x_2^*)$
…			…	…
k			ν_k	$f(x_k^*)$

2. Проверка гипотезы о нормальном распределении генеральной совокупности X, при заданном уровне значимости α, по критерию Пирсона на основании заданного эмпирического распределения.

2.1. Вычисляется выборочное математическое ожидание:

$$m^* = \frac{\sum_{i=1}^{n} x_i}{n}.$$

2.2. Вычисляется выборочное средне - квадратичное отклонение:

$$\sigma_B^* = \sqrt{\frac{\sum_{i=1}^{n} (x_i - m_x^*)^2}{n-1}}.$$

2.3. Вычисляются теоретические частоты:

$$\nu_i' = \frac{n \cdot h}{\sigma_B^*} \cdot \varphi(u_i), \text{ где } \varphi(u_i) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u_i^2}{2}}, \cdots u_i = \frac{x_i - m_x^*}{\sigma_B^*}.$$

2.4. Результат представляется в виде таблицы:

i	x_i	u_i	$\varphi(u_i)$	ν_i'
1				
2				
k				

2.5. Составляется расчетная таблица для сравнения эмпирических и теоретических частот с помощью критерия Пирсона.

i	ν_i	ν_I'	$\nu_I - \nu_I'$	$(\nu_I - \nu_I')^2$	$\frac{(\nu_i - \nu_i')^2}{\nu_i'}$
1					
2					
k					
Σ					$\chi^2_{\text{набл}}$

2.6. По таблице находится наблюдаемое значение критерия:

$$\chi^2_{набл} = \sum_{i=1}^{k} \frac{(\nu_i - \nu_i')^2}{\nu_i'}.$$

2.7. По таблице χ^2 – распределения, по заданному уровню значимости и числу степеней свободы находится критическая точка $\chi^2_{\text{кр}}(\alpha;l)$ правосторонней критической области (Приложение А. Таблица 2).

2.8. Сравниваются значения $\chi^2_{\text{набл}}$ и $\chi^2_{\text{крит}}(\alpha, l)$ и делается вывод.

3. Строиться гистограмма выборки (эмпирическая плотность распределения) и теоретическая нормальная кривая (если гипотеза о нормальном распределении генеральной совокупности из которой взята данная выборка принимается).

2.3. Последовательность выполнения задания

1. Выписать в рабочую тетрадь все основные формулы для расчета статистического распределения выборки и проверки гипотезы о соответствии распределения случайной величины нормальному закону распределения.
2. Выписать согласно номера варианта дискретную выборку полученную в результате квантования непрерывного сигнала.
3. Выполнить задание.

2.4. Задание

1. Произвести расчет статистического распределения выборки.
2. Результат представить в виде таблицы

Таблица 1

i	Интервал	Элементы	ν_i	$f(x_i)$

3. При заданном уровне значимости α проверить гипотезу о нормальном распределении генеральной совокупности X, выборка из которой Вам задана, по критерию Пирсона на основании заданного эмпирического распределения.
4. Результат представить в виде таблиц 2 и 3:

Таблица 2

i	x_i	u_i	$\varphi(u_i)$	ν'_i

Таблица 3

i	ν_i	ν'_i	$\nu_i - \nu'_i$	$(\nu_i - \nu'_i)^2$	$\frac{(\nu_i - \nu'_i)^2}{\nu'_i}$

5. Построить гистограмму выборки – эмпирическую плотность распределения – $f(x_i)$ и теоретическую нормальную кривую, если гипотеза о нормальном распределении генеральной совокупности из которой взята данная выборка принимается.

3. Оценивание характеристик систем случайных величин

3.1. Теория

При неизвестных законах распределения для приближенного описания как случайных величин так и систем случайных величин используют отдельные числовые характеристики, определяющие основные черты закона распределения.

Числовые значения характеристик случайных величин и систем случайных величин, которые относятся к генеральной совокупности, называются параметрами – ***a*.**

Значение параметра, вычисленное по ограниченному объему экспериментальных данных, является случайной величиной, т. е. значение такой величины от выборки к выборке может меняться заранее не предвиденным образом. Следовательно, в результате обработки экспериментальных данных определяется не значение параметра, а только лишь его приближенное значение – статистическая оценка параметра. Оценка есть неточное, неполное, приближенное отображение параметра. Получить статистическую оценку параметра теоретического распределения означает найти функцию от имеющихся результатов наблюдения, которая и даст приближенное значение искомого параметра.

Численные значения характеристик случайных величин, рассчитанные на основании выборки, называют *оценками* этих параметров или *статистиками a**.

Оценка будет тем точнее, чем отчетливее случайный характер выборки и чем больше ее объем. Предполагается, что все наблюдения независимы.

Различают два вида оценок – точечные и интервальные.

Точечными называют такие оценки, которые характеризуются одним числом, рассчитанным по определенной формуле, т.е. находится число («оценка»), способное в некотором смысле, заменить параметр. При малых объемах выборки точечные оценки могут значительно отличаться от истинных значений параметров, поэтому их применяют при большом объеме выборки.

Интервальные оценки задаются двумя числами, определяющими вероятный диапазон возможного значения параметра, т.е. указывается интервал, накрывающий параметр с заданной наперед вероятностью.. Эти оценки применяются для малых и для больших выборок.

Вероятностные свойства произвольной оценки $a^*=a^*(x_1, x_2,...,x_n)$ параметра a можно описать с помощью функции распределения оценки $F(a^*)$ или ее характеристик m_{a^*}, σ_{a^*} .

Для оценивания одного и того же параметра можно использовать в принципе различные оценки. Чтобы выбрать наилучшую из них, необходимо сформулировать некоторые требования к свойствам оценок, желательным с точки зрения практики.

Оценки параметров (“статистики”) должны удовлетворять следующим четырем условиям:

1. Оценки должны быть “*несмещенными*” (unbiased, verzerrungsfrei)

$$M[a^*] = a$$

т.е. при очень большом числе испытаний с одинаковыми выборками среднее значение оценок должно стремиться к истинному значению параметра генеральной совокупности.

2. Оценки должны быть “*состоятельными*” (согласованными) (consistent)

$$\lim_{n\to\infty} a^* = a$$

т.е. с ростом “n” они должны стремиться к соответствующему параметру генеральной совокупности.

3. Оценки должны быть “*эффективными*” (efficient)

$$D[a^*] = \min$$

$n_1=n_2=...=n_k$ - т.е. для выборок равного объема они должны иметь минимальное рассеяние (дисперсию).

Как правило среднеквадратичное отклонение оценки по абсолютной величине и отнесенное к математическому ожиданию уменьшается при увеличении объема выборки.

4. Оценки должны быть “*достаточными*” (sufficient), т.е. не должна возникать необходимость в дополнительной информации об оцениваемом параметре.

Названия состоятельная (consistens), эффективная (efficient); достаточная (sufficient) введены Фишером Р.А. (1925г.).

Для оценивания параметров по выборочным данным разработаны многочисленные методы. Случайная выборка важна, так как только она позволяет распространить выводы на всю генеральную совокупность.

Степень близости статистической оценки a^* к соответствующему параметру, a удобно характеризовать с помощью доверительного интервала.

Поскольку оценка параметра (статистическая надежность) , вычисленная по ограниченной выборке, является случайной величиной, необходимо оценить ее *точность* и *надежность*.

Чтобы дать представление о точности и надежности оценки a^* параметра a в математической статистике пользуются так называемыми доверительными интервалами и доверительными вероятностями. Параметры генеральной совокупности могут быть известны только на основании теоретических соображений.

Понятие “*доверительный интервал*” было введено Дж. Нейманом и Е.С. Пирсоном. (1950г.) Так называют вычисленный по выборочным значениям (x_1, x_2, ... ,x_n) выборки объема “n” интервал J_β, который с заданной вероятностью, доверительной вероятностью β, накрывает истинное, но не известное нам значение параметра a.

В качестве доверительной вероятности в инженерной практике обычно принимают β=95%; эта вероятность говорит о том, что при частых применениях данного метода вычисленный доверительный интервал примерно в 95% случаев будет накрывать параметр a.

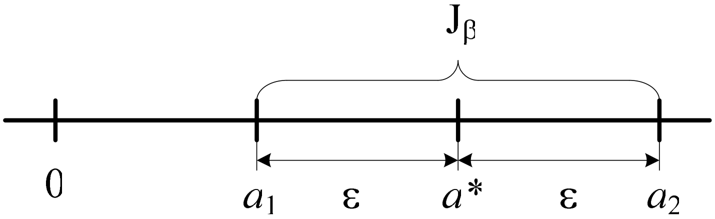


Рис. 3.1

Равенство $P[(a^* - \varepsilon) \le a \le (a^* + \varepsilon)] = \beta$ означает, что с вероятностью β неизвестное значение параметра “a” попадает в интервал $J_\beta = [(a^* - \varepsilon);(a^* + \varepsilon)]$.

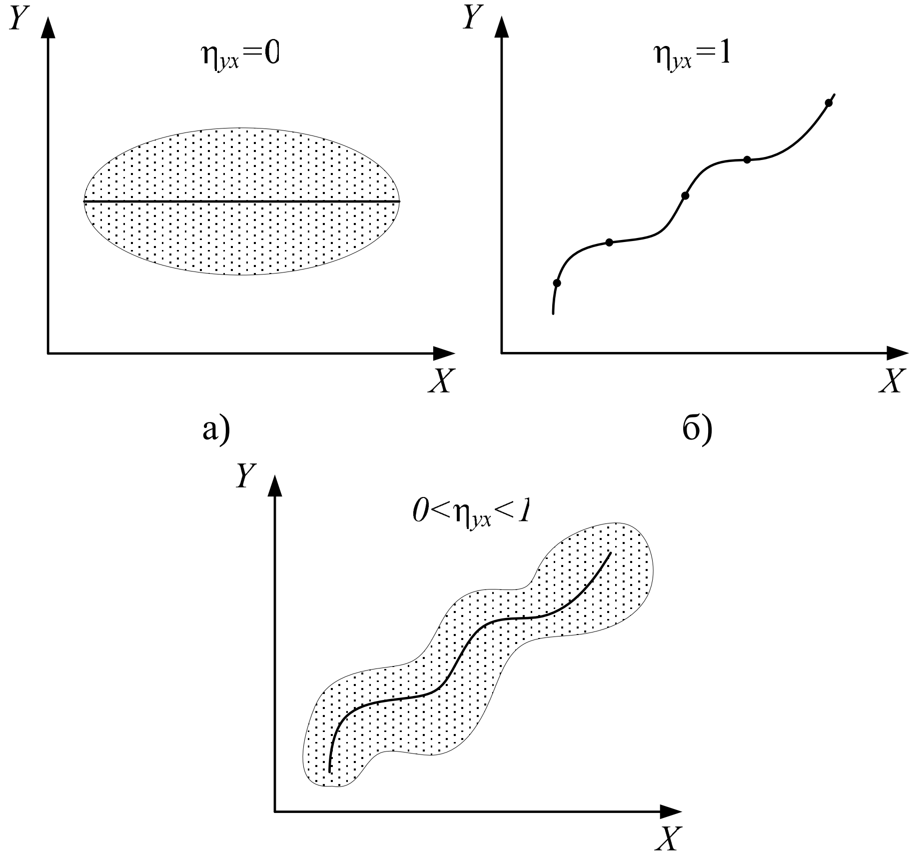
Вероятность β принято называть доверительной вероятностью - статистической надежностью, а интервал J_β – доверительным интервалом.

Границы интервала J_β $a_1 = a^* - \varepsilon$ и $a_2 = a^* + \varepsilon$ называются доверительными границами.

Корреляционное отношение и коэффициенты корреляции

Корреляция двух величин X и Y характеризует степень (тесноту) связи между этими величинами. Она определяется корреляционным отношением. η_{yx} , являющимся безразмерной характеристикой связи.

На рис. 3.2 показаны различные виды корреляционных полей между величинами X и Y. Корреляционное поле представляет собой совокупность точек, полученных в результате измерений пар значений X и Y. Обычно при наличии системы из двух величин одну из них принимают за функцию и ее значения откладывают по оси ординат (Y), другую - за аргумент, откладывая ее значения по оси абсцисс (X).



в)

Рис. 3.2

На рис. 3.2 *а*. не наблюдается никакой зависимости между величинами X и Y. В этом случае говорят об отсутствии корреляции между величинами X и Y. Корреляционное отношение $\eta_{yx}=0$.

На рис. 3.2 *б*. все точки измеренных пар значений X и Y легли на одну кривую. Полученная зависимость считается функциональной, для нее $\eta_{yx}=1$.

Наконец на рис. 3.2 *в*. показана стохастическая зависимость между величинами X и Y, при которой корреляционное отношение принимает промежуточное значение $0< \eta_{yx}<1$.

В практически важном случае линейной зависимости между величинами X и Y корреляционное отношение вырождается в коэффициент корреляции r_{yx}, характеризующий степень линейной связи между величинами X и Y и, так же как η_{yx}, являющейся безразмерной величиной.

На рис. 3.3 показаны различные случаи линейной связи между величинами X и Y от полной функциональной зависимости (рис.3.3 *а* и *е*), до линейной независимости величин (рис.3.3 *в* и *г*)

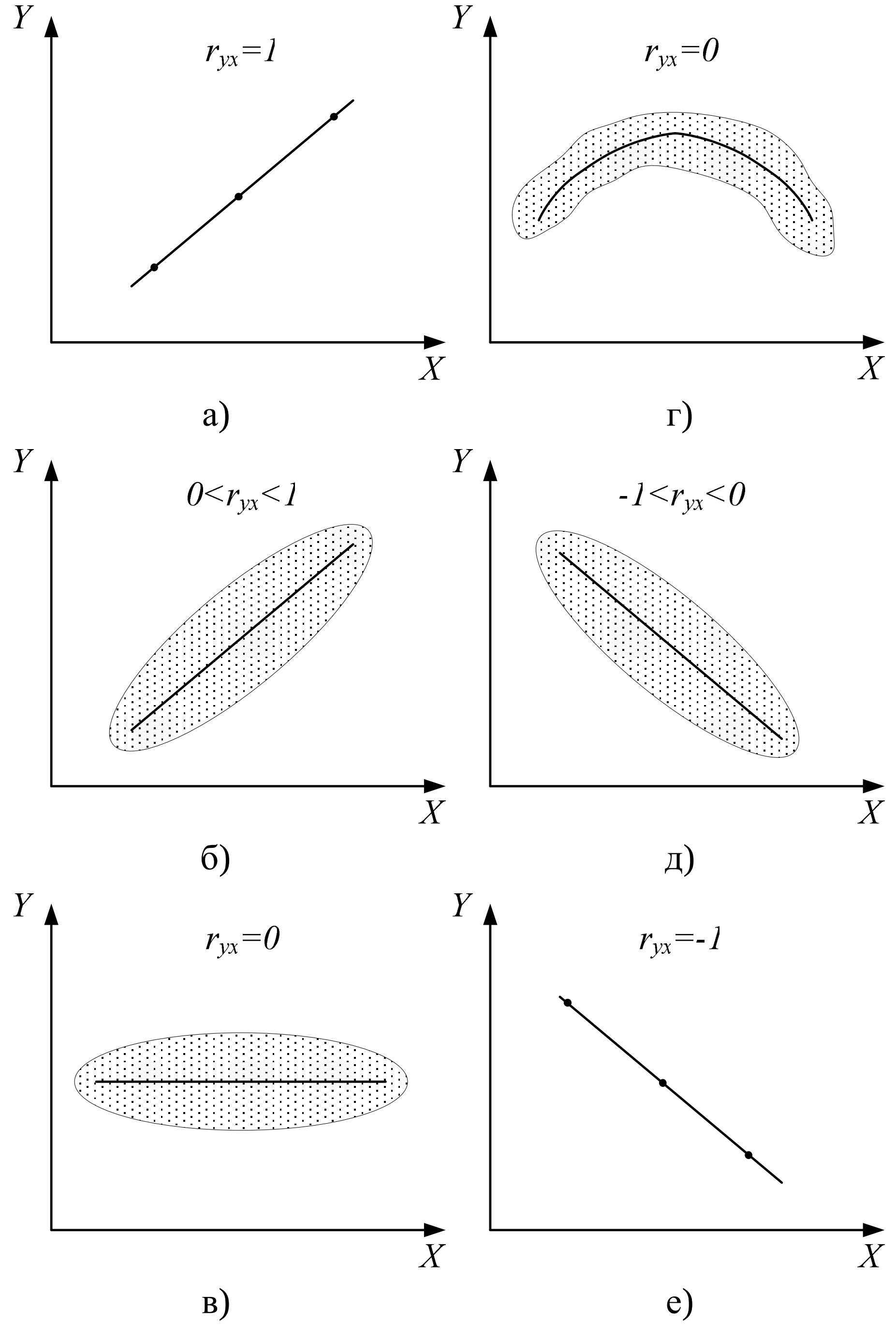


Рис. 3.3

Последовательность оценки характеристик системы двух случайных величин. [X; Y].

1. Строится корреляционное поле по исходной выборке объема “n”.

2. Оцениваются математические ожидания Х и У.

2.1. Точечные оценки рассчитываются по формулам:

$$m_x^* = \frac{\sum_1^n x_i}{n}; \qquad m_y^* = \frac{\sum_1^n y_i}{n};$$

2.2. Интервальные оценки определяются по соотношениям:

$$m_x^* \pm \frac{t\sigma_x^*}{\sqrt{n}}; \qquad m_y^* \pm \frac{t\sigma_y^*}{\sqrt{n}},$$

где $t_{(n-1);\ \alpha}$ - табличное значение коэффициент распределения Стьюдента [(n-1) – число степеней свободы ; α - уровень значимости, при доверительной вероятности β=0,95 , α=0,05].

3. Оцениваются дисперсии и среднеквадратичные отклонения Х и У.

3.1. Точечные оценки рассчитываются по формулам:

$$D_x^* = \frac{Q_x}{n-1}; \qquad \sigma_x^* = \sqrt{\frac{Q_x}{n-1}}$$

$$D_y^* = \frac{Q_y}{n-1}; \qquad \sigma_y^* = \sqrt{\frac{Q_y}{n-1}}$$

$$Q_x = \sum_{i=1}^{n}\left(x_i - m_x^*\right)^2 = \sum_{i=1}^{n} x_i^2 - \frac{\left(\sum x_i\right)^2}{n}$$

$$Q_y = \sum_{i=1}^{n}\left(y_i - m_y^*\right)^2 = \sum_{i=1}^{n} y_i^2 - \frac{\left(\sum y_i\right)^2}{n}$$

3.2. Интервальные оценки определяются по соотношениям:

$$\frac{D_x^* \cdot (n-1)}{\chi^2_{(n-1);\frac{\alpha}{2}}} < D_x < \frac{D_x^* \cdot (n-1)}{\chi^2_{(n-1);\left(1-\frac{\alpha}{2}\right)}}; \qquad \sqrt{\frac{D_x^* \cdot (n-1)}{\chi^2_{(n-1);\frac{\alpha}{2}}}} < \sigma_x < \sqrt{\frac{D_x^* \cdot (n-1)}{\chi^2_{(n-1);\left(1-\frac{\alpha}{2}\right)}}},$$

где $\chi^2_{(n-1);\frac{\alpha}{2}}$ и $\chi^2_{(n-1);(1-\frac{\alpha}{2})}$ – табличное значение *хи* - квадрат- распределения.

4. Оценивается коэффициент корреляции

4.1. Точечная оценка вычисляется по формуле:

$$r_{yx}^{*}=\frac{Q_{xy}}{\sqrt{Q_x\cdot Q_y}}$$

$$Q_{xy}=\sum_{1}^{n}x_i y_i-\frac{\left(\sum x_i\right)\cdot\left(\sum y_i\right)}{n}$$

4.2. Интервальная оценка находится по номограмме. (Л.Закс Статистическое оценивание, Статистика. 1976, стр.390).

5. По характеру корреляционного поля принимается гипотеза о применимости линейного уравнения регрессии:

$$\hat{y}=a_{yx}^{*}+b_{yx}^{*}\cdot x$$

5.1. Вычисляются точечные оценки коэффициентов линейного уравнения регрессии по формулам:

$$b_{yx}^{*}=\frac{Q_{xy}}{Q_x};\quad a_{yx}^{*}=\frac{\sum_{1}^{n}y_i-b_{yx}^{*}\cdot\sum_{1}^{n}x_i}{n}$$

5.2. Определяются интервальные оценки коэффициентов линейного уравнения регрессии:

$b_{yx}^{*}\pm t\sigma_{b_{yx}}^{*};\quad a_{yx}^{*}\pm t\sigma_{a_{yx}}^{*}$ (при (n-2) степенях свободы),

где $t_{(n-2);\ \alpha}$ - табличное значение коэффициент распределения Стьюдента.

При этом среднеквадратичные отклонения коэффициентов уравнения регрессии вычисляются по формулам:

$$\sigma_{b_{yx}}^{*}=\frac{\sigma_{y/x}^{*}}{\sqrt{Q_x}};\qquad \sigma_{a_{yx}}^{*}=\sigma_{b_{yx}}^{*}\sqrt{\frac{\sum_{i=1}^{n}x_i^2}{n}}$$

6. Оцениваются дисперсия и среднеквадратичная ошибка предсказания опытных данных найденным уравнением регрессии.

6.1. Точечные оценки рассчитываются по формулам:

$$D_{y/x}^{*}=\frac{Q_y-\frac{\left(Q_{xy}\right)^2}{Q_x}}{n-2};\qquad \sigma_{y/x}^{*}=\sqrt{\frac{Q_y-\frac{\left(Q_{xy}\right)^2}{Q_x}}{n-2}}$$

6.2. Интервальные оценки определяются по соотношениям:

$$\frac{D^*_{y/x}\cdot(n-2)}{\chi^2_{(n-2);\frac{\alpha}{2}}}<D_{y/x}<\frac{D^*_{y/x}\cdot(n-2)}{\chi^2_{(n-2);\left(1-\frac{\alpha}{2}\right)}};\quad \sqrt{\frac{D^*_{y/x}\cdot(n-2)}{\chi^2_{(n-2);\frac{\alpha}{2}}}}<\sigma_{y/x}<\sqrt{\frac{D^*_{y/x}\cdot(n-2)}{\chi^2_{(n-2);\left(1-\frac{\alpha}{2}\right)}}},$$

где $\chi^2_{(n-2);\frac{\alpha}{2}}$ и $\chi^2_{(n-2);(1-\frac{\alpha}{2})}$ – табличное значение *хи* - квадрат- распределения.

7. На корреляционным поле строится график уравнения регрессии с доверительным интервалом предсказания.

8. Все найденные интервальные и точечные оценки числовых характеристик заданной системы двух случайных величин изображаются на соответствующих числовых осях.

3.2. Алгоритм расчета

Точечные оценки.

Точечные оценки числовых характеристик системы 2-х случайных величин рассчитать на ЭВМ с помощью программы - Оценка.exe. (L:\Study\СМК\Лаб_работа_3\Оценка.exe).

Интервальные оценки.

Найти интервальные оценки всех числовых характеристик системы случайных величин.

1. Построить корреляционное поле по исходной выборке объема “n.

2. Интервальные оценки математических ожиданий X и У определяются по соотношениям:

$$m_x^* \pm \frac{t\sigma_x^*}{\sqrt{n}}; \qquad m_y^* \pm \frac{t\sigma_y^*}{\sqrt{n}},$$

где $t_{(n-1);\ \alpha}$ - табличное значение коэффициент распределения Стьюдента [(n-1) – число степеней свободы ; α - уровень значимости, при доверительной вероятности β=0,95 , α=0,05].(Приложение 1. Таблица 1)

3. Интервальные оценки дисперсий и среднеквадратичные отклонения X и У определяются по соотношениям:

$$\frac{D_x^* \cdot (n-1)}{\chi^2{}_{(n-1);\frac{\alpha}{2}}} < D_x < \frac{D_x^* \cdot (n-1)}{\chi^2{}_{(n-1);\left(1-\frac{\alpha}{2}\right)}}; \quad \sqrt{\frac{D_x^* \cdot (n-1)}{\chi^2{}_{(n-1);\frac{\alpha}{2}}}} < \sigma_x < \sqrt{\frac{D_x^* \cdot (n-1)}{\chi^2{}_{(n-1);\left(1-\frac{\alpha}{2}\right)}}},$$

где $\chi^2_{(n-1);\frac{\alpha}{2}}$ и $\chi^2_{(n-1);(1-\frac{\alpha}{2})}$ – табличное значение *хи* - квадрат- распределения. (Приложение 1. Таблица 2).

4. Интервальная оценка коэффициента корреляции находится по номограмме. (Приложение 2.).

5. Интервальные оценки коэффициентов линейного уравнения регрессии определяются по соотношениям:

$b_{yx}^* \pm t\sigma_{b_{yx}}^*$; $\quad a_{yx}^* \pm t\sigma_{a_{yx}}^*$ (при (n-2) степенях свободы),

где $t_{(n-2);\ \alpha}$ - табличное значение коэффициент распределения Стьюдента. (Приложение 1. Таблица 1).

6. Интервальные оценки дисперсии и среднеквадратичная ошибка предсказания опытных данных определяются по соотношениям:

$$\frac{D^*_{y/x}\cdot(n-2)}{\chi^2_{(n-2);\frac{\alpha}{2}}} < D_{y/x} < \frac{D^*_{y/x}\cdot(n-2)}{\chi^2_{(n-2);\left(1-\frac{\alpha}{2}\right)}}; \quad \sqrt{\frac{D^*_{y/x}\cdot(n-2)}{\chi^2_{(n-2);\frac{\alpha}{2}}}} < \sigma_{y/x} < \sqrt{\frac{D^*_{y/x}\cdot(n-2)}{\chi^2_{(n-2);\left(1-\frac{\alpha}{2}\right)}}},$$

где $\chi^2_{(n-2);\frac{\alpha}{2}}$ и $\chi^2_{(n-2);(1-\frac{\alpha}{2})}$ – табличное значение *хи* - квадрат- распределения. (Приложение А. Таблица 2).

7. На корреляционным поле строится график уравнения регрессии с доверительным интервалом предсказания.

8. Все найденные интервальные и точечные оценки числовых характеристик заданной системы двух случайных величин изображаются на соответствующих числовых осях.

3.3. Последовательность выполнения задания

1. Выписать в рабочую тетрадь все необходимые для оценивания числовых характеристик систем случайных величин формулы.
2. Выписать из таблицы 2, согласно номера варианта, значения входной Х и выходной У величин изучаемого объекта.
3. Создать текстовый файл исходных данных ***имя.txt,*** состоящий из ***n*** пар значений Х; У.

Образец:

356.5 90.5
361.5 91.5
… …
357.6 91.5***размещение курсора (конец файла)***

4. Выполнить задание.

3.4. Задание

На нефтеперерабатывающем заводе для исследования зависимости процента выкипания продукта от температуры выкипания в атмосферно–вакуумной трубчатке было произведено n независимых попарных измерений этих показателей.

Х – температура выкипания;

У – процент выкипания.

Заданы значения входной и выходной величин изучаемого объекта

1. Рассчитать на ЭВМ с помощью программы «Оценка.exe» точечные оценки числовых характеристик системы 2-х случайных величин. (L:\Study\СМК\Лаб_работа_3\оценка.exe).
2. Построить корреляционное поле.
3. Найти интервальные оценки всех числовых характеристик системы случайных величин:

- оценить математические ожидания Х и У - $m_x^*; m_y^*$;
- оценить дисперсии и среднеквадратичные отклонения Х и У - $D_x^*; D_y^*; \sigma_x^*; \sigma_y^*$;

- оценить коэффициент корреляции - r^*_{yx};
- оценить коэффициенты уравнения регрессии $\hat{y} = a^*_{yx} + b^*_{yx} x$ - b^*_{yx}; a^*_{yx};
- оценить дисперсию и среднеквадратичную ошибку предсказания опытных данных - $D^*_{y/x}; \sigma^*_{y/x}$.
- На корреляционным поле, согласно полученному уравнению, построить график уравнения регрессии с доверительным интервалом предсказания.
- Изобразить все найденные интервальные и точечные оценки числовых характеристик заданной системы двух случайных величин на соответствующих числовых осях.

Примечание:

Табличное значение коэффициента распределения Стьюдента (Приложение А. Таблица 1).

Табличные значения *хи* – квадрат – распределения. (Приложение А. Таблица 2).

Доверительные границы для коэффициента корреляции (Приложение Б).

4. Оценка автокорреляционной функции стационарного случайного процесса. Экспоненциальное сглаживание измеряемых сигналов

4.1. Теория

Понятие о стационарном случайном процессе

Исключительно большое значение для практики имеют случайные процессы, протекающие во времени приблизительно однозначно и имеющие вид непрерывных случайных колебаний вокруг некоторого среднего значения, причем ни средняя амплитуда, ни характер этих колебаний не обнаруживают существенных изменений с течением времени. Такие случайные процессы называются ***стационарными.***

Случайный процесс X(t) называется *стационарным*, если все его вероятностные характеристики не зависят от t (точнее, не меняются при любом сдвиге по оси t).

Сформулируем определение стационарного случайного процесса в терминах вероятностных характеристик.

Для стационарного случайного процесса:

- математическое ожидание постоянно

$$m_x(t) = m_x = const$$

- дисперсия постоянна

$$D_x(t) = D_x = const$$

- автокорреляционная функция стационарного случайного процесса X(t) есть функция одного аргумента τ

$$K_x(t_1,t_1+\tau) = K_x(\tau)$$

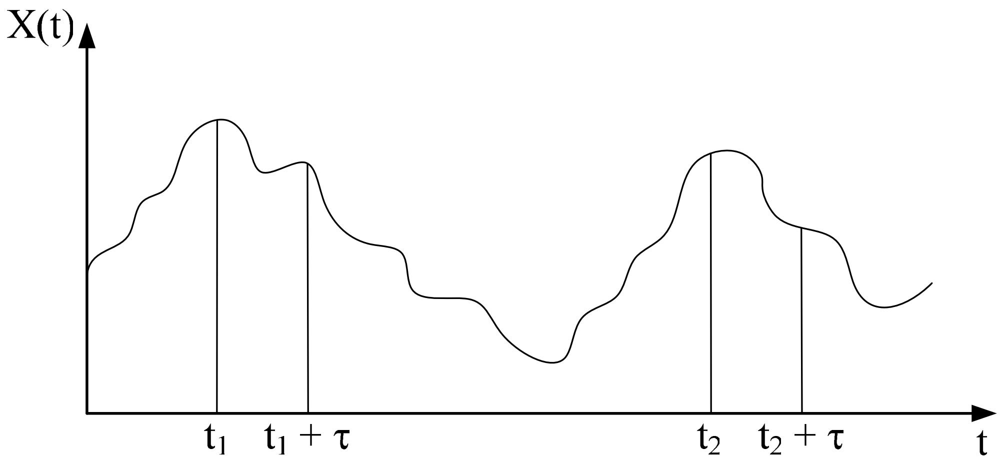


Рис.4.1

Если случайный процесс стационарен, то значение автокорреляционной функции не зависит от того, где именно на оси аргумента t взят участок τ, а только от длины участка.

Оценка автокорреляционной функции эргодического стационарного случайного процесса производится по формуле:

$$K_x^*(\tau) = \frac{1}{T-\tau}\int_0^{T-\tau}\left[x(t)-m_x^*\right]\cdot\left[x(t+\tau)-m_x^*\right]dt$$

При вычислении этой оценки при дискретной обработке можно вместо интеграла (в зависимости от удобства вычислений) пользоваться конечной суммой. Для этого весь интервал записи реализации от 0 до T разбивается на m равных частей, каждая длиной Δt. Значения реализации в моменты времени t_1, t_2, ... t_m равные $x(t_1)$, $x(t_2)$, ..., $x(t_m)$, используется для вычисления соответствующих оценок.

Тогда оценка автокорреляционной функции по дискретной выборке, полученной в результате квантования по времени исходного непрерывного случайного процесса производится по формуле:

$$K_x^*\left(\frac{lT}{m}\right) = \frac{1}{m-l}\sum_{i=1}^{m-l}\left[x(t_i)-m_x^*\right]\cdot\left[x(t_{i+l})-m_x^*\right]$$

где $\tau = l\cdot\Delta t = \frac{lT}{m}$ $l = 0, 1, \ldots, n$; $\underline{\mathbf{n<m.}}$

Относительная величина погрешности определения автокорреляционной функции определяется выражением:

$$\delta_{AK}^2 = \frac{D\left[K_x^*(\tau)\right]}{K_x^*(0)} < 4\frac{\tau_{x\max}}{T}$$

где T – требуемый интервал наблюдения $T \geq \frac{4}{\delta_{AK}^2}\cdot\tau_{\max}$

$$\delta_{AK} \leq 50\% \qquad T \geq 16\tau_{\max}$$

$$\delta_{AK} \leq 30\% \qquad T \geq 45\tau_{\max}$$

$$\delta_{AK} \leq 10\% \qquad T \geq 400\tau_{\max}$$

$$\delta_{AK} \leq 2\% \qquad T \geq 10^4\tau_{\max}$$

Сглаживание измеряемых сигналов

Измеряемый сигнал (например, показания датчика) X(t) в большинстве случаев может быть представлен как сумма некоррелированных между собой случайных процессов: полезного сигнала U(t), функционально связанного с некоторым параметром Y(t) технологического процесса (сигнал с термопары; с дифманометра) и шума S(t), являющегося следствием различного рода помех и ошибок:

$$X(t) = U(t) + S(t)$$

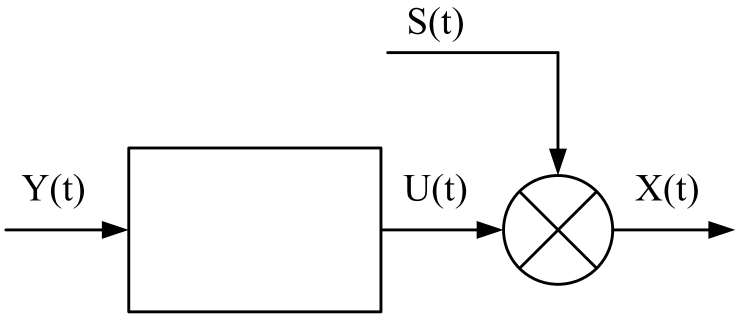


Рис. 4.2

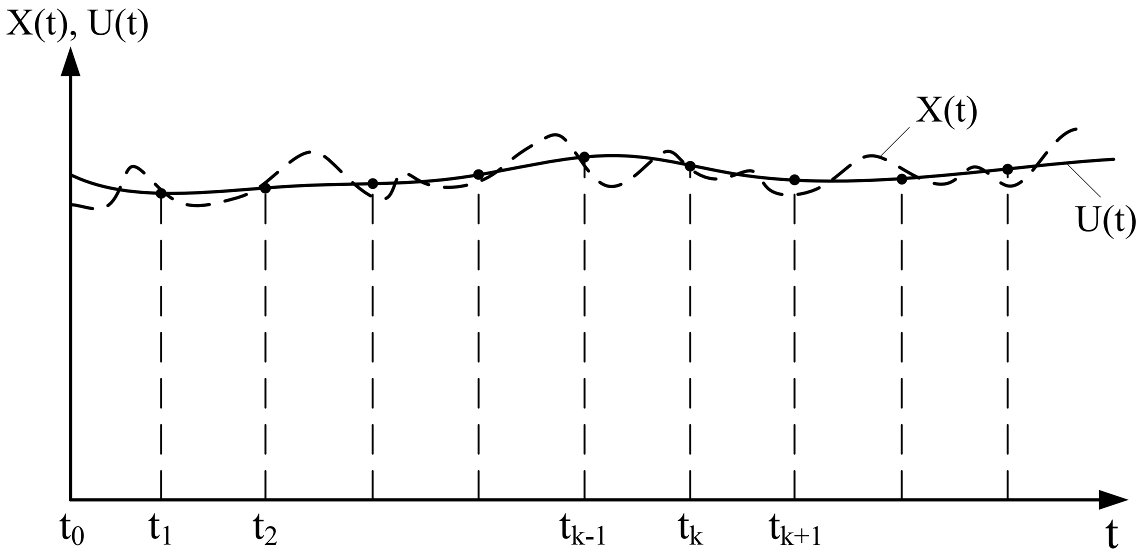


Рис. 4.3

Считаем, что на ограниченном интервале текущих наблюдений параметр Y(t) и соответственно U(t), а также шум S(t) могут считаться стационарными случайными процессами. Кроме того, полагаем, что шум центрирован, т.е. M[S(t)]=0 (систематические ошибки исключены), а спектры полезного сигнала и шума существенно отличаются, так что для интервалов корреляции выполняется условие $\tau_u >> \tau_s$.

Процедура контроля проиллюстрирована на рис 4.3. Измерения величины X(t) производятся дискретно с интервалом $\Delta t = T_ц$. В том случае, когда помеха незначительна ($D_s << D_u$), для дискретного контроля параметра Y(t) не требуется никакой дополнительной обработки измеряемой величины X(t).

Увеличение доли помехи в измеряемом сигнале приводит к необходимости уменьшения ее влияния, фильтрации.

С целью выделения полезного сигнала сглаживание сигнала с датчика можно производить по широко используемой в инженерной практике формуле экспоненциального сглаживания.

Несложный алгоритм фильтрации, учитывающий недостатки арифметического усреднения, основывается на следующей формуле:

$$\widetilde{U}_k = \alpha \cdot X_k + (1-\alpha)\cdot \widetilde{U}_{k-1},$$

где α - коэффициент сглаживания, диапазон его измерения: $0 \le \alpha \le 1$.

Из формулы следует, что для оценки значения полезного сигнала в текущий момент времени $t=t_k$, необходимо взять текущий результат измерения X_k с весом α и предыдущий результат сглаживания U_{k-1} с весом (1 - α). Для реализации операции сглаживания необходимо хранить в памяти всего два параметра - постоянную сглаживания α и результат предыдущего сглаживания U_{k-1}.

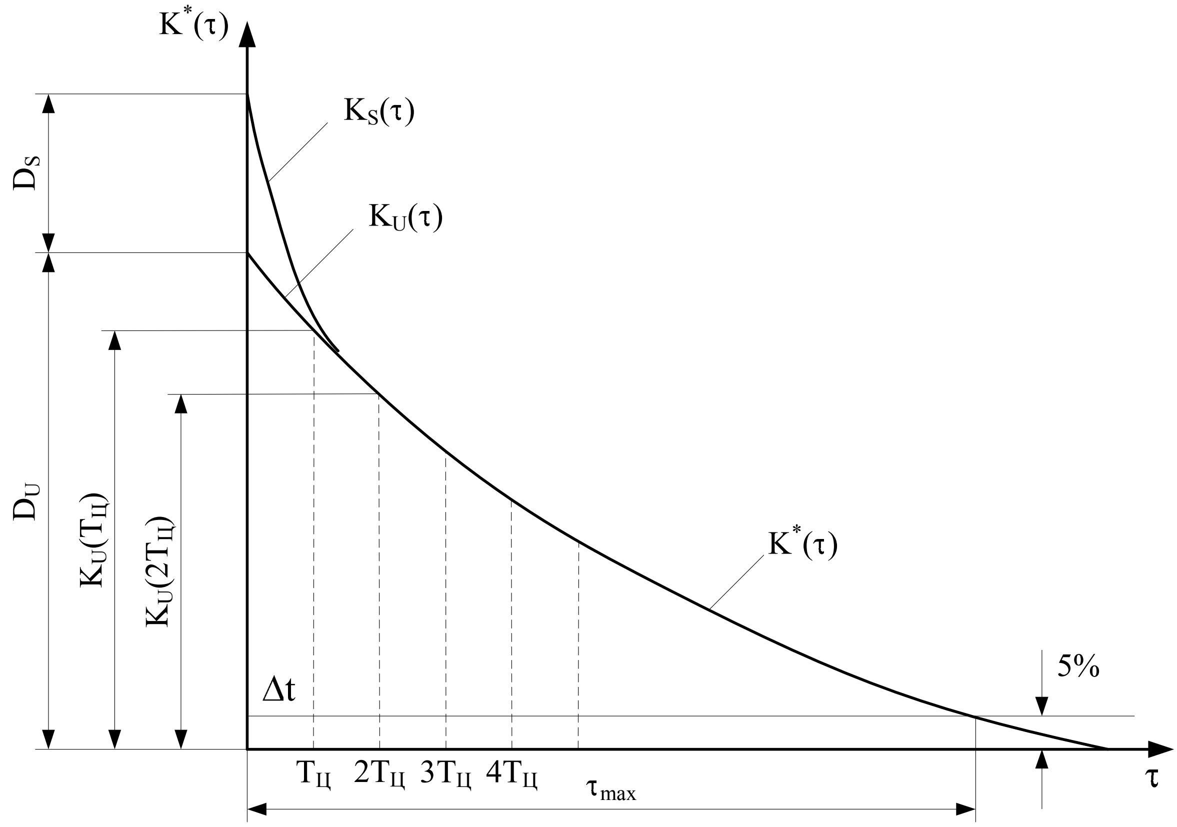


Рис. 4.4.

Расчет коэффициента сглаживания α для случайных процессов с вогнутыми автокорреляционными функциями полезного сигнала может быть произведен по формуле:

$$\alpha = \frac{2 \cdot b \cdot \left[r_{и}(T_{ц}) - r_{и}(2T_{ц}) \right]}{1 + 2 \cdot b \cdot \left[1 - r_{и}(2T_{ц}) \right]},$$

где $b = \dfrac{D_U}{D_S}; r_U = \dfrac{K_U}{D_U}$.- нормированная автокорреляционная функция полезного сигнала (см. рис. 4.4.).

4.2. Алгоритм расчета

1. Автокорреляционную функцию по заданному малому массиву и большому (общему) массиву рассчитать на ЭВМ с помощью программы – (L:\Study\СМК\Лаб_работа_4\автокорреляционная_функция.exe).

2. Строятся автокорреляционные функции по малому и большому (общему) массивам. Оценивается относительная величина погрешности определения автокорреляционной функции для малого и большого (общего) массивов.

3. Выделяется по графику АКФ (526): D_u – дисперсия полезного сигнала; D_s – дисперсия помехи; АКФ полезного сигнала.

4. Вычисляется коэффициент сглаживания – α.

5. Рассчитываются сглаженные значения сигнала для малого (заданного индивидуально) массива данных по формуле экспоненциального сглаживания:

$$\tilde{U}_k = (1-\alpha)\cdot\tilde{U}_{k-1} + \alpha\cdot x_k$$

6. Строятся на графике исходные данные и сглаженные значения сигнала.

4.3. Последовательность выполнения задания

1. Выписать в рабочую тетрадь все необходимые для оценивания автокорреляционной функции формулы.
2. Выписать согласно номера варианта, значения массива экспериментальных точек
3. Перед запуском программы необходимо создать файл исходных данных ***имя.txt,*** состоящий из ***n*** значений X.

Образец:

205 211 206 … 208 204***размещение курсора (конец файла)***

4. Выполнить задание согласно номера варианта.

4.4. Задание

1. Рассчитать на ЭВМ с помощью программы (L:\Study\СМК\Лаб_работа_4\Автокорреляционная_функция.exe) автокорреляционные функции по заданному малому массиву (файл ***имя.txt)*** и большому (общему) массиву (файл **akf526.txt**). Адрес файла **akf526.txt** (L:\Study\СМК\Лаб_работа_4\akf526.txt).
2. Построить автокорреляционные функции по заданному малому массиву и большому (общему) массиву
3. Оценить относительную величину погрешности определения автокорреляционной функции для малого и большого (общего) массивов.
4. Выделить по графику АКФ (526): D_u – дисперсию полезного сигнала; D_s – дисперсию помехи; АКФ полезного сигнала.
5. Вычислить α - коэффициент сглаживания.
6. Построить на графике исходные данные и сглаженные значения заданного малого массива.

5. Планирование экспериментов. планы первого порядка

5.1. Теория

Планирование эксперимента

Для экстремального планирования эксперимента наибольшее применение нашли модели в виде алгебраических полиномов

$$y = \beta_0 + \sum_{j=1}^{n} \beta_j x_j + \sum_{\substack{i,j=1 \\ i \neq j}}^{n} \beta_{ij} x_i x_j + \sum_{j=1}^{n} \beta_{jj} x_j^2 + \ldots \qquad (5.1.)$$

В зависимости от порядка полинома модели (5.1.) различают планы первого порядка, второго порядка, в общем случае “m” порядка, но обычно ограничиваются на практике планами первого либо второго порядков

Если выбрана модель и записано её уравнение, то остаётся в отведённой для исследования области факторного пространства спланировать и провести эксперимент для оценки численных значений коэффициентов этого уравнения. При этом план должен содержать количество опытов не менее числа коэффициентов уравнения модели.

Планы первого порядка

Сущность факторного эксперимента первого порядка состоит в одновременном варьировании всех факторов при его проведении по определенному плану, представлении математической модели (функции отклика) в виде линейного полинома и исследовании последнего методами математической статистики

$$У = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n \qquad (5.2)$$

Введём понятия:

Уровнем фактора называют определенное значение фактора, которое будет фиксироваться при проведении эксперимента.

Уровнями факторов можно назвать и средние значения интервалов. Эти значения называют нулевыми уровнями, обозначим X_{io}.

Интервал варьирования – это такое значение фактора в натуральных единицах, прибавление которого к нулевому уровню даёт верхний, а вычитание – нижний уровень фактора, обозначим ΔX_j .

Экстремальные значения, которые могут принимать факторы, не меняя своих физико-химических свойств и не искажая сути исследуемого процесса, назовём границами существования факторов, а интервал ($X_{j\max} - X_{j\min}$) – интервалом определения фактора (рис. 5.1).

Как правило, область варьирования факторов M должна составлять часть области определения факторов L , если решается задача оптимизации.

В задачах же аппроксимации (или интерполяции) интервал варьирования охватывает всю описываемую область – *L*.

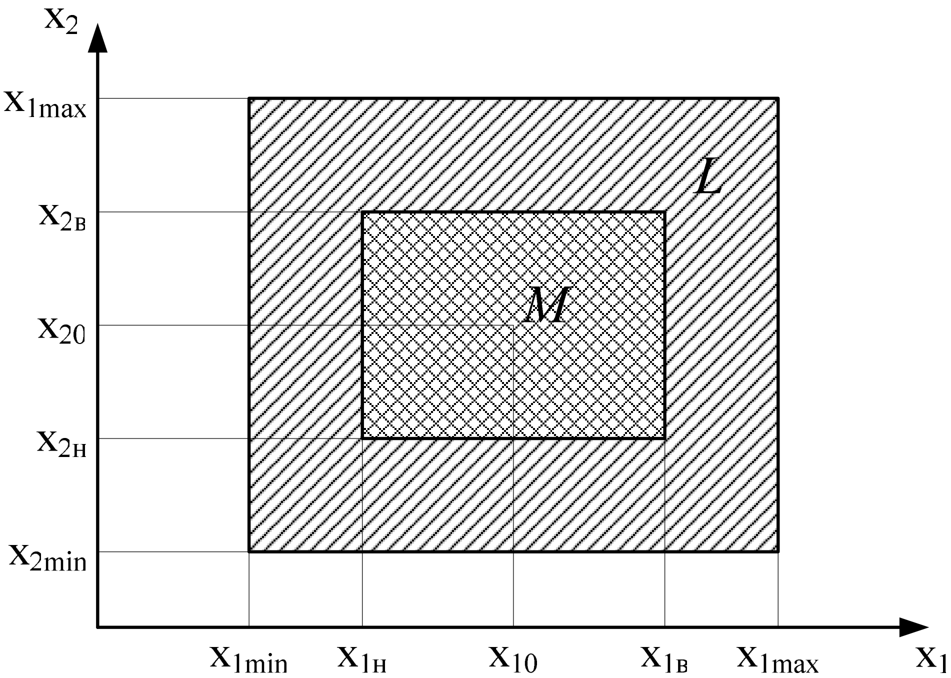


Рис. 5.1

Верхние и нижние уровни факторов обозначают символами “+1”, “–1”, что достигается кодированием факторов по формуле

$$x_j = \frac{X_j - X_{j0}}{\Delta X_j} \qquad (5.3)$$

Кодирование факторов, по сути, означает переход от системы координат в натуральных единицах (рис. 5.2 а) к системе координат в кодированной форме (рис. 5.2 б).

Каждая точка факторного пространства (+1, +1), (–1, +1), (+1, –1), (–1, –1) – это опыт в исследованиях.

Каждая точка факторного пространства (+1, +1), (–1, +1), (+1, –1), (–1, –1) – это опыт в исследованиях.

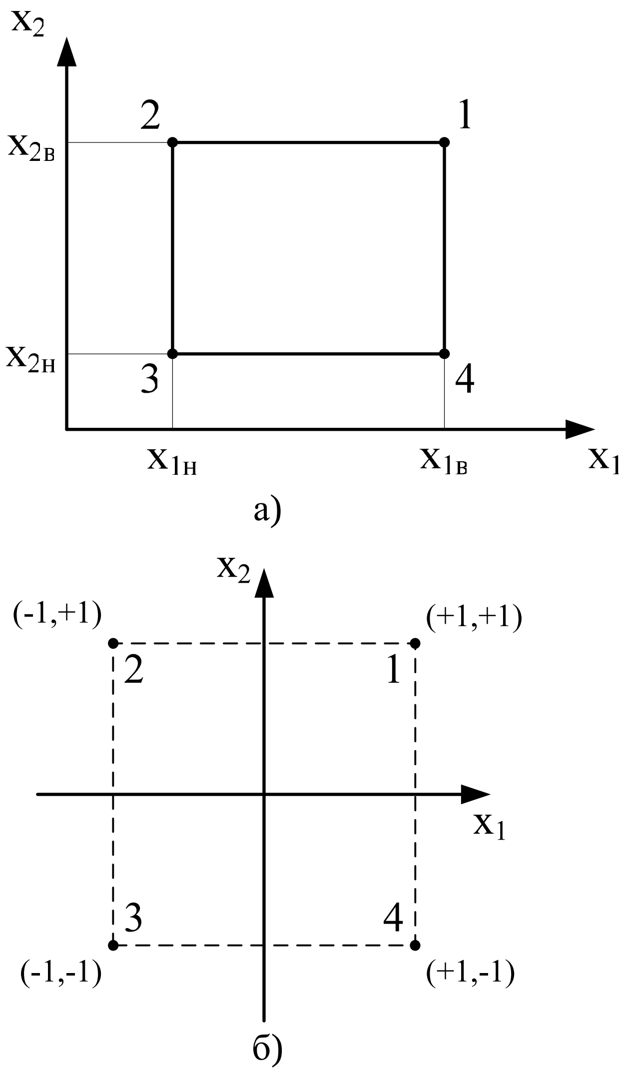


Рис. 5.2

В общем случае эксперимент, в котором реализуются все возможные сочетания уровней факторов, называется полным факторным экспериментом (ПФЭ). Если каждый фактор варьируется на двух уровнях, то получается полный факторный эксперимент типа 2^n. Для двух факторов (n = 2) число опытов $N = 2^2 = 4$. (рис. 5.2).

Можно осуществлять планирование эксперимента на трёх уровнях (верхний, средний, нижний), тогда ПФЭ будет типа 3^n и для n = 2 общее число опытов будет $= 3^2 = 9$. (рис. 5.3).

Если К – число уровней; *n* – число факторов, то количество опытов N равно:

$$N = K^{\,n}$$

Правильный выбор нулевых уровней (центра плана) и интервалов варьирования факторов имеет решающее значение для действенности математической модели.

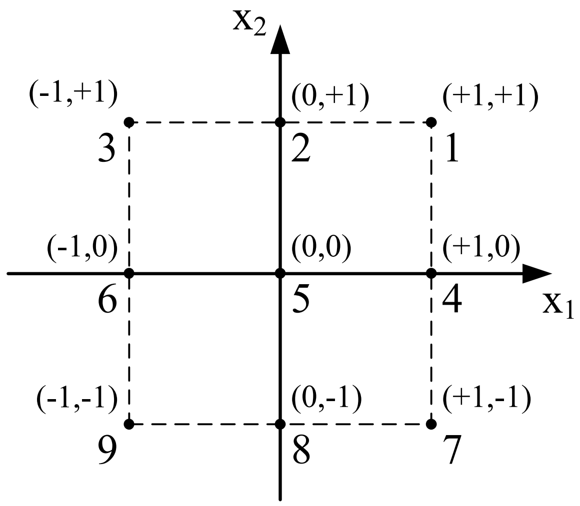


Рис. 5.3

Огромная роль при изучении не исследовавшихся ранее объектов с целью получения их математических моделей принадлежит аналогии и интуиции.

Основное требование к интервалу варьирования состоит в том, чтобы он превышал удвоенную квадратическую ошибку фактора:

$$2\sigma_{xi} < \Delta \mathrm{X}_i < \left(\mathrm{X}_{j\max} - X_{j\min}\right)/2$$

где σ_{xi} – среднеквадратическое отклонение фактора X_i; $\Delta\mathrm{X}_i$ – интервал варьирования; ($\mathrm{X}_{j\max} - \mathrm{X}_{j\min}$) – область определения фактора.

Это требование связано с тем, что интервал между двумя соседними уровнями должен значимо (неслучайно) влиять на переменную состояния. Обычно интервал варьирования выбирается на основании априорной информации (или интуитивно) и затем уточняется (если он выбран неудачно) после получения математической модели. Цена уточнения ощутима, так как повторение эксперимента резко увеличивает число опытов.

Удачный выбор интервала варьирования факторов гарантирует получение достоверной математической модели объекта.

Построение матрицы планирования

План, содержащий запись всех комбинаций факторов или их части в кодированной форме, называется матрицей планирования.

При построении матриц планирования применяется прием чередования знаков. В первом столбце знаки не меняются, во втором меняются поочередно, в третьем они чередуются через два, в четвертом – через 4 и т.д. (по показателям степеней двойки). Вводится столбец x_0 –

столбец значений фиктивной переменной, его участие в матрице планирования делает расчеты коэффициентов математической модели более общими.

Таблица 5.1.

Организация матриц планирования ПФЭ от 2^2 до 2^5

№ п/п	Тип эксперимента				Планирование					
					x_0	x_1	x_2	x_3	x_4	x_5
1.					+1	+1	+1	+1	+1	+1
2.				ПФЭ 2^2	+1	–1	+1	+1	+1	+1
3.					+1	+1	–1	+1	+1	+1
4.			ПФЭ 2^3		+1	–1	–1	+1	+1	+1
5.					+1	+1	+1	–1	+1	+1
6.					+1	–1	+1	–1	+1	+1
7.					+1	+1	–1	–1	+1	+1
8.					+1	–1	–1	–1	+1	+1
9.		ПФЭ 2^4			+1	+1	+1	+1	–1	+1
10.					+1	–1	+1	+1	–1	+1
11.					+1	+1	–1	+1	–1	+1
12.					+1	–1	–1	+1	–1	+1
13.					+1	+1	+1	–1	–1	+1
14.					+1	–1	+1	–1	–1	+1
15.					+1	+1	–1	–1	–1	+1
16.					+1	–1	–1	–1	–1	+1
17.					+1	+1	+1	+1	+1	–1
18.	ПФЭ 2^5				+1	–1	+1	+1	+1	–1
19.					+1	+1	–1	+1	+1	–1
20.					+1	–1	–1	+1	+1	–1
21.					+1	+1	+1	–1	+1	–1
22.					+1	–1	+1	–1	+1	–1
23.					+1	+1	–1	–1	+1	–1
24.					+1	–1	–1	–1	+1	–1
25.					+1	+1	+1	+1	–1	–1
26.					+1	–1	+1	+1	–1	–1
27.					+1	+1	–1	+1	–1	–1
28.					+1	–1	–1	+1	–1	–1
29.					+1	+1	+1	–1	–1	–1
30.					+1	–1	+1	–1	–1	–1
31.					+1	+1	–1	–1	–1	–1

32.			+1	–1	–1	–1	–1	–1

Рассмотренные матрицы планирования обладают такими свойствами, которые позволяют считать, что их построение выполнялось оптимально с точки зрения получаемой по результатам реализации матрицы планирования математической модели. Если мы ищем модель в виде уравнения регрессии, то коэффициенты должны быть наилучшими и точность предсказания значений переменной состояния одинакова в любом направлении факторного пространства. Эти требования формулируются как условия **ортогональности и ротатабельности.**

Из построения матрицы планирования вытекают следствия, которые математически можно представить так:

свойство симметричности $$\sum_{U=1}^{N} x_{ju} = 0 \qquad (5.4)$$

т.е. сумма элементов любого столбца матрицы планирования равна нулю;

свойство нормировки $$\sum_{u=1}^{N} x^2{}_{ju} = N \quad (j = 1, 2, ..., n;\ u - 1, 2, ..., N) \qquad (5.5)$$

т.е. сумма квадратов элементов каждого столбца равна числу опытов.
Где n – число факторов;

N – число опытов (или строк матрицы планирования).

Эти условия легко проверить по таблице 5.1.

Условие ортогональности предполагает равенство нулю суммы по членных произведений любых двух вектор–столбцов матрицы:

$$\sum_{u=1}^{N} x_{iu} x_{ju} = 0 \qquad (i, j = 1, 2, \ldots, n, \quad i \neq j) \qquad (5.6)$$

это условие также легко проверить по табл. 5.1. Действительно, полный факторный эксперимент типа 2^n является ортогональным. Ортогональные планы ПФЭ (для линейных моделей) имеют свойство – ротатабельность.

Последнее предполагает равенство и минимальность дисперсий предсказанных значений переменной состояния для всех точек факторного пространства.

Полный факторный эксперимент типа 2^n (при равном числе параллельных опытов в каждой точке факторного пространства)

Если априорные сведения предполагают невысокую воспроизводимость результатов, то в матрицу планирования эксперимента включают параллельные опыты, в результате получают экспериментальный материал для последующей обработки (табл. 5.2).

Таблица 5.2.

Матрица ПФЭ 2^2 с параллельными опытами

Номер опыта	x_0	Планирование		Результаты наблюдений. Параметр оптимизации (функция отклика)				Результаты предварительной обработки	
		x_1	x_2	y_{u1}	y_{u2}	...	y_{um}	$\bar{y}_u$	D_u^*
1	+ 1	+ 1	+ 1	y_{11}	y_{12}	...	y_{1m}	$\bar{y}_1$	D_1^*
2	+ 1	– 1	+ 1	y_{21}	y_{22}	...	y_{2m}	$\bar{y}_2$	D_2^*
3	+ 1	+ 1	– 1	y_{31}	y_{32}	...	y_{3m}	$\bar{y}_3$	D_3^*
4	+ 1	– 1	– 1	y_{41}	y_{42}	...	y_{4m}	$\bar{y}_u$	D_4^*

Расчет ошибки опыта (дисперсии воспроизводимости)

Разброс параметра оптимизации обусловлен наличием ошибки опыта (ошибки воспроизводимости). Ошибка опыта является суммарной величиной, состоящей из ошибок при измерении факторов и параметра оптимизации; а также ошибок при проведении опыта.

Знание ошибки воспроизводимости необходимо для анализа данных эксперимента.

Воспроизводимость оценивают по результатам параллельных опытов.

Перед расчетом ошибки опыта необходимо убедиться, что рассеяние опытов в каждой точке факторного пространства не превышает некоторой величины. С этой целью рассчитывают построчные дисперсии и проверяют их однородность.

$$D_u^*\{y\} = \frac{1}{m-1}\sum_{k=1}^{m}\left(y_{uk} - \bar{y}_u\right)^2 \qquad (5.7)$$

где
$$\bar{y}_u = \frac{\sum_{k=1}^{m} y_{uk}}{m};$$

Требование однородности дисперсий является одним из требований регрессионного анализа.

Для проверки однородности построчных дисперсий (при равном числе параллельных опытов на каждой строчке матрицы планирования) применяют критерий Кохрена. Согласно критерию Кохрена, образуют отношение максимальной дисперсии к сумме всех дисперсий

$$G_p = \frac{D^*_{u\max}}{\sum\limits_{u=1}^{N} D^*_u}$$

где $D^*_{u\,max}$ – максимальная из рассчитанных построечных дисперсий;

$\sum\limits_{u=1}^{N} D^*_u$ – сумма всех дисперсий по N строкам матрицы планирования.

Если выполняется условие:

$$G_p \le G_T \qquad (5.8)$$

то гипотеза об однородности дисперсий принимается.

$G_т$ – табличное значение для степеней свободы $\upsilon_1 = m - 1$ и $\upsilon_2 = N$ и уровня значимости α.

При выполнении условия (5.8) дисперсия воспроизводимости находится путем усреднения построчных дисперсий по формуле

$$D^*_0 = \frac{1}{N}\sum_{u=1}^{N} D^*_u = \frac{1}{N(m-1)}\sum_{u=1}^{N}\sum_{k=1}^{m}\left(y_{uk} - \bar{y}_u\right)^2 \qquad (5.9)$$

где $\upsilon = N(m - 1)$ – число степеней свободы.

Таким образом получают ошибку опыта.

Неоднородные дисперсии усреднять нельзя.

Проверка значимости различия $\bar{y}_{\max}$ и $\bar{y}_{\min}$

Реализовав эксперимент и вычислив средние значения откликов для каждой точки плана, нужно проверить, значимо ли отличаются друг от друга $\bar{y}_{max}$ *и* $\bar{y}_{min}$. Значимость различия двух средних можно проверить с помощью t–критерия (критерия Стьюдента) по формуле

$$t_p = \left(\bar{y}_{max} - \bar{y}_{min}\right) \big/ \sigma^* \sqrt{1/m_{max} + 1/m_{min}} > t_T \qquad (5.10)$$

если $t_p > t_т$ – то $\bar{y}_{max}$ *и* $\bar{y}_{min}$ значимо отличаются.

Средние значимо отличаются, если расчетное значение t – критерия превосходит табличное (для $\upsilon = m_{\max} + m_{\min}$ степеней свободы и заданного уровня значимости α).

Если расчетное значение $t_р$ будет меньше табличного, то с вероятностью Р = 1 – α можно считать, что разницы между результатами двух опытов нет.

Если статистически незначима разница между максимальным и минимальным результатами плана, вряд ли удастся получить, сколько нибудь ценную информацию, применив к обработке таких практически идентичных данных метод регрессионного анализа.

Расчет коэффициентов уравнения регрессии

Основная задача обработки результатов эксперимента сводится к получению математической модели процесса в виде (5.2).

Искомая модель представляет собой уравнение регрессии, т.е. при обработке стоит задача определения коэффициентов уравнения регрессии.

При обработке экспериментальных данных, полученных при планировании эксперимента, используют метод наименьших квадратов – эффективный и простой способ получения оценок коэффициентов уравнения регрессии.

При этом предполагают выполнение следующих предпосылок: независимые переменные x_j достаточно точно поддерживаются на определенных уровнях, а наблюдаемые значения отклика y_1, y_2, ...y_N,…, y_{N+m} по данным N; (N·m) опытов плана представляют собой независимые и нормально распределенные случайные величины.

Полученный в результате опытов ограниченный статистический материал даёт возможность определить лишь оценки b_0, b_1, ..., b_n теоретических коэффициентов регрессии β_0, β_1, ..., β_n в (5.2), справедливых для некоторой гипотетической совокупности, состоящей из всех мыслимых опытов. Тогда уравнение регрессии, полученное на основании N; (N·m) опытов, запишется следующим образом

$$\hat{y} = b_0 + b_1 x_1 + \ldots + b_n x_n = \sum_{j=0}^{n} b_j x_j \qquad (5.11)$$

Любой коэффициент уравнения регрессии (5.11) b_j определяется скалярным произведением столбца $\overline{Y} = \{y_u\}$ на соответствующий столбец $\overline{X}_j = \{x_{uj}\}$ и делением его на число опытов в матрице планирования N:

$$b_j = \frac{\sum_{u=1}^{N} x_{uj} y_u}{N} \qquad (5.12)$$

т.е. коэффициенты уравнения регрессии определяются независимо друг от друга. Для расчета b_0 по формуле (5.12) в матрицу планирования вводят вектор-столбец фиктивной переменной $\overline{X}_0 = \{x_{u0}\}$, которая принимает во всех опытах значение +1.

Коэффициенты при независимых переменных указывают на силу влияния факторов. Если коэффициент имеет знак плюс, то с увеличением значения фактора параметр оптимизации увеличивается и наоборот. Величина коэффициента соответствует вкладу данного фактора в величину параметра оптимизации при переходе фактора с нулевого уровня на верхний или нижний.

Если коэффициенты уравнения регрессии рассчитываются по матрице планирования с m параллельными опытами, то

$$b_j = \frac{\sum_{u=1}^{N} x_{uj}\bar{y}_u}{N}, \ \left(j = 0,1,\ldots,n\right), \quad (5.13)$$

где $\bar{y}_u = \frac{\sum_{k=1}^{m} y_{uk}}{m}$ среднее значение по параллельным опытам *u*-й строки матрицы планирования; формулу (5.12) можно переписать в виде

$$b_j = \frac{\sum_{u=1}^{N}\sum_{k=1}^{m} x_{uj}y_{uk}}{N \cdot m} \quad (5.14)$$

Проверка значимости коэффициентов уравнения регрессии

Очевидно, что один фактор больше влияет на параметр оптимизации, другой – меньше. Для оценки этого влияния используют проверку значимости каждого коэффициента по критерию Стьюдента.

1. Находят дисперсии коэффициентов уравнения регрессии по формуле

$$D^*_{b_j} = \frac{D^*_0}{N \cdot m} \quad (5.15)$$

Для ПФЭ дисперсии коэффициентов регрессии равны и определяются независимо друг от друга, зависят только от ошибки опыта D^*_0, числа строк матрицы планирования N и числа параллельных опытов на каждой строчке матрицы планирования *m*.

2. Для проверки значимости коэффициентов регрессии используют доверительный интервал Δb_j, который одинаков для всех b_j вследствие равенства $\sigma^*_{b_j}$ для всех коэффициентов.

$$\Delta b_j = \pm t_T \sigma^*_{b_j} \quad (5.16)$$

t_T – табличное значение критерия Стьюдента, которое находят по числу степеней свободы $\upsilon_0 = N\,(m-1)$ и уровню значимости α;

$\sigma^*_{b_j} = \sqrt{D^*_{b_j}}$ – оценка среднеквадратичного отклонения b_j.

Если выполняется условие:

$$| b_j | > | \Delta b_j | \qquad (5.17)$$

то b_j признаётся значимым,

Если для какого-то коэффициента условия (5.17) не выполняются, то соответствующий фактор можно признать незначимым и исключить его из уравнения регрессии.

Проверка адекватности модели

(линейного уравнения регрессии $\hat{y} = \sum_0^n b_j x_j$)

Проверку адекватности линейного уравнения регрессии можно провести сравнением двух дисперсий – одна показывает рассеяние средних опытных данных параметра оптимизации $\bar{y}_u$ относительно предсказанных уравнением регрессии $\hat{y}$.

Эта дисперсия называется дисперсией адекватности и рассчитывается по формуле

$$D^*_{ad.} = \frac{m}{N-l} \sum_{u=1}^{N} \left(\bar{y}_u - \hat{y}_u \right)^2 \qquad (5.18)$$

где m – число параллельных опытов;

N – число строк матрицы планирования;

l – число членов в уравнении регрессии, оставшихся после оценки значимости его коэффициентов.

$\hat{y}$ – значения выходной координаты, вычисленные по уравнению регрессии полученного с учетом значимых коэффициентов.

Другая дисперсия D^*_0 показывает рассеивание отдельных значений результатов эксперимента от средних значений откликов для каждой точки плана и вычисляется по формуле (5.9).

Адекватность модели проверяется по критерию Фишера.

Если выполняется условие

$$F_P = \frac{D^*_{ad.}}{D^*_0} < F_T \qquad (5.19)$$

(F_T – коэффициент Фишера находится для степеней свободы $\upsilon_{ад.} = N - l$; $\upsilon_0 = N\ (m - 1)$ и заданного уровня значимости – α), то линейное уравнение регрессии (5.11) признается адекватным, т.е. рассеяние экспериментальных данных параметра оптимизации относительно уравнения регрессии того же порядка, что и рассеяние, вызванное случайными изменениями в объекте исследования (ошибка опыта).

При расчете F_P предполагается, что $D^*_{ад.} > D^*_0$. Однако на практике бывает, что $D^*_{ад.} \le D^*_0$, тогда вывод об адекватности модели может быть сделан без проверки условия (5.19).

Уравнение регрессии относительно факторов в натуральных величинах

Коэффициенты уравнения регрессии с кодированными факторами отличаются от коэффициентов уравнения регрессии с натуральными значениями факторов x_j и определяются из **полного факторного эксперимента** (ПФЭ), проведённого в рассматриваемой ограниченной области.

Для расчёта натуральных значений коэффициентов, в кодированное уравнение регрессии (5.11) вместо кодированных факторов x_j (j = 1, … n) следует подставить выражения для последних через натуральные значения факторов X_j (j = 1, … n) (5.3). В результате получим следующее уравнение, представляющее собой уравнение регрессии относительно факторов в натуральных величинах:

$$\hat{y} = B_0 + B_1 X_1 + \ldots + B_n X_n = \sum_{j=0}^{n} B_j X_j \qquad (5.20)$$

5.2. Алгоритм расчета

1. Кодирование факторов

$$x_i = \frac{X_i - X_{i0}}{\Delta X_i};$$

2. Вычисляются средние значения отклика $\overline{y}_u$ по строкам матрицы

$$\overline{Y_u} = \frac{\sum_{k=1}^{m} Y_{uk}}{m}$$

3. Вычисляются построечные дисперсии

$$D_u = \frac{\sum_{k=1}^{m} \left(Y_{uk} - \overline{Y_u}\right)^2}{(m-1)}$$

4. Проверяется однородность построечных дисперсий

$$G_p = \frac{D_u^{\max}}{\sum_{1}^{N} D_u} \le G_T$$ – критерий Кохрена

Табличное значение, коэффициента Кохрена для степеней свободы υ_1 = m-1, υ_2 = N, α =0,05 находится по таблице1 (Приложение В).

5. Вычисляется ошибка опыта (дисперсия воспроизводимости)

$$D_0 = \frac{\sum_{1}^{N} D_u}{N}; \qquad \sigma_0 = \sqrt{D_0};$$

6. Проверяется значимость различия между $\overline{y}_{\max}$ и $\overline{y}_{\min}$ по критерию Стьюдента:

$$t_p = \frac{\overline{y}_{\max} - \overline{y}_{\min}}{\sigma_0 \sqrt{\frac{1}{m_{\max}} + \frac{1}{m_{\min}}}} > t_T$$

где $m_{\max} = m_{\min}$; t_T – коэффициент Стьюдента для $\nu = m_{\max} + m_{\min}$; α = 0,05. (Приложение В. Таблица 2).

7. Рассчитываются коэффициенты уравнения регрессии:

$$\hat{y}_u = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + b_4 x_4 + b_5 x_5$$

$$b_j = \frac{\sum_{u=1}^{N} x_{uj} \overline{y}_u}{N}$$

8. Рассчитываются дисперсии коэффициентов уравнения регрессии:

$$D_{b_j} = \frac{D_0}{N \cdot m}; \qquad \sigma_{b_j} = \sqrt{D_{bj}}$$

9. Оценивают значимость коэффициентов уравнения регрессии: $\Delta b_j = t_{CT} \cdot \sigma_{b_j}$; $t_{ст}$ – коэффициент Стьюдента для $\upsilon_0 = N(m-1)$; $\alpha = 0{,}05$. (Приложение В. Таблица 2).

10. Проверяется адекватность линейного уравнения регрессии по критерию Фишера:

$$F_P = \frac{D_{aq}}{D_0} < F_T$$

F_T - критерий Фишера для $\begin{cases} \upsilon_{aq} = N - \ell \\ \upsilon_0 = N(m-1) \end{cases}$; α = 0,05. (Приложение В. Таблица 3).

$$D_{aq} = \frac{m}{N-\ell} \sum_{u=1}^{N} \left(\overline{y}_u - \hat{y}_u \right)^2$$

ℓ - число значимых коэффициентов построенного уравнения регрессии;

$\hat{y}_u$ – значения выходной координаты, вычисленные по уравнению регрессии полученного с учетом значимых коэффициентов;

$$D_0 = \frac{\sum_{u=1}^{N} D_u}{N}.$$

Искомое уравнение регрессии принимает вид:

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + b_4 x_4 + b_5 x_5.$$

11. Проводится анализ полученного уравнения регрессии.

12. Записывается полученное уравнение регрессии относительно факторов в натуральных величинах.

5.3. Последовательность выполнения задания

1. Выписать в рабочую тетрадь все необходимые для расчета факторного эксперимента формулы.
2. Выполнить задание согласно номера варианта.

5.4. Задание

1. Спланировать столбцы матрицы при факторах x_4 и x_5 (если это необходимо) по заданным формулам: ($x_4 = x_1 \cdot x_2 \cdot x_3$), ($x_5 = x_1 \cdot x_2$)
2. Представить зависимость параметра оптимизации от факторов (уравнение регрессии) в виде линейного полинома.
3. Выполнить все этапы расчета факторного эксперимента:
   - вычислить средние значения отклика *y* по строкам матрицы;
   - вычислить построечные дисперсии;
   - проверить однородность построечных дисперсий;
   - вычислить ошибку опыта (дисперсия воспроизводимости);
   - проверить значимость различия между $\overline{y}_{\max}$ и $\overline{y}_{\min}$ по критерию Стьюдента;
   - рассчитать коэффициенты уравнения регрессии;
   - рассчитать дисперсии коэффициентов уравнения регрессии;
   - оценить значимость коэффициентов уравнения регрессии;
   - проверить адекватность линейного уравнения регрессии по критерию Фишера.
4. Проанализировать полученную полиномиальную модель выщелачивания восстановленного барийсодержащего сырья (характер влияния факторов на целевую функцию).
5. Записать полученное уравнение регрессии относительно факторов в натуральных величинах.

Примечание:

Табличное значение коэффициента Кохрена. (Приложение В. Таблица 1).

Табличное значение коэффициента Стьюдента. (Приложение В. Таблица 2).

Табличное значение коэффициента Фишера. (Приложение В. Таблица 3).

6. Планирование экспериментов. планы второго порядка

6.1. Теория

После достижения области оптимума перед исследователем встает задача детального изучения поверхности отклика. Описать область оптимума линейным уравнением регрессии не удается из-за значимости эффектов взаимодействия факторов и квадратичных эффектов.

Поэтому область оптимума описывается полиномами более высоких порядков, среди которых самые распространенные уравнения второго порядка:

$$\hat{y} = b_0 + \sum_{j=1}^{n} b_j x_j + \sum_{\substack{j,i=1 \\ j \neq i}}^{n} b_{ij} x_i x_j + \sum_{1}^{n} b_{jj} x_j^2 \qquad (6.1)$$

Для нахождения коэффициентов уравнения (6.1) нужны планы, в которых каждая переменная принимала бы хотя бы три разных значения.

Однако полный факторный эксперимент типа 3^n содержит слишком большое число опытов.

Сократить число опытов можно, если воспользоваться так называемыми композиционными или последовательными планами, предложенными Боксом и Уилсоном.

“Ядро” таких планов составляет ПФЭ типа 2^n (при $n < 5$ или дробная реплика от него при $n \geq 5$); “звездные” точки, расположенные на координатных осях факторного пространства ($\pm\alpha$, 0, ..., 0); (0, $\pm\alpha$, 0, ..., 0) ,..., (0, 0, ..., $\pm\alpha$).

Где α – расстояние от центра плана до “звездной” точки, т.к. называемое “звездное” плечо; и эксперименты в центре плана – m_0.

Общее число опытов в матрице композиционного плана при факторах составит

$$N = 2^n + 2n + m_0 \qquad (6.2)$$

В таблице 6.1 приводится сравнение количества точек и числа степеней свободы для ПФЭ типа 3^n и композиционного плана (6.2) при $m_0 = 1$.

Таблица 6.1

Число факторов n	ПФЭ 3^n	Числ. коэфф.	Число степеней Свободы, ν	Композиционный план, N	Число степеней свободы, ν_N
2	9	6	3	9	3
3	27	10	17	15	5

4	81	15	66	25	10
5	243	21	222	43	22
5*	243	21	222	27	6

* с полурепликой.

Рассмотрим построение композиционных планов при n = 2 (рис. 6.1).

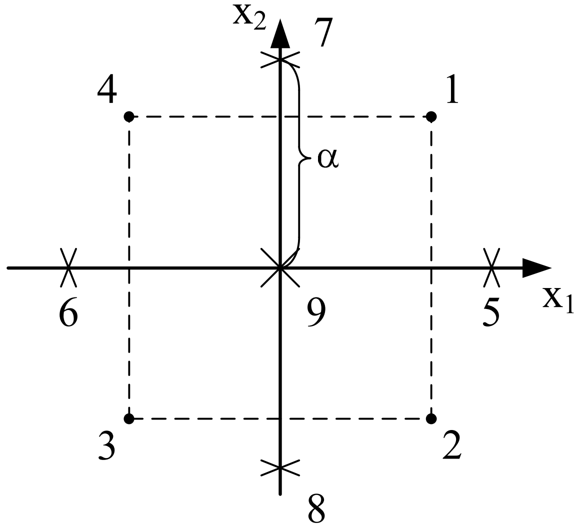


Рис. 6.1 Схема построения композиционного плана

Точки 1, 2, 3, 4 образуют ПФЭ 2^2, точки 5, 6, 7, 8 – “звездные” точки с координатами (±α, 0) и (0, ±α). Координаты m_0 опытов в центре плана нулевые – (0, 0). Величина звездного плеча α и количество опытов в центре плана m_0 зависит от выбранного плана (чаще всего m_0 = 1).

Композиционные планы легко приводятся к ортогональным выбором звездного плеча α. На количество опытов в центре плана m_0 не накладывается никаких ограничений. m_0 обычно принимается равным единице.

Рассмотрим матрицу композиционного планирования для n = 3, положив m_0 = 1 (табл. 6.2).

Таблица 6.2

		План									
N	x_0	x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3	x_1^2	x_2^2	x_3^2	y
1	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1	y_1
2	+1	–1	+1	+1	–1	–1	+1	+1	+1	+1	y_2

3	+1	+1	–1	+1	–1	+1	–1	+1	+1	+1	y_3
4	+1	–1	–1	+1	+1	–1	–1	+1	+1	+1	y_4
5	+1	+1	+1	–1	+1	–1	–1	+1	+1	+1	y_5
6	+1	–1	+1	–1	–1	+1	–1	+1	+1	+1	y_6
7	+1	+1	–1	–1	–1	–1	+1	+1	+1	+1	y_7
8	+1	–1	–1	–1	+1	+1	+1	+1	+1	+1	y_8
9	+1	$+\alpha$	0	0	0	0	0	α^2	0	0	y_9
10	+1	$-\alpha$	0	0	0	0	0	α^2	0	0	y_{10}
11	+1	0	$+\alpha$	0	0	0	0	0	α^2	0	y_{11}
12	+1	0	$-\alpha$	0	0	0	0	0	α^2	0	y_{12}
13	+1	0	0	$+\alpha$	0	0	0	0	0	α^2	y_{13}
14	+1	0	0	$-\alpha$	0	0	0	0	0	α^2	y_{14}
15	+1	0	0	0	0	0	0	0	0	0	y_{15}

В общем виде матрица неортогональна, так как

$$\left.\begin{aligned} &\sum_{u=1}^{N} x_{u0} x_{uj}^2 \neq 0 \\ &\sum_{u=1}^{N} x_{ui}^2 x_{uj}^2 \neq 0 \end{aligned}\right\} \qquad (6.3)$$

Чтобы получить ортогональное планирование второго порядка необходимо:

1) произвести некоторое преобразование квадратичных переменных;

2) специальным образом выбрать величину звездного плеча.

Введем преобразование

$$x_{uj}^1 = x_{uj}^2 - \frac{\sum_{u=1}^{N} x_{uj}^2}{N} = x_{uj}^2 - \overline{x_{uj}^2} \qquad (6.4)$$

Тогда будут равны нулю скалярные произведения

$$\sum_{u=1}^{N} x_{u0} x_{uj}^1 = \sum_{u=1}^{N} x_{uj}^2 - N\overline{x_{uj}^2} = 0$$

$$\sum_{u=1}^{N} x_{u0} x_{uj}^1 = \sum_{u=1}^{N} x_{u0} x_{uj}^2 - \sum_{u=1}^{N} x_{u0} \overline{x_{uj}^2} = 0$$

т.к. $\sum_{u=1}^{N} x_{u0} x_{uj}^1 = \sum_{u=1}^{N} x_{uj}^2 - N\overline{x_{uj}^2} = N - N \cdot 1$

Неортогональными останутся только вектор - столбцы для вновь введённых квадратичных членов.

$$\sum_{u=1}^{N} x_{ui}^1 x_{uj}^1 \neq 0$$

Чтобы сделать матрицу планирования полностью ортогональной, величину звездного плеча α выбирают из условия равенства нулю недиагонального члена матрицы $(X^T X)^{-1}$. В табл. 6.3 приведены значения α и общее число опытов для различного числа независимых переменных n.

Таблица 6.3

Число опытов	Число независимых переменных			
	2	3	4	5*
В ядре плана $N_я$	4	8	16	16
В “звездных” точках N_α	4	6	8	10
В центре плана m_0	1	1	1	1
Общее число опытов N	9	15	25	27
Величина “звездного” плеча α	1,00	1,215	1,414	1,547

* с полурепликой

Ортогональная матрица второго порядка для n = 2 имеет вид табл. 6.4.

Таблица 6.4

Опыты	x_0	План					Параметр оптимизации
		x_1	x_2	$x_1 x_2$	x_1^1	x_2^1	
1	+1	+1	+1	+1	1/3	1/3	y_1
2	+1	–1	+1	–1	1/3	1/3	y_2
3	+1	+1	–1	–1	1/3	1/3	y_3
4	+1	–1	–1	+1	1/3	1/3	y_4
5	+1	+1	0	0	1/3	–2/3	y_5
6	+1	–1	0	0	1/3	–2/3	y_6
7	+1	0	+1	0	–2/3	1/3	y_7
8	+1	0	–1	0	–2/3	1/3	y_8
9	+1	0	0	0	–2/3	–2/3	y_9

Коэффициенты уравнения регрессии:

$$\hat{y} = b_0^1 + b_1 x_1 + b_2 x_2 + \ldots + b_n x_n + b_{12} x_1 x_2 + \ldots + b_{n-1} x_{n-1} x_n + \\ + b_{11}(x_1^2 - \overline{x_1^2}) + \ldots + b_{nn}(x_n^2 - \overline{x_n^2}) \qquad (6.5)$$

по центральному композиционному ортогональному плану вычисляют в соответствии с формулами

$$b_0^1 = \frac{\sum_{u=1}^{N} y_u}{N} \qquad (6.6)$$

$$b_j = \frac{\sum_{u=1}^{N} x_{uj} y_u}{\sum_{u=1}^{N} x_{uj}^2} \qquad (6.7)$$

$$b_{ij} = \frac{\sum_{u=1}^{N} x_{ui} x_{uj} y_u}{\sum_{u=1}^{N} \left(x_{ui} x_{uj} \right)^2} \qquad (6.8)$$

$$b_{jj} = \frac{\sum_{u=1}^{N} x_{uj}^1 y_u}{\sum_{u=1}^{N} \left(x_{uj}^1 \right)^2} \qquad (6.9)$$

Дисперсии коэффициентов регрессии рассчитываются по формулам:

$$\left. \begin{aligned} D_{b_o^1}^* &= \frac{D_0^*}{N \cdot m} \\ D_{b_j}^* &= \frac{D_0^*}{\left(\sum_{u=1}^{N} x_{uj}^2 \right) \cdot m} \\ D_{b_{ij}}^* &= \frac{D_0^*}{\left[\sum_{u=1}^{N} \left(x_{ui} \cdot x_{uj} \right)^2 \right] \cdot m} \\ D_{b_{jj}}^* &= \frac{D_0^*}{\left[\sum_{u=1}^{N} \left(x_{uj}^1 \right)^2 \right] \cdot m} \end{aligned} \right\} \qquad (6.10)$$

где m – число параллельных опытов в каждой точке матрицы планирования (при их наличии).

D^*_0 – ошибка опыта (дисперсия воспроизводимости), оценивается также как и по полному факторному эксперименту;

$$D_0^* = \frac{\sum_{k=1}^{m_0} \left(y_{ok} - \bar{y}_0 \right)^2}{m_0 - 1}$$

индекс “0” означает центр плана.

Где $y_{ок}$ – значения параметра оптимизации в центре плана;

$\bar{y}_0$ – среднее значение параметра оптимизации в центре плана;

m_0 – число параллельных опытов в центре плана.

Т.е. коэффициенты регрессии оцениваются с разными ошибками.

$$t_{b_0' p} = \frac{b_0'}{\sigma_{b_0'}} > t_T; \quad t_{b_j p} = \frac{b_j}{\sigma_{b_j}} > t_T; \quad t_{b_{ij} p} = \frac{b_{ij}}{\sigma_{b_{ij}}} > t_T; \quad t_{b_{jj} p} = \frac{b_{jj}}{\sigma_{b_{jj}}} > t_T$$

Оценка значимости коэффициентов как и ранее проводится по критерию Стьюдента.

$t_т$ – коэффициент Стьюдента для $\upsilon_0 = m_0 - 1; \quad \alpha = 0{,}05$.

Если расчетное значение критерия Стьюдента меньше табличного, то коэффициент не значим.

Дисперсия величины b_0 оценивается по равенству

$$D^*_{b_0} = D^*_{b_o'} + \left(\overline{x_1^2}\right) D^*_{b_{11}} + \ldots + \left(\overline{x_n^2}\right) D^*_{b_{nn}} \tag{6.11}$$

Дисперсию адекватности определяют по формуле:

$$D^*_{ad.} = \frac{m}{N-l} \sum_{u=1}^{N} \left(\bar{y}_u - \hat{y}_u\right)^2 \tag{6.12}$$

ℓ - число значимых коэффициентов построенного уравнения регрессии.

$\hat{y}_u$ - значения выходной координаты, вычисленные по уравнению регрессии полученного с учетом значимых коэффициентов:

$$\hat{y} = b_0^1 + b_1 x_1 + b_2 x_2 + \ldots + b_n x_n + b_{12} x_1 x_2 + \ldots + b_{n-1} x_{n-1} x_n +$$
$$+ b_{11}(x_1^2 - \overline{x_1^2}) + \ldots + b_{nn}(x_n^2 - \overline{x_n^2})$$

(6.13)

От уравнения (6.5) переходят к уравнению (6.13) с помощью расчета величины свободного члена по формуле

$$b_0 = b_0' - b_{11} \overline{x_1^2} - \ldots - b_{nn} \overline{x_n^2} \tag{6.14}$$

где $\overline{x_j^2} = x_j^2 - x_j^1 \;\; x_{uj}^1 = x_{uj}^2 - \dfrac{\sum\limits_{u=1}^{N} x_{uj}^2}{N} = x_{uj}^2 - \overline{x_{uj}^2}$.

Адекватность уравнения регрессии проверяется по критерию Фишера :

$$F_P = \frac{D^*_{ad.}}{D^*_0} < F_T$$

F_T – коэффициент Фишера находится для $\begin{cases} \upsilon_{ad} = N - \ell \\ \upsilon_0 = m_0 - 1 \end{cases}$ $\alpha = 0{,}05$.

если $F_P < F_T$, то уравнение регрессии признается адекватным.

Дальнейшее совершенствование планов второго порядка привело (Бокс и Хантер) к построению рототабельных планов второго порядка, которые обеспечивают равные оценки дисперсий коэффициентов регрессии. При этом информация, содержащаяся в уравнении регрессии равномерно располагается по сфере в различных направлениях факторного пространства.

6.2. Алгоритм расчета

1. Кодирование факторов

$$x_i = \frac{X_i - X_{i0}}{\Delta X_i};$$

2. Производится преобразование квадратичных переменных:

$$x_{uj}^1 = x_{uj}^2 - \frac{\sum_{u=1}^{N} x_{uj}^2}{N} = x_{uj}^2 - \overline{x_{uj}^2}$$

и специальным образом выбирается величина звездного плеча (таблица 6.3).

3. Вычисляется дисперсия воспроизводимости.

$$D_0^* = \frac{\sum_{k=1}^{m_0} \left(y_{ok} - \bar{y}_0\right)^2}{m_0 - 1}$$

4. Рассчитываются коэффициенты уравнения регрессии:

$$\hat{y} = b_0' + b_1 x_1 + b_2 x_2 + \ldots + b_n x_n + b_{12} x_1 x_2 + \ldots + b_{n-1,n} x_{n-1} x_n + b_{11}\left(x_1^2 - \overline{x}_1^2\right) + \ldots + b_{nn}\left(x_n^2 - \overline{x}_n\right.$$

$$b_0^1 = \frac{\sum_{u=1}^{N} y_u}{N};\ b_j = \frac{\sum_{u=1}^{N} x_{uj} y_u}{\sum_{u=1}^{N} x_{uj}^2};\ b_{ij} = \frac{\sum_{u=1}^{N} x_{ui} x_{uj} y_u}{\sum_{u=1}^{N} \left(x_{ui} x_{uj}\right)^2};\ b_{jj} = \frac{\sum_{u=1}^{N} x_{uj}^1 y_u}{\sum_{u=1}^{N} \left(x_{uj}^1\right)^2}$$

5. Рассчитываются дисперсии коэффициентов уравнения регрессии:

$$\left.D_{b_o^1}^* = \frac{D_0^*}{N \cdot m}; D_{b_j}^* = \frac{D_0^*}{\left(\sum_{u=1}^{N} x_{uj}^2\right) \cdot m}; D_{b_{ij}}^* = \frac{D_0^*}{\left[\sum_{u=1}^{N} (x_{ui} \cdot x_{uj})^2\right] \cdot m}; D_{b_{jj}}^* = \frac{D_0^*}{\left[\sum_{u=1}^{N} \left(x_{uj}^1\right)^2\right] \cdot m}\right\}$$

(в данном эксперименте m=1).

6. Оценивается значимость коэффициентов уравнения регрессии по критерию Стьюдента:

$$t_{b_{0p}^1} = \frac{b_0^1}{\sigma_{b_0^1}} > t_T;\; t_{b_{jp}} = \frac{b_j}{\sigma_{b_j}} > t_T;\quad t_{b_{ijp}} = \frac{b_{ij}}{\sigma_{b_{ij}}} > t_T;\; t_{b_{jjp}} = \frac{b_{jj}}{\sigma_{b_{jj}}} > t_T$$

$t_т$ – коэффициент Стьюдента для $\upsilon_0 = m_0 - 1;\quad \alpha = 0{,}05$. (Приложение 3. Таблица 2).

7. Оценивается дисперсия свободного члена b_0:

$$D^*_{b_0} = D^*_{b'_o} + \left(\overline{x_1^2}\right)D^*_{b_{11}} + \ldots + \left(\overline{x_n^2}\right)D^*_{b_{nn}}$$

8. Проверяется адекватность построенного уравнения регрессии по критерию Фишера:

$$F_P = \frac{D^*_{ad.}}{D^*_0} < F_T$$

F_T находится для $\begin{cases} \upsilon_{ad} = N - \ell \\ \upsilon_0 = m_0 - 1 \end{cases}$ α = 0,05. (Приложение 2. Таблица 4).

$D^*_{ad.} = \dfrac{m}{N-l}\sum\limits_{u=1}^{N}\left(\bar{y}_u - \hat{y}_u\right)^2$; (в данном эксперименте m=1).

ℓ - число значимых коэффициентов построенного уравнения регрессии.

$\hat{y}_u$ - значения выходной координаты, вычисленные по уравнению регрессии полученного с учетом значимых коэффициентов:

$$\hat{y} = b'_0 + b_1x_1 + b_2x_2 + \ldots + b_nx_n + b_{12}x_1x_2 + \ldots + b_{n-1,n}x_{n-1}x_n + b_{11}\left(x_1^2 - \overline{x}_1^2\right) + \ldots$$
$$+ b_{nn}\left(x_n^2 - \overline{x}_n^2\right)$$

6.3. Последовательность выполнения задания

1. Выписать в рабочую тетрадь все необходимые для расчета факторного эксперимента формулы.
2. Выполнить задание согласно номера варианта.

6.4. Задание

2. Заполнить матрицу планирования для парных взаимодействий и модифицированных квадратичных членов:

$x_1 x_2$	$x_1 x_3$	$x_1 x_4$	$x_2 x_3$	$x_2 x_4$	$x_3 x_4$	x_1^1	x_2^1	x_3^1	x_4^1

3. Представить зависимость параметра оптимизации от факторов (уравнение регрессии) в виде полинома второго порядка.
4. Выполнить все этапы расчета факторного эксперимента.

- произвести преобразование квадратичных переменных и специальным образом выбрать величину звездного плеча (Приложение Г. Таблица 1);
- вычислить дисперсию воспроизводимости;
- рассчитать коэффициенты уравнения регрессии:

$$\hat{y} = b_0' + b_1 x_1 + b_2 x_2 + \ldots + b_n x_n + b_{12} x_1 x_2 + \ldots + b_{n-1,n} x_{n-1} x_n + b_{11}\left(x_1^2 - \overline{x}_1^2\right) + \ldots + b_{nn}\left(x_n^2 - \overline{x}_n^2\right);$$

- рассчитать дисперсии коэффициентов уравнения регрессии (в данном эксперименте m=1);
- оценить значимость коэффициентов уравнения регрессии по критерию Стьюдента;
- оценить дисперсию свободного члена b_0;
- проверить адекватность построенного уравнения регрессии по критерию Фишера.

5. Проанализировать полученную полиномиальную модель зависимости степени разложения боратов смесью серной и фосфорной кислот (характер влияния факторов на целевую функцию).
6. Записать полученное уравнение регрессии относительно факторов в натуральных величинах.

Примечание:

Табличное значение коэффициента Стьюдента. (Приложение В. Таблица 2).

Табличное значение коэффициента Фишера. (Приложение В. Таблица 3).

Библиографический список

1. Вентцель Е.С., Овчаров Л.А. Теория вероятностей и ее инженерные приложения. М., Наука, 1988. – 480 с.
2. Вентцель Е.С., Овчаров Л.А. Теория случайных процессов и ее инженерные приложения. М, Наука, 1991. – 384 с.
3. Ицкович Э.Л. Статистические методы при автоматизации производства. М–л, Энергия, 1964. – 192 с.
4. Ицкович Э.Л. Контроль производства с помощью вычислительных машин. М., Энергия, 1975. – 416 с.
5. Закс Л. Статистическое оценивание. М., Статистика, 1976. – 598 с.
6. Дейч А.М. Методы идентификации динамических объектов. М., Энергия, 1979. – 240 с.
7. Анисимов С.А. и др. Типовые линейные модели объектов управления. Под ред. Н.С. Райбмана. М., Энергоатомиздат, 1983. – 264 с.
8. Бондарь А.Г., Статюха Г.А. Планирование эксперимента в химической технологии. Киев, “Вища школа”, 1976. – 183 с.
9. Бондарь А.Г., Статюха Г.А., Потяженко И.А. Планирование эксперимента при оптимизации процессов химической технологии. Киев, “Вища школа”, 1980. – 263 с.
10. Астапов Ю.М., Медведев В.С. Статистическая теория систем автоматического регулирования и управления. М., Наука, 1982. – 304 с.
11. Дядик В.Ф., Байдали Т.А. Учебно-методическое пособие к лабораторному практикуму по курсу "Статистические методы контроля и управления". Томск, ТПУ, 2004.
12. Дядик В.Ф., Байдали Т.А. Сборник заданий к выполнению лабораторных работ по курсу "Статистические методы контроля и управления". Томск, ТПУ, 2006.
13. http://e-le.lcg.tpu.ru. Login: Student_KEA, password: student_KEA. Дядик В.Ф. Статистические методы контроля
14. Бендат Д., Пирсол А. Прикладной анализ случайных данных. М.. Мир, 1989. – 540 с.
15. Егоров А.Е., Азаров Г.Н., Коваль А.В. Исследование устройств и систем автоматики методом планирования эксперимента. Харьков, “Вища школа”, 1986. – 238 с.
16. Рузинов Л.П., Слободчикова Р.И.. Планирование эксперимента в химии и химической технологии. М., Химия, 1980. – 280 с.

17. Реймаров Г.А. Первичная переработка информации в АСУ ТП. Учебно-методическое пособие. М., ЦНИИатоминформ, 1980.– 95 с.
18. Райбман Н.С. и Чадеев В.М. Построение моделей процессов производства. М., Энергия, 1975.

ПРИЛОЖЕНИЕ А

Таблица 1 – Распределение Стьюдента при вероятности ошибки α для двухстороннего критерия, равной 0,05

Число степеней свободы	$t_{ст}$	Число степеней свободы	$t_{ст}$
1	12,706	23	2,069
2	4,303	24	2,064
3	3,182	25	2,060
4	2,776	26	2,056
5	2,571	27	2,052
6	2,447	28	2,048
7	2,365	29	2,045
8	2,306	30	2,042
9	2,262	35	2,030
10	2,228	40	2,021
11	2,201	45	2,014
12	2,179	50	2,009
13	2,160	60	2,000
14	2,145	70	1,994
15	2,131	80	1,990
16	2,120	90	1,987
17	2,110	100	1,984
18	2,101	120	1,980
19	2,093	200	1,972
20	2,086	500	1,965
21	2,080	1000	1,962
22	2,074	∞	1,960

Продолжение приложения А

Таблица 2 – X^2 – распределение (β = 0,95)

Число степеней свободы	α	$\alpha/2$	$1-\frac{\alpha}{2}$
1	3,8	5,02	0,00098
2	6,0	7,38	0,0506
3	7,8	9,35	0,216
4	9,5	11,14	0,484
5	11,1	12,83	0,831
6	12,6	14,45	1,24
7	14,1	16,01	1,69
8	15,5	17,53	2,18
9	16,9	19,02	2,70
10	18,3	20,48	3,25
11	19,7	21,92	3,82
12	21,0	23,34	4,40
13	22,4	24,74	5,01
14	23,7	26,12	5,63
15	25,0	27,49	6,26
16	26,3	28,35	6,91
17	27,6	30,19	7,56
18	28,9	31,53	8,23
19	30,1	32,85	8,91
20	31,4	34,17	9,59
22	33,9	36,78	10,98
24	36,4	39,36	12,40
26	38,9	41,92	13,84
28	41,3	44,46	15,31
30	43,8	46,98	16,79
40		59,34	24,43
50		71,42	32,36
60		83,30	40,48
80		106,63	57,15
100		129,56	74,22

ПРИЛОЖЕНИЕ Б.

Доверительные границы для коэффициента корреляции

95%-ый доверительный интервал для **r** определяется по рис. 1. Как расстояние между точками пересечения вертикали, соответствующей значению **r***, с кривыми, соответствующей значению **n** . Если доверительный интервал не включает значение **r** = 0 , то можно говорить о наличии корреляции (**r** ≠ 0).

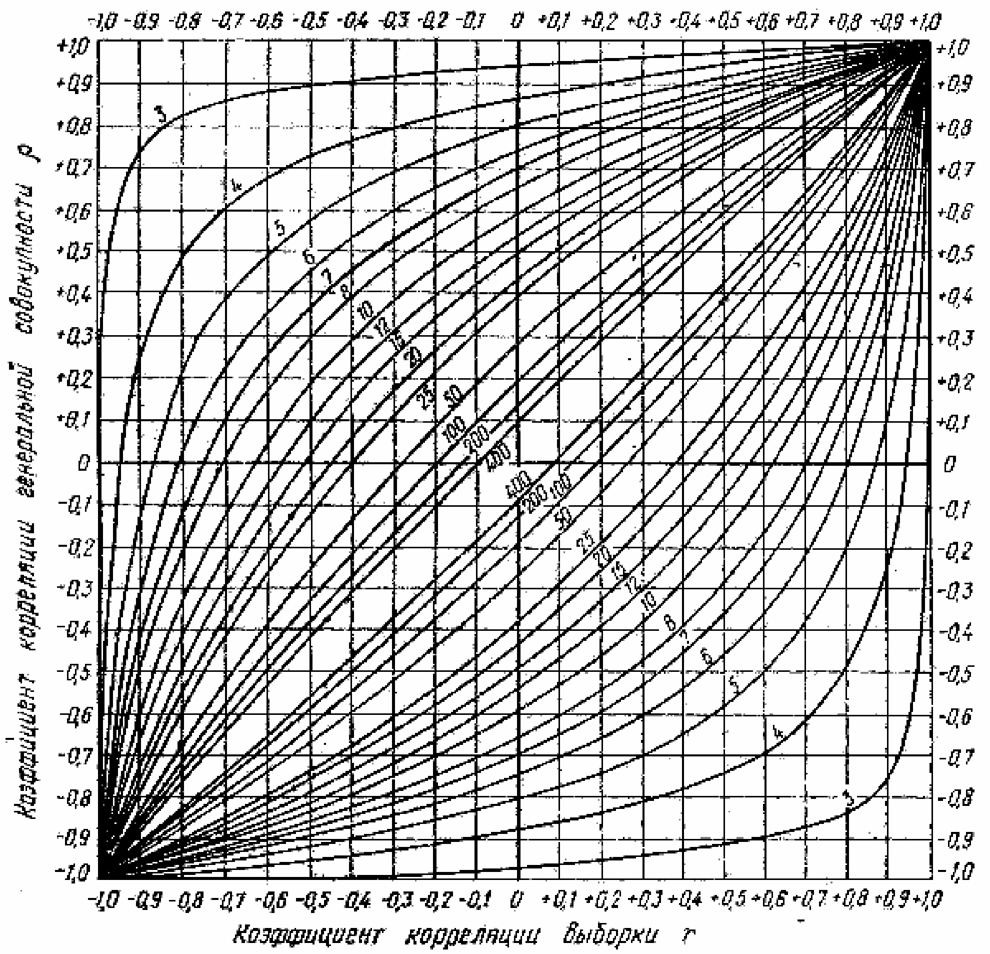


Рис. 1. Доверительные границы для коэффициента корреляции: 95%-ый доверительный интервал для **r**; числа на кривых обозначают объем выборки.

ПРИЛОЖЕНИЕ В

Таблица 1 – Критерий Кохрена при q=0,05

$\upsilon_1 \rightarrow$ $\upsilon_2 \downarrow$	1	2	3	4	5	6	7	8	9	10	16	36	144	∞
2	0,9985	0,9750	0,9392	0,9057	0,8584	0,8534	0,8332	0,8159	0,8010	0,7880	0,7341	0,6602	0,5813	0,5000
3	0,9669	0,8709	0,7977	0,7457	0,7071	0,6771	0,6530	0,6333	0,6167	0,6025	0,5466	0,4748	0,4031	0,3333
4	0,9065	0,7679	0,6941	0,6287	0,5895	0,5598	0,5365	0,5175	0,5017	0,4884	0,4366	0,3720	0,3093	0,2500
5	0,8412	0,6838	0,5981	0,5440	0,5063	0,4783	0,4564	0,4387	0,1241	0,4118	0,3645	0,3066	0,2513	0,2000
6	0,7808	0,6161	0,5321	0,4803	0,4447	0,4184	0,3980	0,3817	0,3682	0,3508	0,3135	0,2612	0,2119	0,1667
7	0,7271	0,5612	0,4800	0,4307	0,3907	0,3726	0,3555	0,3384	0,3254	0,3154	0,2756	0,2277	0,1833	0,1429
8	0,6798	0,5157	0,4377	0,3910	0,3595	0,3362	0,3185	0,3043	0,2926	0,2829	0,2462	0,2022	0,1616	0,1250
9	0,6385	0,4775	0,4027	0,3584	0,3286	0,3067	0,2901	0,2768	0,2659	0,2568	0,2226	0,1820	0,1446	0,1111
10	0,6020	0,4450	0,3733	0,3311	0,3029	0,2823	0,2606	0,2541	0,2439	0,2353	0,2032	0,1655	0,1308	0,1000
12	0,5410	0,3924	0,3204	0,2880	0,2624	0,2439	0,2299	0,2187	0,2098	0,2020	0,1737	0,1403	0,1100	0,0833
15	0,4709	0,3346	0,2758	0,2419	0,2195	0,2034	0,1911	0,1815	0,1736	0,1671	0,1429	0,1144	0,0889	0,0677
20	0,3894	0,2705	0,2205	0,1921	0,1735	0,1602	0,1501	0,1422	0,1357	0,1303	0,1108	0,0879	0,0675	0,0500
24	0,3434	0,2354	0,1907	0,1656	0,1493	0,1374	0,1286	0,1216	0,1160	0,1113	0,0942	0,0743	0,0567	0,0417
30	0,2929	0,1980	0,1593	0,1377	0,1237	0,1137	0,1061	0,1002	0,0958	0,0921	0,0771	0,0604	0,0457	0,0333
40	0,2370	0,1576	0,1259	0,1082	0,0958	0,0887	0,0827	0,0780	0,0745	0,0713	0,0595	0,0462	0,0347	0,0250
60	0,1737	0,1131	0,0895	0,0766	0,0682	0,0623	0,0583	0,0552	0,0520	0,0497	0,0411	0,0316	0,0234	0,0167
120	0,0998	0,0632	0,0495	0,0419	0,0371	0,0337	0,0312	0,0292	0,0279	0,0266	0,0218	0,0165	0,0120	0,0083
∞	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000

Продолжение приложения В

Таблица 2 – Критерий Стьюдента

q	10%	5%	2%	1%	q	10%	5%	2%	1%
1	6,31	12,71	31,82	63,66	18	1,73	2,10	2,55	2,88
2	2,92	4,30	6,96	9,92	19	1,73	2,09	2,54	2,86
3	2,35	3,18	4,54	5,84	20	1,72	2,09	2,53	2,85
4	2,13	2,78	3,75	4,6	21	1,72	2,08	2,52	2,83
5	2,02	2,57	3,36	4,03	22	1,72	2,07	2,51	2,82
6	1,94	2,45	3,14	3,71	23	1,71	2,07	2,50	2,81
7	1,89	2,31	3,00	3,50	24	1,71	2,06	2,49	2,80
8	1,86	2,26	2,90	3,36	25	1,71	2,06	2,49	2,79
9	1,83	2,36	2,82	3,25	26	1,71	2,06	2,48	2,78
10	1,81	2,23	2,76	3,17	27	1,70	2,05	2,47	2,77
11	1,80	2,20	2,72	3,11	28	1,70	2,05	2,47	2,76
12	1,78	2,18	2,68	3,05	29	1,70	2,05	2,46	2,76
13	1,77	2,16	2,65	3,01	30	1,70	2,04	2,46	2,75
14	1,76	2,14	2,62	2,98	40	1,68	2,02	2,42	2,70
15	1,75	2,13	2,60	2,95	60	1,67	2,00	2,39	2,66
16	1,75	2,12	2,58	2,92	120	1,66	1,98	2,36	2,62
17	1,74	2,11	2,57	2,90	∞	1,64	1,96	2,33	2,58

Продолжение приложения В

Таблица 3 – Критерий Фишера при q=0,05

$\upsilon_{ад}\rightarrow$ $\upsilon_0\downarrow$	1	2	3	4	5	6	8	12	24	∞
1	161,4	199,5	215,7	224,6	230,2	234,0	238,9	243,9	249,0	254,3
2	18,51	19,00	19,16	19,25	19,30	19,33	19,37	19,41	19,45	19,50
3	10,13	9,55	9,28	9,12	9,01	8,94	8,84	8,74	8,64	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,04	5,91	5,77	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,82	4,68	4,53	4,36
6	5,99	5,14	4,76	4,53	4,39	4,28	4,15	4,00	3,84	3,67
7	5,39	4,74	4,35	4,12	3,97	3,87	3,73	3,57	3,41	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,44	3,28	3,12	2,93
9	5,12	4,26	3,85	3,63	3,48	3,37	3,23	3,07	2,90	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,07	2,91	2,74	2,54
11	4,84	3,98	3,59	3,36	3,20	3,09	2,95	2,79	2,61	2,40
12	4,75	3,88	3,49	3,26	3,11	3,00	2,85	2,69	2,50	2,30
13	4,67	3,80	3,41	3,18	3,02	2,92	2,77	2,60	2,42	2,21
14	4,60	3,74	3,34	3,11	2,96	2,85	2,70	2,53	2,35	2,13
15	4,54	3,68	3,29	3,06	2,90	2,79	2,64	2,48	2,29	2,07
16	4,49	3,63	3,24	3,01	2,85	2,74	2,59	2,42	2,24	2,01
17	4,45	3,59	3,20	2,96	2,81	2,70	2,55	2,38	2,19	1,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,51	2,34	2,15	1,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,48	2,31	2,11	1,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,45	2,28	2,08	1,84
21	4,32	3,47	3,07	2,84	2,68	2,57	2,42	2,25	2,05	1,81
22	4,30	3,44	3,05	2,82	2,66	2,55	2,40	2,23	2,03	1,78
23	4,28	3,42	3,03	2,80	2,64	2,53	2,38	2,20	2,00	1,76
24	4,26	3,40	3,01	2,78	2,62	2,51	2,35	2,18	1,98	1,73
25	4,24	3,38	2,99	2,76	2,60	2,49	2,34	2,16	1,96	1,71
26	4,22	3,37	2,98	2,74	2,59	2,47	2,32	2,15	1,95	1,69
27	4,21	3,35	2,96	2,73	2,57	2,46	2,30	2,13	1,93	1,67
28	4,20	3,34	2,95	2,71	2,56	2,44	2,29	2,12	1,91	1,65
29	4,18	3,33	2,93	2,70	2,54	2,43	2,28	2,10	1,90	1,64
30	4,17	3,32	2,92	2,69	2,53	2,42	2,27	2,09	1,89	1,62
40	4,08	3,23	2,84	2,61	2,45	2,34	2,18	2,00	1,79	1,52
60	4,00	3,15	2,76	2,52	2,37	2,25	2,10	1,92	1,70	1,39
120	3,92	3,07	2,68	2,45	2,29	2,17	2,02	1,83	1,61	1,25
∞	3,84	2,99	2,60	2,37	2,21	2,09	1,94	1,75	1,52	1,00

ПРИЛОЖЕНИЕ Г

Таблица 1 – Значения α и общее число опытов

Число опытов	Число независимых переменных			
	2	3	4	5*
В ядре плана $N_я$	4	8	16	16
В “звездных” точках N_α	4	6	8	10
В центре плана m_0	1	1	1	1
Общее число опытов N	9	15	25	27
Величина “звездного” плеча α	1,00	1,215	1,414	1,547

Учебное издание

ДЯДИК Валерий Феодосиевич
БАЙДАЛИ Тамара Алексеевна
БАЙДАЛИ Сергей Анатольевич

СТАТИСТИЧЕСКИЕ МЕТОДЫ КОНТРОЛЯ И УПРАВЛЕНИЯ

Учебно-методическое пособие

Научный редактор
д-р физ.-мат. наук,
профессор *И.В. Шаманин*

Подписано к печати 21.11.2008. Формат 60x84/8. Бумага «Снегурочка».
Печать XEROX. Усл.печ.л. 000. Уч.-изд.л. 000.
Заказ XXX. Тираж XXX экз.

Томский политехнический университет
Система менеджмента качества
Томского политехнического университета
сертифицирована
NATIONAL QUALITY ASSURANCE по стандарту ISO 9001:2000

ИЗДАТЕЛЬСТВО ТПУ. 634050, г. Томск, пр. Ленина, 30.