



МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
РОССИЙСКОЙ ФЕДЕРАЦИИ

КЕМЕРОВСКИЙ ТЕХНОЛОГИЧЕСКИЙ
ИНСТИТУТ
ПИЩЕВОЙ ПРОМЫШЛЕННОСТИ

Кафедра «Технологическое проектирование пищевых
производств»

**ОСНОВЫ МАТЕМАТИЧЕСКОГО АНАЛИЗА
ТЕХНОЛОГИЧЕСКИХ ПРОЦЕССОВ**

УЧЕБНОЕ ПОСОБИЕ

Кемерово 2017

УДК 664.62(075)

ББК 36.81 – 5а7

Б 83

Составители:

Бородулин Д.М., д-р техн.наук, доцент

Сухоруков Д.В., к-т техн.наук, старший преподаватель

Рецензенты:

П.Т. Петрик, доктор технических наук, профессор, ФГБОУ ВПО «Кузбасский государственный технический университет», заведующий кафедрой «Энергоресурсосберегающие процессы в химической и нефтегазовой технологиях»

С.Л. Тихонов, доктор технических наук, профессор, ФГБОУ ВПО «Уральский государственный экономический университет», заведующий кафедрой пищевой инженерии

*Рекомендовано редакционно-издательским советом
Кемеровского технологического института пищевой
промышленности (университета)*

Б 83 Основы математического анализа технологических процессов: учебное пособие / [Д.М. Бородулин, Д.В. Сухоруков], Кемеровский технологический институт пищевой промышленности (университет). - Кемерово, 2017. - 55 с.

ISBN 978-5-89289-987-1

В данном учебном пособии описаны статистические методы анализа экспериментальных данных

Предназначено для магистрантов, обучающихся по направлению подготовки 15.04.02. «Технологические машины и оборудование» по программе «Процессы и аппараты пищевых производств»

ISBN 978-5-89289-987-1

*Охраняется законом об авторском
праве, не может быть использовано
любым незаконным способом
без письменного договора*

КемТИПП, 2017

Введение

Экспериментальное исследование является необходимым этапом процесса познания. Имея дело с различными процессами и объектами в своей деятельности, человек постепенно накапливает практические сведения об этих процессах и объектах, что помогает ему познать их и на основе этого знания усовершенствовать, оптимизировать. Зачастую сведений явно недостаточно, чтобы сделать какие-либо выводы. Тогда тот или иной объект, процесс подвергают планомерному экспериментальному исследованию.

1. Статические методы анализа экспериментальных данных

1.1 Критерий оптимальности, факторы, ошибки опыта

Экспериментом называется совокупность опытов, объединенных целью, единой системой ограничений в пространстве и времени (i – номер опыта, $i = 1/N$).

Опытом можно считать реализацию на каком-либо объекте некоторых условий, правил. В результате опыта появляется то или иное **событие** (например, событием будет накопление фрагмента в результате жизнедеятельности культивируемого микроорганизма). Появление события регистрируется при помощи какого-либо параметра, имеющего, как правило, численное выражение и наиболее полно характеризующего эффективность исследуемого процесса в i -ом опыте с позиции цели исследования. Такой параметр обычно обозначается буквой x_i и называется **критерием оптимальности**.

Всеобщим критерием оптимальности является экономический критерий. **Частным** критерием оптимальности может быть любой технологический параметр (накопление фермента, биомассы, освобождение препарата от сопутствующей нежелательной активности и т.д.) – такой критерий обычно называют **выходом процесса**.

Результат процесса зависит от условий его протекания, характеризующихся значениями параметров, влияющих на процесс (например, температуры, давления, величины рН и т.д.) эти параметры называют **факторами**. Численное значение любого фактора X должно устанавливаться и реализовываться независимо от значений других факторов (i – номер фактора, $i = 1/N$).

Условия проведения каждого опыта эксперимента в виде конкретного значения исследуемых факторов, а иногда и очередность проведения опытов регламентируется **планом эксперимента**. По результатам опытов, следуя определенному алгоритму, получают соответствующее **уравнение**, характеризующее влияние факторов на эффективность исследуемого процесса.

При выборе типа плана следует учитывать его возможные недостатки или преимущества перед другими типами планов, имеющих одно или тоже число опытов, один и тот же диапазон изменения факторов, но разное расположение экспериментальных точек.

Планы, обеспечивающие получение наименьшей величины максимальной дисперсии предсказания, называют **G-оптимальными**.

Планы, обеспечивающие получение одинаковых дисперсий предсказания для точек, равноотстоящих от центра эксперимента, называют **ротабельными**.

Планы, обеспечивающие получение одинаковой величины дисперсии предсказания для любой точки в пределах изучаемой области, называются **униформ-ротабельными**.

D-оптимальными называют планы, обеспечивающие получение минимальной обобщенной дисперсии коэффициентов уравнения.

Планы, обеспечивающие получение уравнения, коэффициенты которого ортогональны друг другу (независимы друг от друга) и имеют одинаковую дисперсию, называются **ортогональными**.

Составители планов находят зависимость этих свойств от расположения экспериментальных точек и оптимизируют по заданному свойству параметры плана. Для начинающих осваивать планирование привлекательными обычно являются ортогональные планы, позволяющие получить уравнение с коэффициентами, характеризующими лишь те эффекты, которые описывает данный коэффициент.

Оптимизация процесса (объекта) сводится к отысканию таких условий его протекания (таких конструктивных размеров, технологических параметров и т.д.), при которых критерий оптимальности будет иметь экспериментальное значение: максимум, если увеличение численного значения критерия оптимальности означает увеличение эффективности процесса (объекта), или минимум, если его увеличение означает снижение эффективности процесса (объекта).

Очевидно, исследователь должен быть вооружен приборами для измерения критерия оптимальности или его составляющих, для измерения и стабилизации величин исследуемых факторов. Последнее требование связано с тем, что опыт проводится при определенных заранее заданных условиях.

Изучение всех влияющих на исследуемый объект факторов одновременно провести невозможно, поэтому в эксперименте рассматриваются их ограниченное число. Остальные активные факторы стабилизируются, т.е. устанавливаются на каких-то одинаковых для всех опытов уровнях.

Некоторые фрагменты не могут быть обеспечены системами стабилизации (например, погодные условия, самочувствие оператора и т.д.), другие же стабилизируются с какой-то погрешностью (например, содержание какого-либо компонента в среде зависит от ошибки при взятии навески и приготовления раствора). Учитывая также, что изменение параметра x осуществляется прибором, обладающим какой-то погрешностью, зависящей от класса точности прибора, можно прийти к выводу, что результаты повторностей одного и того же опыта x_s будут приближенными и должны отличаться один от другого и от истинного значения выхода процесса.

Неконтролируемое, случайное изменение и множества других влияющих на процесс факторов вызывает **случайные** отклонения измеряемой величины x_s от ее истинного значения. Поэтому результаты повторностей одного и того же опыта $s=1/m$ образуют набор случайных величин, анализ точности и достоверности которых осуществляется методами математической статистики и теории вероятности. Отличие полученных результатов повторностей опыта x_s от истинного результата x характеризуется величиной абсолютной погрешности

$$\Delta_s = x_s - x . \quad (1.1)$$

Несовпадение результатов может вызываться и другими чрезвычайными причинами. Ими могут быть неквалифицированные действия оператора, неисправность измерительных приборов, резкое нарушение основных условий протекания процесса и т.д. Такие результаты выделяются из остальных по величине, и их называют **грубыми** ошибками. Их появление может исказить нормальность закона распределения вероятности случайной величины x_s и привести к ошибочным выводам. Грубые ошибки следует исключить из массива экспериментальных данных, доказав предварительно правоту такого действия.

В исследовательской практике может возникнуть необходимость убедиться в надежности применяемой методики и исправности приборов. В этом случае проводят серию контрольных опытов, в которых осуществляют процесс с известным заранее результатом x (эталонный процесс). Если получаемые в контрольном опыте результаты x_s смещены относительно x преимущественно в одну сторону, то такие результаты искажены т.н. **систематической ошибкой**. Оценкой систематической ошибки может быть средняя арифметическая величина абсолютного отклонения результатов всех повторностей какого-либо опыта:

$$\bar{\Delta}_s = \frac{1}{m} \sum_{s=1}^m (x_s - x) \quad (1.2)$$

Если выявить и исключить причины появления систематической ошибки не удается, то в середине результаты:

$$\bar{x}'_s = \bar{x}_s + \bar{\Delta}, \quad (1.3)$$

Где

$$\bar{x}_s = \frac{1}{m_s} \sum_{s=1}^m x_s \quad (1.4)$$

1.2 Распределение случайной величины

Процедура статистического анализа экспериментальных данных основывается на предположении о том, что эти данные являются случайной величиной, распределенной по нормативному закону.

Под **распределением случайной величины** понимают совокупность всех возможных ее значений и соответствующих им вероятностей появления этих значений, формализуется распределение в виде **закона распределения случайной величины**.

В соответствии с теоремой А. М. Ляпунова (1901 г.) случайная величина x_s будет распределена по нормальному закону, если она подвержена влиянию большого числа неучтенных, влияющих на процесс факторов, неконтролируемое (случайное) изменение которых вызывает малое и соразмерное для каждого фактора изменение измеряемого параметра x_s .

При экспериментальном исследовании какого-либо технологического процесса измеряемый результат последнего является случайной величиной, на которую оказывает влияние огромное число факторов (изменение погодных условий, самочувствие оператора, неоднородность сырья, влияние износа измерительной и стабилизирующей аппаратуры и т.д. и т.п.). Именно поэтому результат исследования является случайной величиной, распределенной по нормальному закону. Однако если исследователь какой-либо активный фактор не заметил или отнес его к неактивным, а неконтролируемое изменение этого фактора может вызвать несоразмерно большое изменение эффективности процесса и параметра, характеризующего эту эффективность, то распределение вероятности последнего может нормальному закону не подчиниться.

Точно также приведет к нарушению нормальности закона распределения наличие в массиве экспериментальных данных грубых ошибок. Именно поэтому в первую очередь проводят анализ на наличие в экспериментальных данных грубых ошибок с принятой доверительной вероятностью.

При достаточно большом числе повторностей ($m \rightarrow \infty$) нормальный закон распределения проявляется в том, что:

1) Абсолютные отклонения одинаковой величины, но разного знака встречаются одинаково часто;

2) Большие по модулю абсолютные отклонения встречаются реже, чем меньшие, т.е. вероятность появления той или иной погрешности уменьшается с увеличением модуля этой погрешности.

В соответствии с (1) для всех m повторностей можно записать:

$$\sum_{s=1}^m \Delta_s = \sum_{s=1}^m x_s - mx \quad (1.5)$$

Разделив полученное уравнение на число повторностей m , имеем:

$$x = \bar{x} - \frac{1}{m} \sum_{s=1}^m \Delta x_s \quad (1.6)$$

За экспериментальную оценку истинного значения (математического ожидания) критерия оптимальности x принимается **среднеарифметическая оценка** результатов всех m повторностей:

$$\bar{x} = \frac{1}{m} \sum_{s=1}^m x_s \quad (1.7)$$

Если случайная величина является распределенной по нормальному закону, то при достаточно большом числе повторностей ($m \rightarrow \infty$) будет справедливо равенство:

$$\lim_{m \rightarrow \infty} \left(\frac{1}{m} \sum_{s=1}^m x_s \right) = 0 \quad (1.8)$$

Иными словами, лишь при $m \rightarrow \infty$ среднеарифметическое x .

Как уже отмечалось, более часто будут встречаться меньшие по модулю абсолютные отклонения, т.е. их появление более вероятно, и будет соответствовать величине плотности распределения вероятности появления этих результатов:

$$P_x = \frac{dP}{dx_s} \quad (1.9)$$

Плотность нормального распределения (рис. 1.1), во-первых, симметрична относительно x , во-вторых, достигает максимального значения при $x_s=x$ и, в-третьих, стремится к нулю при увеличении $|\Delta x_s| = |x_s - x|$.

Функция плотности нормального распределения случайной величины (функция Гаусса) задается двумя параметрами: истинным значением x и **среднеквадратичным отклонением σ** :

$$P_x = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x_s-x)^2}{2\sigma^2}\right] \quad (1.10)$$

Квадрат среднеквадратичного отклонения σ^2 называется **дисперсией** случайной величины и является количественной характеристикой разброса результатов x_s вокруг истинного значения x .

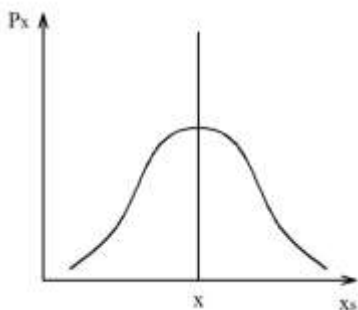


Рис. 1.1 – Кривая плотности нормального распределения

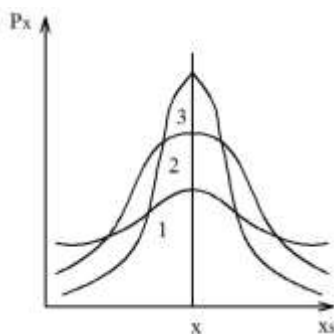


Рис. 1.2 – Кривая плотности вероятности нормально распределенной случайной величины при различных дисперсиях

На рис. 1.2 изображены кривые плотности нормального распределения при различных значениях дисперсии, причем σ_1^2

› σ^2_2 › σ^2_3 . Очевидно, кривая 3 характеризует плотность распределения случайной величины, воспроизводимость которой в повторных измерениях лучше, чем воспроизводимость случайных величин, имеющих плотность распределения 1 и 2.

1.3 Экспериментальные оценки истинного значения измеряемой случайной величины и ее дисперсии.

Если в распоряжении исследователя находится конечное число независимых результатов повторностей одного и того же опыта, то он может получить лишь экспериментальные оценки истинного значения результата опыта x в виде единичных оценок x_s и среднеарифметического результата, $S^2(x_s)$.

Следует стремиться к тому, чтобы оценки обладали следующими свойствами:

1. Несмещенностью, проявляющейся в том, что теоретическое среднее из генеральной совокупности результатов совпадает с истинным значением измеряемого параметра;
2. Состоятельностью, когда оценки при неограниченном увеличении числа измерений могут иметь сколь угодно малый интервал при вероятности, стремящийся к единице;
3. Эффективностью, проявляющейся в том, что из всех несмещенных оценок данная оценка будет иметь наименьшее рассеяние (наименьшую дисперсию).

Среднеарифметическое значение \bar{x} , если x_s является случайной величиной, распределенной по нормальному закону, будет несмещенной, состоятельной и эффективной оценкой истинного значения x .

Экспериментальная оценка среднеквадратичного отклонения обозначается буквой S с указанием в скобках символа анализируемой величины. Например, $S(x_s)$ – оценка среднеквадратичного отклонения единичного результата $S(\bar{x})$ – оценка среднеквадратичного отклонения среднего результата.

Квадрат экспериментальной оценки среднеквадратичного отклонения, S^2 является **экспериментальной оценкой дисперсии**. Если истинное

значение x известно, то по m повторностям x_s оценку дисперсии находят как среднее из квадратов абсолютных отклонений:

$$S^2(x_s) = \frac{\sum_{s=1}^m (x_s - x)^2}{m} \quad (1.11)$$

Если по m результатам рассчитывают оценку истинного значения \bar{x} по формуле 1.11, а затем, используя те же результаты, рассчитывают оценки абсолютных отклонений:

$$\Delta_s = x_s - \bar{x}, \quad (1.12)$$

То оценку дисперсии единичного результата находят по зависимости:

$$S^2(x_s) = \frac{1}{m-1} \sum_{s=1}^m (x_s - \bar{x})^2 \quad (1.13)$$

Разность между числом m независимых результатов x_s и числом уравнений, в которых эти результаты уже были использованы для расчета неизвестных оценок (в рассматриваемом случае полученные результаты использовались лишь для расчета \bar{x}), называют **числом степеней свободы f** :

$$f = m - 1 \quad (1.14)$$

Для оценки дисперсии эталонного процесса $f = m$.

Поскольку средняя оценка \bar{x} является более точной, чем единичная x_s , дисперсия (разброс) средних будет меньше дисперсии единичных результатов.

В статистике доказывается, что оценка дисперсии среднего результата будет меньше оценки единичного в m раз, если \bar{x} рассчитано по всем m единичным результатам x_s , теорема о дисперсии среднеарифметической величины

$$S^2(\bar{x}) = \frac{1}{m} S^2(x_s) \quad (1.15)$$

Примеры и контрольные задачи

Прежде чем приступить к рассмотрению изложенного ниже иллюстративного материала и, тем более к решению контрольных задач, следует основательно ознакомиться с последней главой, посвященной некоторым правилам работы с приближенными числами при их использовании в различных вычислительных операциях.

К приближенным числам следует отнести результаты прямых измерений, а также результаты. Полученные при вычислении каких-либо функций от них.

Пример 1. Рассчитать оценку дисперсии единичного $S^2(x_s)$ и среднего $S^2(\bar{x})$ результатов по $m = 9$ повторностям опыта.

s	1	2	3	4	5	6	7	8	9	Σ
$X_s, \%$	35	36	37	40	36	39	35	38	37	333
$\Delta_s, \%$	-2,0	-1,0	0,0	+3,0	-1,0	+2,0	-2,0	+1,0	0,0	0,0
$\Delta_s, (\%)$	4,0	1,0	0,0	9,0	1,0	4,0	4,0	1,0	0,0	24,0

Решение:

1. Рассчитаем средний результат опыта:

$$\bar{x} = \frac{1}{m} \sum_{s=1}^m x_s = \frac{333}{9} = 37,0\%.$$

2. Рассчитаем абсолютные отклонения:

$$\Delta_1 = x_1 - \bar{x} = 35 - 37,0 = 2,0 \text{ и т.д. (см. таблицу).}$$

Проверка:

$$\sum_{s=1}^m \Delta_s = -6,0 + 6,0 = 0 - \text{ошибок нет.}$$

$$\Delta_1^2 = 2,0^2 = 4,00 = 4,0 (\%)^2 \text{ (см. таблицу).}$$

3. Рассчитаем оценки дисперсии:

$$S^2(x_s) = \frac{\sum_{s=1}^m \Delta_s^2}{m-1} = \frac{24,0}{9-1} = 3,00 (\%)^2.$$

$$S^2(\bar{x}) = \frac{S^2(x_s)}{m} = \frac{3,00}{9} = 0,333 (\%)^2.$$

4. Число степеней свободы:

$$f = m - 1 = 9 - 1 = 8.$$

Ответ:

$$S^2(x_s) = 3,00(\%)^2, f = 8, S^2(\bar{x}) = 0,333 (\%)^2.$$

Пример 2 В таблице представлены результаты реализации эталонного процесса $x=38,4$ мг\мл.

Найти оценки дисперсии $S^2(x_s)$ и $S^2(\bar{x})$

s	1	2	3	4	5	6	7	8	Σ
Δs мг/мл	9,6	7,2	8,4	9,2	6,8	7,4	8,8	9,0	06,4
Δs мг/мл	+1,	-1,2	0,0	+0,	-1,6	-1,0	+0,	+0,	-
Δs (мг/мл) ²	1,4	1,4	0,0	0,6	2,6	1,0	0,2	0,4	7,6

Решение:

1. Абсолютные отклонения:

$$\Delta_1 - x = 39,6-38,4=1,2 \text{ мг/мл (см. таблицу)}$$

$$2. \quad S^2(x_s) = \frac{\sum_{s=1}^m (x_s - \bar{x})^2}{m} = \frac{7,6}{8} = 0,95 (\text{мг/мл})^2$$

$$3. \quad S^2(\bar{x}) = S^2(x_s)/m = 0,95/8 = 0,12 (\text{мг/мл})$$

Число степеней свободы $f = m = 8$

Ответ:

$$S^2(x_s) = 0,95 (\text{мг/мл})^2; f = 8; S^2(\bar{x}) = 0,12 (\text{мг/мл})^2$$

Контрольная задача 1.

Рассчитать оценки дисперсий по следующим экспериментальным данным.

s	1	2	3	4	5	6	7	8
x ₁ , кг/ч	12	115	122	117	119	125	121	124
x ₂ , мл/с	25	28	32	31	30	27	26	29
x ₃ , кг/м ³	12	125	130	115	120	110	140	130
x ₄ , %	68	72	73	69	70	72	71	69
s	9	10	11	12	13	14	15	16
x ₁ , кг/ч	11	118	-	-	-	-	-	-
x ₂ , мл/с	32	30	26	29	28	29	30	32
x ₃ , кг/м ³	13	125	120	115	135	125	-	-
x ₄ , %	72	74	73	69	68	-	-	-

Ответы:

$$S^2(x_1) = 10,4 (\text{кг/ч})^2, f_1 = 9; S^2(\bar{x}_1) = 1,04 (\text{кг/ч})^2,$$

$$S^2(x_2) = 4,9 (\text{мл/с})^2, f_2 = 15; S^2(\bar{x}_2) = 0,3 (\text{мл/с})^2,$$

$$S^2(x_3) = 74,7 (\text{кг/м}^3)^2, f_3 = 13; S^2(\bar{x}_3) = 5,3 (\text{кг/м}^3)^2,$$

$$S^2(x_4) = 4,3 (\%)^2, f_4 = 12; S^2(\bar{x}_4) = 0,33 (\%)^2.$$

Пример 3. Однородны ли дисперсии? Принять $P = 0,95$

	1	2	3	4	5	6
$S^2(x_i), \text{кг}^2$	230	140	180	212	80	150

f_i	24	12	10	8	16	30
-------	----	----	----	---	----	----

Решение:

Поскольку $f_i \neq \text{const}$, следует применить критерий Фишера:

$$1. \quad F = \frac{S^2(x_{si})_{max}}{S^2(x_{si})_{min}} = \frac{230}{80} = 2,9$$

2. По таблице приложения 3

$$F_T(0,95; 24; 16) = 2,3 < 2,9$$

Ответ:

Оценки дисперсии не однородны.

Пример 4. Можно ли представленные в таблице оценки дисперсии объединить в единую средневзвешенную оценку? Если можно, то рассчитать такую оценку. Принять $P=0,95$

i	1	2	3	4	5	Σ
$S^2(x_{si}), (\%)^2$	120	55	85	90	65	415
f_i	8	8	8	8	8	40

Решение:

Поскольку $f_i = \text{const}$, то для проверки однородности следует применить критерий Кохрена:

$$3 \quad G = \frac{S^2(x_{si})_{max}}{\sum_{i=1}^N S^2(x_{si})} = \frac{120}{415} = 0,2892.$$

По таблице приложения 2

$$4 \quad G_{sp}(0,95; 8; 5) = 0,4387$$

5 Средневзвешенная оценка:

$$S^2(x_s) = \frac{\sum_{i=1}^N S^2(x_{si}) f_i}{\sum_{i=1}^N f_i} = \frac{\sum_{i=1}^N S^2(x_{si})}{N} = \frac{415}{5} = 83,0(\%)^2$$

$$f = \sum_{i=1}^N f_i = N \cdot f = 8 \cdot 5 = 40.$$

Ответ:

$$S^2(x_s) = 83,0 (\%)^2, f = 40.$$

Контрольная задача 2

Рассчитать средневзвешенные оценки дисперсий по следующим экспериментальным оценкам дисперсий.

№	i	1	2	3	4	5	6	7	8
I	$S^2(x_{si}), (\%)$	86	30	45	54	68	70	34	-
	f_i	5	14	20	10	6	8	12	-
II	$S^2(x_{si}), (\text{мг/м})$	180	80	150	90	140	130	160	100
	f_i	6	6	6	6	6	6	6	6
III	$S^2(x_{si}), \text{ВТ}^2$	590	470	320	430	190	270	-	-
	f_i	4	21	16	12	30	25	-	-

Ответы:

I: $S^2(x_s) = 48,9 (\%)^2, f = 75;$

II: $S^2(x_s) = 128,8 (\%)^2, f = 48;$

III: Оценки неоднородны ($F=3,1; F_T=2,7$ при $P=0,95$), перейти к единой средневзвешенной оценке не представляется возможным.

Пример 5. С надежностью $P=0,95$ обеспечить однородность представленных в таблице данных, исключив грубые ошибки.

s	1	2	3	4	5	6	7	8	Σ
$x_s, \%$	54	53	54	30	46	52	55	54	398
$\Delta s, \%$	4,2	3,2	4,2	-19,8	-3,8	2,2	5,2	4,2	0,4

$\Delta_s^2, (\%)$	18	10	18	392	14	4,8	27	18	501
--------------------	----	----	----	-----	----	-----	----	----	-----

Решение:

$$6 \quad \bar{x} = \frac{1}{m} \sum_{s=1}^m x_s = 398/8 = 49,75 = 49,8 \%$$

2. Планирование и обработка результатов однофакторных экспериментов

2.1 Критерий оптимальности, факторы, ошибки опыта случайные, грубые, систематические

Влияние какого-либо фактора на выход процесса может быть выражено зависимостью $k=f(C)$. Если конкретному значению C_i соответствует единственное значение x_i , то такая зависимость называется **функциональной**. Эту зависимость получают путем строгих логических доказательств, не нуждающихся в опытной проверке. Например, площадь квадрата ω может быть представлена функциональной зависимостью от размера стороны квадрата $a: \omega = a^2$.

Если x_i остается неизвестным в то время как C_i изменяется, то x не зависит от C . Например, угол при вершине квадрата равный $\frac{\pi}{2}$, не зависит от размера стороны a_i .

Если для оценки величин x_i и C_i используются данные наблюдений, величины случайные, то функциональная зависимость между ними существовать не может.

Измерив отдельно сторону \bar{a} и площадь $\bar{\omega}$ квадрата, можно убедиться, что полученные результаты не могут быть представлены с абсолютной точностью $\bar{\omega} = \bar{a}^2$.

К формализации экспериментальных данных, т.е. построению по ним описывающей процесс зависимости, исследователь прибегает, когда не может составить **эвристическую (детерминированную) математическую модель** из-за недостаточного понимания механизма процесса или его чрезмерной сложности.

Получения в результате формализации экспериментальных данных **эмпирическая математическая модель** имеет меньшую ценность, чем отражающая механизм процесса эвристическая математическая модель, которая может предсказать поведение объекта за пределами изученного диапазона изменения переменных.

Приступая к эксперименту с целью получения эмпирической математической модели, исследователь должен определить необходимый объем опытных данных с учетом количества принятых к исследованию факторов, воспроизводимости процесса, предполагаемой структуры модели и обеспечения возможности проверки адекватности уравнения.

Если по результатам эксперимента, состоящего из двух опытов, получено линейное однофакторное уравнение $x = b_0 + b_1 C$, то построенная по этому уравнению прямая обязательно пройдет через эти экспериментальные точки. Следовательно, для того чтобы поверить, насколько хороша эта зависимость описывает данный процесс, надо поставить опыт еще хотя бы в одной точке. Этот дополнительный опыт дает возможность осуществить корректную процедуру проверки пригодности уравнения. Однако проверку обычно проводят не по одной дополнительной точке, которая не участвовала в определении коэффициентов уравнения, а по всем экспериментальным точкам, число которых (N) должно превышать число коэффициентов уравнения (N').

Так как $N > N'$, решение такой системы требует специального подхода, что видно из следующего простейшего примера.

В результате исследования влияния на процесс вакуум выпаривания температуры получены следующие данные (средние из m повторностей):

i	1	2	3	4	5
$t_i, ^\circ\text{C}$	25	28	30	32	35
$\bar{x}_i, \%$	22	24	27	30	33

Требуется по этим данным получить линейное уравнение

$$x = b_0 + b_1 t$$

Для нахождения двух неизвестных коэффициентов можно ограничиться результатами двух опытов.

Рассчитаем b_0 и b_1 по результатам первых двух опытов:

$$\begin{array}{l} \bar{x}_1 = b_0 + 25b_1 = 22 \\ \bar{x}_2 = b_0 + 28b_1 = 24 \end{array} \left| \begin{array}{l} \\ 0,7 \end{array} \right. \quad \begin{array}{l} 3b_1 = 2; \quad b_1 = 0,67 \approx \\ \\ b_0 = 24 - 0,7 \cdot 28 = -4,4 \\ x_{12} = -4,4 + 0,7t \end{array}$$

Рассчитаем b_0 и b_1 по результатам 2-го и 3-го опытов:

$$\begin{array}{l} \bar{x}_2 = b_0 + 28b_1 = 24 \\ \bar{x}_3 = b_0 + 30b_1 = 27 \end{array} \left| \begin{array}{l} \\ \\ \end{array} \right. \quad \begin{array}{l} 2b_1 = 3; \quad b_1 = 1,5 \\ \\ b_0 = 24 - 1,5 \cdot 28 = -18,0 \\ x_{23} = -18,0 + 1,5t \end{array}$$

Если взять крайние точки, то

$$\begin{array}{l} \bar{x}_1 = b_0 + 25b_1 = 22 \\ \bar{x}_5 = b_0 + 35b_1 = 33 \end{array} \left| \begin{array}{l} \\ \\ \end{array} \right. \quad \begin{array}{l} 10b_1 = 11; \quad b_1 = 1,1 \\ \\ b_0 = 22 - 1,1 \cdot 25 = -5,5 \\ x_{15} = -5,5 + 1,1t. \end{array}$$

Следовательно, оценки коэффициентов уравнения по результатам двух опытов зависят от выбора этих опытов.

Совершенно очевидно, что вычислительная процедура должна основываться на использовании всех экспериментальных данных. Алгоритм нахождения неизвестных коэффициентов уравнения в такой «переопределенной» задаче можно построить, если базироваться на признании вероятной природы экспериментальных данных.

Ошибка в предсказании по искомому уравнению результата i -го опыта характеризуется величиной невязки:

$$\Delta_i = \hat{x}_i - \bar{x}_i, \quad (2.1)$$

где \hat{x}_i - предсказанное значение выхода процесса в i -ом опыте; \bar{x}_i - полученное в i -ом опыте значение выхода процесса.

Плотность вероятности появления любой i -ой невязки, величины случайной, подчиняющейся нормальному закону распределения:

$$p(\Delta_i) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{\Delta_i^2}{2\sigma^2}\right) \quad (2.2)$$

Вся совокупность невязок может быть представлена в виде кривой плотности распределения, подобно кривой на рис. 1.1.

Если эта кривая соответствует кривой 3 на рис. 1.2, то можно говорить о большей точности описания процесса данным уравнением по сравнению с уравнением, по которому получена кривая 2 и особенно 1. Для 3-ей кривой произведение ординат $\prod_{i=1}^N p(\Delta_i)$ для конечного числа невязок будет больше соответствующего произведения ординат 2-ой и 1-ой кривой.

Произведение плотностей вероятности появления невязок, вычисленных для каждого из N опытов эксперимента, называется **функцией правдоподобия**:

$$\prod_{i=1}^N p(\Delta_i) = \prod_{i=1}^N \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{\Delta_i^2}{2\sigma^2}\right) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^N \frac{1}{\exp\left(\frac{\sum_{i=1}^N \Delta_i^2}{2\sigma^2}\right)}. \quad (2.3)$$

Чем больше величина функции правдоподобия, тем более точно уравнение описывает экспериментальные данные.

Совокупность оценок коэффициентов уравнения, которая максимизирует функцию правдоподобия, будет удовлетворять условиям максимального правдоподобия, будет наилучшей из всех других совокупностей при заданной структуре уравнения.

Функция правдоподобия достигает максимума при минимизации суммы квадратов $\sum_{i=1}^N \Delta_i^2$ невязок. Следовательно,

минимизация суммы квадратов невязок и будет условием получения максимально правдоподобных оценок коэффициентов аппроксимирующего уравнения при нормальном законе распределения вероятности результатов x_{si} .

Существует еще один метод определения коэффициентов уравнения. Процедура реализации этого условия при получении оценок коэффициентов уравнения названа **методом наименьших квадратов** (правильнее было бы назвать эту процедуру методом наименьшей суммы квадратов невязок).

Метод наименьших квадратов позволяет сгладить влияние случайных причин на экспериментальные оценки коэффициентов и достаточно просто получить по экспериментальным данным \bar{x}_i ($i = 1 \div N$) математическую модель процесса в виде полинома той или иной степени.

Предположим, что следует получить оценки коэффициентов линейного уравнения:

$$x = b_0 + b_1 C.$$

Минимизируемая функция:

$$\Phi = \sum_{i=1}^N (\hat{x}_i - \bar{x}_i)^2 = \sum_{i=1}^N [(b_0 + b_1 C_i) - \bar{x}_i]^2 = \sum_{i=1}^N (b_0^2 + 2b_0 b_1 C_i + b_1^2 C_i^2 - 2b_0 \bar{x}_i - 2b_1 C_i \bar{x}_i + \bar{x}_i^2) \quad (2.4)$$

Реализация $\frac{\partial \Phi}{\partial b_0} = 0$ и $\frac{\partial \Phi}{\partial b_1} = 0$, получим возможность рассчитать наилучшие оценки b_0 и b_1 , обеспечивающие минимизацию суммы квадратов невязок:

$$\begin{aligned} \frac{\partial \Phi}{\partial b_0} &= \sum_{i=1}^N (2b_0 + 2b_1 C_i - 2\bar{x}_i) = 0; \\ \frac{\partial \Phi}{\partial b_1} &= \sum_{i=1}^N (2b_0 C_i + 2b_1 C_i^2 - 2C_i \bar{x}_i) = 0. \end{aligned}$$

Система уравнений запишется следующим образом:

$$\sum_{i=1}^N \bar{x}_i = b_0 \sum_{i=1}^N 1 + b_1 \sum_{i=1}^N C_i \quad (2.5)$$

$$\sum_{i=1}^N C_i \bar{x}_i = b_0 \sum_{i=1}^N C_i + b_1 \sum_{i=1}^N C_i^2$$

В этом простейшем случае решение системы не вызывает больших трудностей:

$$b_0 = \frac{\sum_{i=1}^N \bar{x}_i \sum_{i=1}^N C_i^2 - \sum_{i=1}^N C_i \sum_{i=1}^N C_i \bar{x}_i}{\sum_{i=1}^N 1 \sum_{i=1}^N C_i^2 - (\sum_{i=1}^N C_i)^2},$$

$$b_1 = \frac{\sum_{i=1}^N 1 \sum_{i=1}^N C_i \bar{x}_i - \sum_{i=1}^N \bar{x}_i \sum_{i=1}^N C_i}{\sum_{i=1}^N 1 \sum_{i=1}^N C_i^2 - (\sum_{i=1}^N C_i)^2}$$

Используя данные рассмотренного выше примера, определим оценки коэффициентов линейной зависимости методом наименьших квадратов (число повторностей в каждом из 5-ти опытов одинаково).

Обозначив $C_i = t_i$, получим по (2.5):

$$Nb_0 + b_1 \sum_{i=1}^N t_i = \sum_{i=1}^N \bar{x}_i;$$

$$b_0 \sum_{i=1}^N t_i + b_1 \sum_{i=1}^N t_i^2 = \sum_{i=1}^N t_i \bar{x}_i.$$

Подсчитаем необходимые суммы и подставив их в систему уравнений; получим искомые оценки коэффициентов линейного уравнения:

$$\sum_{i=1}^N t_i = 150; \sum_{i=1}^N \bar{x}_i = 136; \sum_{i=1}^N t_i^2 = 4558; \sum_{i=1}^N t_i \bar{x}_i = 4147.$$

$$\begin{array}{l} 5b_0 + 150b_1 = 136 \\ 150b_0 + 4558b_1 = 4147 \end{array} \quad \left| \quad \begin{array}{l} 58b_1 = 67; \\ b_1 = 1,2 \end{array} \right.$$

$$b_0 = \frac{136 - 1,2 \cdot 150}{5} = -8,4$$

$$x = -8,4 + 1,2t.$$

Очевидно, это уравнение должно наилучшим образом представлять экспериментальные данные.

2.2 Проверка адекватности полученного уравнения и его использование для оптимизации процесса

Термин «адекватность» в данном случае предусматривает соответствие неточности предсказания экспериментальных данных с помощью полученного уравнения той неточности, с которой получены сами экспериментальные данные.

Адекватность уравнения экспериментальным данным можно проверить по критерию Фишера $F_T(P; f_a; f)$, сравнивая с ним отношение:

$$F = \frac{S_a^2}{S^2(\bar{x})} > 1,0, \quad (2.6)$$

где S_a^2 - дисперсия неадекватности (неточности) предсказания экспериментальных данных уравнением; $S^2(\bar{x})$ - средневзвешенная оценка дисперсии воспроизводимости среднего результата с числом степеней свободы $f=N(m-1)$.

S_a^2 - рассчитывается по формуле:

$$S_a^2 = \frac{\sum_{i=1}^N (x_i - \bar{x}_i)^2}{N - N'} \quad (2.7)$$

$$\text{где} \quad N - N' = f_a \quad (2.8)$$

есть число степеней свободы при определении дисперсии неадекватности (N – число экспериментальных точек, N' - число коэффициентов аппроксимирующей зависимости; $N > N'$).

Если $F < F_T$, то уравнение адекватно описывает исследуемый процесс в реализованном диапазоне изменения

аргумента, т.к. оценки неточности предсказания S_a^2 и неточности самих экспериментальных данных $S^2(\bar{x})$ однородны.

Если $S_a^2 > S^2(\bar{x})$ и $F > F_T(P; f_a; f)$, то уравнение неадекватно (недостаточно точно) описывает исследуемый процесс. В этом случае повышают степень аппроксимирующего полинома, увеличивая при необходимости число опытов эксперимента.

Поскольку в числителе F – отношения всегда записывается большая дисперсия, при $S^2(\bar{x}) > S_a^2$, F – отношение будет вычисляться не по формуле (2.6), а по формуле (2.8), и сравнение его будет осуществляться с критерием $F_T(P; f; f_a)$, где $f_1 = f$, а $f_2 = f_a$:

$$F = \frac{S^2(\bar{x})}{S_a^2} > 1,0 \quad (2.8)$$

Если в этом случае $F > F_T$, то уравнение неоправданно точно, неадекватно описывает экспериментальные данные, полученные с большой доверительной ошибкой. Для получения адекватного уравнения можно попробовать уменьшить степень уравнения. Если же в этом случае уравнение станет слишком грубым, то рекомендуется за математическую модель принять уравнение неадекватное, слишком точное. Однако в такой ситуации надо с большой осторожностью подходить к интерпретации полученного уравнения.

3. Двухуровневые планы многофакторных экспериментов

3.1 Методы наименьших квадратов при обработке результатов многофакторного эксперимента

Многофакторные уравнения получают по результатам опытов многофакторных экспериментов методом наименьших квадратов (наименьшей суммы квадратов невязок, см. 2.1).

Для получения линейного уравнения по результатам двухфакторного эксперимента требуется получить оценку трех коэффициентов

$$x = b_0 + b_1 C_1 + b_2 C_2. \quad (3.1)$$

Тогда минимальная функция:

$$\Phi = \sum_{i=1}^N [(b_0 + b_1 C_{1i} + b_2 C_{2i}) - \bar{x}_i]^2 = \sum_{i=1}^N (b_0^2 + 2b_0 b_1 C_{1i} + b_1^2 C_{1i}^2 + 2b_0 b_2 C_{2i} - 2b_0 \bar{x}_i + 2b_1 b_2 C_{1i} C_{2i} - 2b_1 C_{1i} \bar{x}_i + b_2^2 C_{2i}^2 - 2b_2 C_{2i} \bar{x}_i + \bar{x}_i^2).$$

Система уравнений максимального правдоподобия запишется следующим образом:

$$\begin{aligned} \frac{\partial \Phi}{\partial b_0} = 0 & \quad \sum_{i=1}^N (b_0 + b_1 C_{1i} + b_2 C_{2i}) = \sum_{i=1}^N \bar{x}_i; \\ \frac{\partial \Phi}{\partial b_1} = 0 & \quad \sum_{i=1}^N (b_0 C_{1i} + b_1 C_{1i}^2 + b_2 C_{1i} C_{2i}) = \sum_{i=1}^N C_{1i} \bar{x}_i; \\ \frac{\partial \Phi}{\partial b_2} = 0 & \quad \sum_{i=1}^N (b_0 C_{2i} + b_2 C_{1i} C_{2i} + b_2 C_{2i}^2) = \sum_{i=1}^N C_{2i} \bar{x}_i. \end{aligned}$$

Представим полученную систему в несколько ином виде (3.2):

$$\begin{aligned} \frac{\partial \Phi}{\partial b_0} = 0 & \quad N b_0 + b_1 \sum_{i=1}^N C_{1i} + b_2 \sum_{i=1}^N C_{2i} = \sum_{i=1}^N \bar{x}_i; \\ \frac{\partial \Phi}{\partial b_1} = 0 & \quad b_0 \sum_{i=1}^N C_{1i} + b_1 \sum_{i=1}^N C_{1i}^2 + b_2 \sum_{i=1}^N C_{1i} C_{2i} = \\ & \quad \sum_{i=1}^N C_{1i} \bar{x}_i; \\ \frac{\partial \Phi}{\partial b_2} = 0 & \quad b_0 \sum_{i=1}^N C_{2i} + b_2 \sum_{i=1}^N C_{2i}^2 + b_1 \sum_{i=1}^N C_{1i} C_{2i} = \\ & \quad \sum_{i=1}^N C_{2i} \bar{x}_i. \end{aligned}$$

Если при составлении плана эксперимента удалось бы обеспечить равенство нулю трех сумм:

$$\sum_{i=1}^N C_{1i} = 0, \quad \sum_{i=1}^N C_{2i} = 0, \quad \sum_{i=1}^N C_{1i} C_{2i} = 0, \quad (3.3)$$

то система уравнений распалась бы на три независимых уравнения:

$$Nb_0 = \sum_{i=1}^N \bar{x}_i; b_1 \sum_{i=1}^N C_{1i}^2 = \sum_{i=1}^N C_{1i} \bar{x}_i; b_2 \sum_{i=1}^N C_{2i}^2 = \sum_{i=1}^N C_{2i} \bar{x}_i.$$

с очень простым решением, причем полученные оценки коэффициентов b_0, b_1 и b_2 были бы независимы друг от друга (ортогональны друг другу). Однако выполнить эти условия в натуральной размерности факторов невозможно.

При рассмотрении однофакторного планирования такое же затруднение уже встречалось и было преодолено переходом к безразмерному выражению величины факторов по состоянию:

$$x_i = \frac{c_i - c_0}{\lambda}.$$

При составлении равномерных планов многофакторного и многоуровневого исследования безразмерную величину факторов получают по формуле:

$$x_{ij} = \frac{c_{ij} - c_{i0}}{\lambda_i}, \quad (3.4)$$

c_{ij} – величина i – го фактора в j – ом опыте;

$$c_{ij} = \frac{c_{imax} - c_{imin}}{2}, \quad (3.5)$$

c_{i0} – центр эксперимента; λ_i – интервал варьирования (изменения) фактора от 1-го уровня фактора до последующего $(l+1)$ -го уровня $(l=1 \div L)$:

$$\lambda_i = C_{i(l+1)} - C_{il} = const. \quad (3.6)$$

Перенос точки отсчета в центр эксперимента позволяет получить безразмерное выражение факторов в виде отрицательного числа, если $C_{ij} < C_{i0}$, равное нулю, если $C_{ij} = C_{i0}$,

и продолжительного числа, если $C_{ij} > C_{i0}$. Остается убедиться, что

$$\sum_{i=1}^N x_{ij} = 0 \text{ и } \sum_{i=1}^N x_{ij} x_{iz} = 0 \quad (i \neq z) \quad (3.7)$$

Планы, удовлетворяющие условию $\sum_{i=1}^N x_{ij} = 0$ называются симметричными, удовлетворяющие условию $\sum_{i=1}^N x_{ij} x_{iz} = 0$, при $(i \neq z)$ – линейно ортогональными, линейных коэффициентов т.е. обеспечивающими получение независимых оценок b_1 (см. 3.9).

Система уравнений (3.2) для планов симметричных и линейно-ортогональных распадется на независимые уравнения с очень простым определением оценок коэффициентов уравнения:

$$x = b_0 + b_1 x_1 + b_2 x_2 \quad (3.8)$$

$$\begin{aligned} N b_0 &= \sum_{i=1}^N \bar{x}_i & b_0 &= \frac{\sum_{i=1}^N \bar{x}_i}{N}, \\ b_1 \sum_{i=1}^N x_{1i}^2 &= \sum_{i=1}^N x_{1i} \bar{x}_i ; & b_1 &= \frac{\sum_{i=1}^N x_{1i} \bar{x}_i}{\sum_{i=1}^N x_{1i}^2}, \\ b_2 \sum_{i=1}^N x_{2i}^2 &= \sum_{i=1}^N x_{2i} \bar{x}_i & b_2 &= \frac{\sum_{i=1}^N x_{2i} \bar{x}_i}{\sum_{i=1}^N x_{2i}^2}. \end{aligned} \quad (3.9)$$

Возможность использовать в исследовании планы, обеспечивающие получение математической модели многофакторного процесса с независимыми оценками эффекта влияния фактора на величину критерия оптимальности, привлекают исследователя. Этим объясняется тот факт, что ортогональные планы применяются достаточно часто.

3.2 Двухуровневый план полного факторного эксперимента ПФЭ 2^n

Самым простым системным ($\sum_{i=1}^N x_{ii}=0$) и ортогональным ($\sum_{i=1}^N x_{ii} x_{ji} = 0$) планом является двухуровневый план полного факторного эксперимента ПФЭ 2^n .

План предусматривает реализацию всех возможных опытов, условия проведения которых соответствуют любому сочетанию величин исследуемых факторов при их изменении лишь на двух уровнях: верхнем C_i^+ и нижнем C_i^- (см. рис. 3.1).

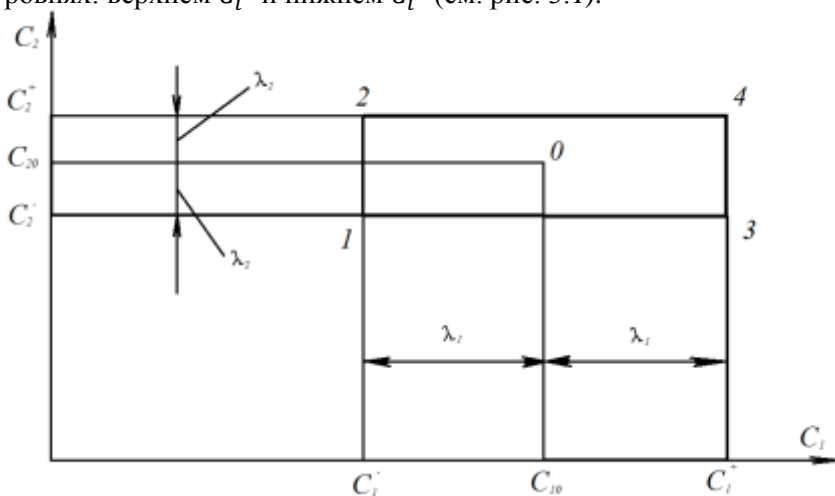


Рис. 3.1 План ПФЭ 2^2 в натуральной размерности факторов

Центр эксперимента:

$$C_{i0} = \frac{C_i^+ + C_i^-}{0} \quad (3.10)$$

соответствует нулевому ($x_i = 0$) уровню величины факторов, от него рассчитывают интервалы варьирования λ_i :

$$\lambda_i = C_i^+ - C_{i0} = C_{i0} - C_i^-. \quad (3.11)$$

Приняв нумерацию опытов плана с рисунка, можно план записать в виде таблицы:

Таблица 3.1
План ПФЭ² в натуральной размерности факторов

i	C _{1i}	C _{2i}	\bar{x}_i
1		C_2^-	\bar{x}_1
2	C_1^-	C_2^+	\bar{x}_2
3	C_1^+	C_2^-	\bar{x}_3
4	C_1^+	C_2^+	\bar{x}_4

Число опытов планов ПФЭⁿ соответствует числу сочетаний из n элементов при их изменении на двух уровнях:

$$N=2^n. \quad (3.12)$$

Для двухфакторного эксперимента $N = 2^2=4$, трехфакторного - $N = 2^3=8$ и т.д.

В соответствии с 3.4 и 3.11 найдем числовые значения верхнего x_i^+ и нижнего x_i^- уровней факторов в безразмерном выражении:

$$\begin{aligned} x_i^+ &= \frac{C_i^+ - C_{i0}}{\lambda_i} = \frac{C_i^+ - C_{i0}}{C_i^+ - C_{i0}} = +1; \\ x_i^- &= \frac{C_i^- - C_{i0}}{\lambda_i} = \frac{C_i^- - C_{i0}}{C_{i0} - C_i^-} = -1. \end{aligned} \quad (3.13)$$

Следовательно, любой фактор на нижнем уровне в безразмерном выражении характеризуется числом минус единица (-1 или -), на верхнем - плюс единица (+1 или +).

Сначала записывают планы ПФЭⁿ в безразмерном выражении величины факторов, а потом по ним составляют рабочий план в натуральной размерности факторов. Перевод величины фактора в натуральную размерность осуществляют с учетом 3.4 по формуле:

$$C_{ii} = C_{i0} + \lambda_i x_{ij} \quad (3.14)$$

В приведенной ниже таблице и на рис. 3.2 изображен план ПФЭ² в безразмерном выражении величины факторов. Целесообразно в дополнительных строках записать параметры плана C_{i0} и λ_i , чтобы иметь возможность пользоваться формулой 3.14.

Таблица 3.2
План ПФЭ² в безразмерном выражении факторов

i	x_{1j}	x_{2j}	\bar{x}_i
1	-	-	\bar{x}_1
2	-	+	\bar{x}_2
3	+	-	\bar{x}_3
4	+	+	\bar{x}_4
C_{i0}	C_{10}	C_{20}	\bar{x}_0
λ_i	λ_1	λ_2	

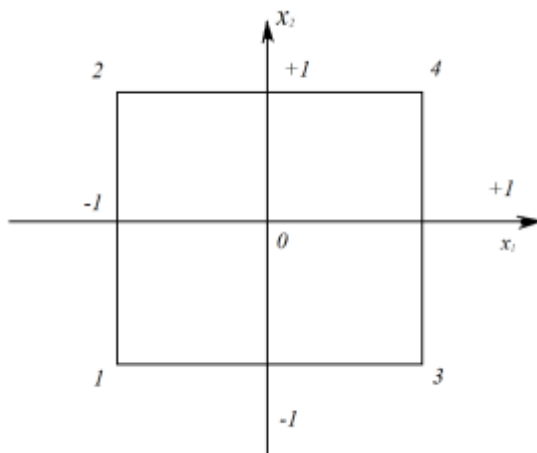


Рис. 3.2 План ПФЭ² в безразмерном выражении факторов

Центр эксперимента назначается после изучения литературы и должен соответствовать самым благоприятным

условиям протекания исследуемого процесса. Иногда, конечно, приходится назначать центр эксперимента по одному или нескольким факторам интуитивно или пользуясь советом более опытного специалиста. Следует сказать, что необоснованное назначение центра эксперимента может лишь увеличить трудоемкость исследования, но не снизить эффективность полученных конечных результатов.

Величины интервалов варьирования назначают опять-таки по результатам изучения литературных данных с учетом разрешающей способности применяемых методик и приборов. Следует стремиться к тому, чтобы интервал варьирования был как можно меньше, на вызывал такое изменение критерия оптимальности, которое можно достоверно зафиксировать при проведении эксперимента. Ошибка в назначении интервала варьирования может негативно отразиться на конечном результате исследования, т.к. слишком малый интервал варьирования может создать уверенность, что данный фактор значимо не влияет на процесс, а слишком большой интервал варьирования активного фактора вызовет такое большое изменение величины критерия оптимальности, что будет маскировать влияние других факторов, некоторые из которых в связи с этим могут быть отнесены к неактивным (незначимым).

Алгоритм построения планов при $n > 2$ сводится к следующему:

1. составляем таблицу, в которой число свободных строк $N=2^n$, число свободных столбцов $2n$ (при $n=3$ число строк 8, число столбцов 6 – см. табл. 3.3);
2. в левом верхнем углу таблицы записываем план ПФЭ 2^{n-1} в безразмерном выражении факторов (при $n=3$ записываем план ПФЭ 2^2);
3. повторяем план ПФЭ 2^{n-1} , заполняя оставшиеся строки в $n-1$ левых столбцах таблицы;
4. планируем величину последнего фактора в опытах верхней половины таблицы на нижнем уровне, нижней половины на верхнем уровне;

5. в оставшихся свободных столбцах записываем условия проведения каждого опыта C_{ii} в натуральной размерности факторов

$$C_{ij} = C_{i0} + x_{ij}\lambda_i.$$

Величины C_{i0} и λ_i удобно записывать в последних двух дополнительных строках, \bar{x}_i – в дополнительном столбце таблицы.

Для построения плана ПФЭ⁴ число строк будет $2^4+2=18$, число столбцов $4 \cdot 2 + 1 = 9$, план ПФЭ^{2ⁿ⁻¹} будет соответствовать ПФЭ³ (см. пример 1).

После построения плана в безразмерном выражении факторов x_{ii} следует проверить по всем столбцам $\sum_{i=1}^N x_{ij} = 0$ и $\sum_{i=1}^N x_{ij} x_{iz} = 0$, т.е. симметричность плана и его ортогональность (см. примеры 2,3,4,6).

Таблица 3.3
План ПФЭ³

	x_{1i}	x_{2i}	x_{3i}	C_{1i}	C_{2i}	C_{3i}	\bar{x}_i
1	-	-	-	C_1^-	C_2^-	C_3^-	\bar{x}_1
2	-	+	-	C_1^-	C_2^+	C_3^-	\bar{x}_2
3	+	-	-	C_1^+	C_2^-	C_3^-	\bar{x}_3
4	+	+	-	C_1^+	C_2^+	C_3^-	\bar{x}_4
5	-	-	+	C_1^-	C_2^-	C_3^+	\bar{x}_5
6	-	+	+	C_1^-	C_2^+	C_3^+	\bar{x}_6
7	+	-	+	C_1^+	C_2^-	C_3^+	\bar{x}_7
8	+	+	+	C_1^+	C_2^+	C_3^+	\bar{x}_8
C_{i0}	C_{10}	C_{20}	C_{30}				
λ_i	λ_1	λ_2	λ_3				

Для планов ПФЭⁿ

$$\sum_{i=1}^N x_{ij}^2 = N \tag{3.15}$$

где N – число опытов, такие планы называются нормированными.

Тогда формула (101) для определения коэффициентов линейного уравнения запишется следующим образом:

$$b_0 = \frac{\sum_{i=1}^N \bar{x}_i}{N}; \quad b_i = \frac{\sum_{i=1}^N x_{ij} \bar{x}_i}{N}. \quad (3.16)$$

Двухуровневые планы полного факторного эксперимента помимо линейных коэффициентов позволяют определить независимым образом оценки эффектов межфакторных взаимодействий (см. 3.3).

3.3 Уравнения, полученные по результатам реализации планов ПФЭ 2^n

По результатам двухфакторного эксперимента можно составить уравнение, в котором помимо линейных членов (3.8) будет член, учитывающий эффект парного межфакторного взаимодействия:

$$x = b_0 + b_1 x_1 + b_2 x_2 + b_{12} x_1 x_2. \quad (3.17)$$

Такая возможность появилась в связи с тем, что эксперимент по плану ПФЭ 2^2 насчитывает $N=2^2=4$ опыта, а число коэффициентов N' в линейном уравнении всего три. Рассматривая МНК примечательно к плану ПФЭ 2^2 , надо было невязки формулировать с учетом эффекта межфакторного взаимодействия, получить не три, а четыре уравнения в системе (3.2) и конечном счете – формулу для расчета независимой оценки эффекта парного межфакторного взаимодействия b_{12} :

$$b_{12} = \frac{\sum_{i=1}^N x_{1i}x_{2i}\bar{x}_i}{N} \quad (3.18)$$

План ПФЭ³ дает возможность рассчитать $8(N^3=8)$ коэффициентов:

$$x = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_{12}x_1x_2 + b_{13}x_1x_3 + b_{23}x_2x_3 + b_{123}x_1x_2x_3 \quad (3.19)$$

Коэффициент b_{123} будет рассчитываться независимо от других коэффициентов (так как столбец $x_{1i}x_{2i}x_{3i}$ ортогонален любому другому столбцу) по формуле:

$$b_{123} = \frac{\sum_{i=1}^N x_{1i}x_{2i}x_{3i}\bar{x}_i}{N} \quad (3.20)$$

3.4 Статистический анализ значимости оценок коэффициентов управления, его адекватности и работоспособности

Статистический анализ значимости оценок коэффициентов управления имеет своей целью показать с заранее заданной вероятностью P , что полученные оценки коэффициентов управления по модулю либо больше (тогда они значимо отличаются от нуля), либо меньше ошибки в их определении (тогда они незначимо отличаются от нуля и должны быть из уравнения исключены).

Определение оценки дисперсии единичного результата в каждом опыте при $m_i = const = m$ осуществляется по формуле:

$$S^2(x_{ij}) = \frac{\sum_{j=1}^{m_i} (x_{ij} - \bar{x}_i)^2}{m-1} \quad (3.21)$$

Средневзвешенная оценка дисперсии единичного результата в каждом опыте при $m = const$ рассчитывается по однородным оценкам $S^2(x_{ik})$:

$$S^2(x_{ij}) = \frac{\sum_{i=1}^N \sum_{j=1}^m (x_{ij} - \bar{x}_i)^2}{N(m-1)} \quad (3.22)$$

Величина \bar{x}_i определяется по m повторностям, следовательно, она будет ближе к истинному значению выхода, чем результаты единичных повторностей в соответствующем опыте.

В соответствии с теоремой об оценке дисперсии среднего результата средневзвешенная оценка дисперсии среднего результата любого опыта эксперимента будет в m раз меньше средневзвешенной оценки дисперсии единичного результата $S^2(x_j)$:

$$S^2(\bar{x}) = \frac{S^2(x_j)}{m} = \frac{\sum_{i=1}^N \sum_{j=1}^m (x_{ij} - \bar{x}_i)^2}{N(m-1)m} \quad (3.23)$$

Число степеней свободы средневзвешенных оценок дисперсии равна сумме чисел степеней свободы исходных оценок дисперсии, т.е.

$$f = \sum_{i=1}^N f_i = N(m-1). \quad (3.24)$$

В определении оценки любого из коэффициентов уравнения участвуют все N средних результатов опытов, оценкой дисперсии которых будет одна и та же величина $S^2(\bar{x})$. Так как структура формул для определения оценок коэффициентов будет соответствовать структуре формулы для расчета среднеарифметического, в соответствии с теоремой о дисперсии среднего можно записать:

$$S^2(b_i) = \frac{S^2(\bar{x})}{N} = \frac{\sum_{i=1}^N \sum_{j=1}^m (x_{ij} - \bar{x}_i)^2}{N(m-1)mN}. \quad (3.25)$$

Следовательно, оценка дисперсии любого из независимо определяемых коэффициентов будет иметь одну и ту же величину.

Планы, обеспечивающие получение одинаковой и минимальной оценки дисперсии коэффициентов полученного по ним уравнения, называются –оптимальными (буквой D может обозначаться дисперсии, отсюда и название). Это свойство ценится специалистами достаточно высоко, в связи с чем при прочих равных свойствах планов предпочтение отдается – оптимальным.

Доверительная ошибка коэффициентов рассчитывается по критерию Стьюдента

$$\varepsilon(b_i) = t(P; f)S(b_i). \quad (3.26)$$

Если $|b_i| > \varepsilon(b_i)$, то оценка коэффициента b_i значимо отличается от нуля. В противном случае оценка b_i считается значимо от нуля не отличающейся, и ее приравнивают нулю, т.е. исключают из уравнения (см. пример 4а).

Правомерность этих действий подтверждается тем, что оценки всех коэффициентов независимы друг от друга (ортогональны) (см. пример 3).

При получении незначимого линейного коэффициента какого-либо фактора следует найти этому объяснение, проанализировав следующие ситуации:

1. данный фактор на исследуемый процесс не влияет;
2. выбран слишком малый интервал варьирования, в связи с чем изменение выхода процесса, обусловленное изменением фактора, соразмерно случайным отклонениям, вызываемым влиянием неучтенных факторов и погрешностью измерительных приборов;
3. значение данного фактора в центре в центре эксперимента соответствует вершине отклонения отклика, в связи с чем одинаковое его уменьшение или увеличение понизит выход процесса приблизительно на одну и ту же величину.

Подтвердить наличие той или иной ситуации поможет следующее мероприятие:

1. в плане ПФЭ² выделить результаты двух опытов, условия которых отличаются лишь уровнем i -го фактора, давшего незначимую оценку линейного эффекта (остальные факторы находятся на одинаковых в обоих опытах уровнях);

2. осуществить два дополнительных опыта при $C_{ij} = C_{i0} \pm \lambda_i$ и $C_i = C_{i0} + (2 \div) \lambda_i$ при прежнем значении всех других факторов.

Если в уравнении после проверки значимость коэффициентов останутся все N коэффициентов, то проверка адекватности уравнения теряет смысл. Рассчитанный по такому уравнению выход процесса для условий какого-либо i -го опыта должен с учетом округления при расчете оценок коэффициентов уравнения быть равным \bar{x}_i , а само уравнение характеризуется как неадекватное, слишком точное.

Если число значимых коэффициентов хотя бы на единицу меньше числа опытов, то появляется необходимость (и возможность) статистической проверки адекватности уравнения экспериментальным данным. Эта проверка осуществляется по критерию Фишера:

1. рассчитывают \hat{x}_i для каждого опыта по уравнению, из которого исключены незначимые члены;

2. находят величины невязок $\Delta_i = \hat{x}_i - \bar{x}_i$, причем $\sum_{i=1}^N \Delta_i \approx 0$, что является проверкой отсутствия арифметических ошибок в предыдущих расчетах;

3. рассчитывают дисперсию неадекватности:

$$S_a^2 = \frac{\sum_{i=1}^N |\hat{x}_i - \bar{x}_i|^2}{N - N'} \quad (3.27)$$

с числом степеней свободы $f_a = N - N'$; N' - число значимых коэффициентов в уравнении регрессии;

4. рассчитывают F -отношение по одной из формул:

$$F_1 = \frac{S_a^2}{S^2(\bar{x})} > 1,0; F_2 = \frac{S^2(\bar{x})}{S_a^2} > 1,0.$$

5. сравнивают полученное значение F -отношение со значением $F_m(P; f_1; f_2)$ критерия Фишера (см. приложение 3).

Критерий Фишера всегда больше единицы. Поэтому в числитель F -отношение ставится большая оценка дисперсии. Если дисперсия воспроизводимости $S^2(\bar{x})$ больше S_a^2 , то в числителе должна стоять $S^2(\bar{x})$. Тогда $f_1 = N(m - 1) = f$ и $f_2 = N - N' = f_a$. Если $S_a^2 > S^2(\bar{x})$, то $f_1 = N - N' = f_a$ и $f_2 = N(m - 1) = f$.

Если $F_1 < F_m(f_a; f)$ и $F_2 < F_1(f; f_a)$, то уравнение адекватно экспериментальным данным и может служить хорошей математической моделью (если еще будет доказана работоспособность уравнения в центре эксперимента – см. далее).

Если $F_1 > F_m(f_a; f)$, то следует сделать вывод о том, что уравнение неадекватно описывает экспериментальные данные, точность описания процесса данным уравнением значимо ниже той точности, с которой получены экспериментальные результаты. Однако, такой ситуации при анализе уравнения, полученного по результатам плана ПФЭ²ⁿ, не должно быть. Если она реально создалась, то это свидетельствует об ошибке или ошибках: при определении оценок коэффициентов уравнения, при исключении незначимых оценок коэффициентов, при расчете предсказанных по уравнению результатных опытов \hat{x}_i , при расчете оценки дисперсии адекватности S_a^2 и F -отношения. Это утверждение основывается на следующих соображениях: если до исключения незначимых оценок некоторых коэффициентов $\hat{x}_i = \bar{x}_i$, (т.к. $N' = N$), то исключение оценок коэффициентов, действительно значимо от нуля не отличающихся, не может привести уравнение в разряд слишком грубых. Поэтому, следует обязательно найти ошибку в расчетах и ее исправить.

Если $F_2 > F_m(f_a; f)$, то уравнение и в этом случае также будет неадекватным, но неадекватность его будет означать неоправданно точное описание экспериментальных данных этим уравнением. Уравнение в этом случае может служить основой для отыскания оптимальных условий, но использовать его при интерпретации уравнения надо очень осторожно (см. пример 4б).

Однако даже если уравнение оказывается адекватным, то это еще не значит, что в приятном диапазоне изменения факторов оно будет достаточно точно описывать поверхность отклика при $|x_i|$, отличающихся от единицы.

Свободный член уравнения является оценкой выхода процесса в центральной точке эксперимента. Расчетным столбцом для определения b_0 по формуле

$$b_0 = \frac{\sum_{i=1}^N \bar{x}_i}{N} = \frac{\sum_{i=1}^N x_{0i} \bar{x}_i}{N} \quad (3.28)$$

будет столбец фиктивной переменной $x_{0i} = 1$. Но такой же столбец $x_{ij}^2 = (\pm 1)^2 = 1$ надо было бы использовать и при расчете оценки квадратичного эффекта b_i .

$$b_i = \frac{\sum_{i=1}^N x_{ij}^2 \bar{x}_i}{\sum_{i=1}^N (x_{ij}^2)^2} = \frac{\sum_{i=1}^N x_{ij}^2 \bar{x}_i}{N} = \frac{\sum_{i=1}^N \bar{x}_i}{N}. \quad (3.29)$$

Следовательно, по результатам реализации плана ПФЭ 2^n оценка свободного члена является смешанной с суммарной оценкой квадратичных эффектов всех факторов, т.к. $\sum_{i=1}^N x_{0i} x_{ij}^2 \neq 0$:

$$b_0 \rightarrow \beta_0 + \sum_{i=1}^n \beta_i. \quad (3.30)$$

Если квадратичные эффекты будут значимы, то и прогнозируемые результаты \hat{x}_0 опытов при $x_{i0} = 0$ будут

значимо отличаться от их экспериментальных значений, т.к. в уравнении квадратичные эффекты не представлены.

Если дополнительно к плану поставить в нескольких ($m_0 > m$) повторностях опыт в центре эксперимента, то, не приступая даже к расчету всех (кроме b_0) оценок коэффициентов уравнения, можно судить о возможности описания процесса уравнением без включения в последнее квадратичных членов.

Оценкой суммарного квадратичного эффекта факторов будет являться разность:

$$|\bar{x}_0 - b_0| = \sum_{i=1}^n b_i. \quad (3.31)$$

Если эта разность при заданной вероятности значима, то переходят к планам второго порядка, позволяющим получить квадратное уравнение процесса.

Значимость разности $|\hat{x}_0 - b_0|$ устанавливается при помощи доверительной ошибки этой разности:

$$\varepsilon(\bar{x}_0 - b_0) = t(P; f)S(\bar{x}_0 - b_0) \quad (3.32)$$

и сравнения ее с величиной самой разности.

Процедура этой проверки основывается на дисперсионном анализе, осуществляемом в следующем порядке.

1. Рассчитывают среднеарифметическую оценку результата дополнительного опыта в центре эксперимента:

$$\bar{x}_0 = \frac{1}{m_0} \sum_{j=1}^{m_i} x_{j0}$$

и оценку дисперсии с числом степеней свободы $f_0 = m_0 - 1$

$$S^2 x_{0j} = \frac{\sum_{j=1}^{m_0} (x_{j0} - \bar{x}_0)^2}{m_0 - 1} \quad (3.33)$$

2. Оценку дисперсии среднего результата, полученного в центре плана, находим в соответствии с теоремой о дисперсии среднего по формуле:

$$S^2(\bar{x}_0) = \frac{S^2(x_{0j})}{m_0} \quad (3.34)$$

3. Оценка дисперсии коэффициента b_0 в соответствии с (3.25) равна:

$$S^2(b_0) = \frac{S^2(\bar{x})}{N} = \frac{S^2(x_j)}{Nm}$$

4. В соответствии с законом накопления ошибок (42) можно записать:

$$S^2(\bar{x}_0 - b_0) = S^2(\bar{x}_0) + S^2(b_0) = \frac{S^2(x_{0j})}{m_0} + \frac{S^2(x_j)}{Nm}$$

5. По однородным оценкам $S^2(x_{0j})$ и $S^2(x_j)$ рассчитывают средневзвешенную оценку дисперсии единичного результата $S^2(x_j)'$:

$$S^2(x_j)' = \frac{S^2(x_{0j})(m_0 - 1) + S^2(x_j)N(m - 1)}{(m_0 - 1) + N(m - 1)} \quad (3.35)$$

с числом степеней свободы $f = (m_0 - 1) + N(m - 1)$.

6. Окончательно:

$$S^2(\bar{x}_0 - b_0) = S^2(x_j)' \cdot \left(\frac{1}{m_0} + \frac{1}{Nm} \right) = S^2(x_j)' \cdot \frac{Nm + m_0}{Nmm_0} \quad (3.36)$$

7. Доверительная ошибка разности $|\bar{x}_0 - b_0|$ равна:

$$\varepsilon(\bar{x}_0 - b_0) = t(P; f)S(\bar{x}_0 - b_0), \quad (3.37)$$

где $t(P; f)$ – критерий Стьюдента (табл. приложение 7).

8. Если $\varepsilon(\bar{x}_0 - b_0) > |\bar{x}_0 - b_0|$, то с заданной вероятностью P можно считать эту разность значимой. Это равносильно доказательству необходимости включения в уравнение оценок квадратичных эффектов факторов и признанию невозможности по результатам реализации плана ПФЭ² получить работоспособную математическую модель процесса.

Если $|\bar{x}_0 - b_0| < t(P; f) S(\bar{x}_0 - b_0)$, то квадратичные эффекты в уравнении можно не представлять. Лишь в этом случае уравнение, адекватность которого ранее была доказана с помощью критерия Фишера, можно считать математической моделью исследуемого процесса (см. пример 4б).

Для прогнозирования результатов процесса при заданных величинах факторов в натуральной размерности C_i целесообразно математическую модель представить в виде $x = f(C_i)$, заменив $x_i \frac{C_i - C_{i0}}{\lambda_i}$, где C_{i0} и λ_i – параметры плана эксперимента.

3.5 Расчет программы оптимизации по линейному уравнению

Если предстоящий эксперимент имеет своей целью определение оптимальных условий функционирования исследуемого процесса, то следует себя настроить на неоднократное повторение выбранного на начальном этапе исследования простого плана с постепенным (или резким, если повезет) приближением к около оптимальной области. Когда возможности выбранного плана эксперимента обеспечивать дальнейшее усложнение уравнения, а, следовательно, и получение более точного описания процесса исчерпаны,

исследователь должен перейти к более сложному плану и получить модели в виде более сложного уравнения.

План ПФЭ 2^n позволяет получить линейное:

$$x = b_0 + \sum_{i=1}^n b_i x_i$$

или неполное квадратное уравнение:

$$x = b_0 + \sum_{i=1}^n b_i x_i + \sum_{i,j=1}^n b_{ij} x_i x_j + \sum_{i,j,k=1}^n b_{ijk} x_i x_j x_k + \dots \quad (3.38)$$

По мере приближения к около оптимальной области уравнение будет содержать все большее число значимых коэффициентов, но наступит момент, когда будет доказана неработоспособность уравнения по опыту в центре эксперимента и, как следствие, значимость квадратичных эффектов факторов. Вот тогда приходится переходить к планам второго порядка, позволяющим получить полное квадратное уравнение и большие возможности сформулировать оптимальные условия процесса. Именно планы второго порядка и квадратные уравнения, описывающие процесс в около оптимальной области, являются, как правило, конечным результатом предпринятого исследования.

На первом этапе исследования более вероятно, чем на следующих, получить описание участка поверхности отклика, находящегося вдали от около оптимальной области (мы еще только хотим к ней приблизиться), в виде простого линейного уравнения. Предположим, нам это удалось, и некое линейное уравнение, как адекватное и работоспособное, принято за математическую модель процесса в реализованном диапазоне изменения факторов.

Линейная модель описывает наклонную гиперплоскость в $(n+1)$ -мерном факторном пространстве. Если не принимать во внимание ограничения, то можно, анализируя линейную модель, прийти к абсурдному выводу о возможности увеличивать эффективность данного процесса беспредельно. Однако линейное уравнение будет отражать с достаточной точностью

поверхность отклика лишь в некоторой локальной области, соответствующий, строго говоря, лишь изученному диапазону изменения факторов. Можно заранее сказать, что характер поверхности отклика на некотором удалении от изученного участка будет все больше отходить от плоскости, описанной линейным уравнением.

Если найти направление самого круглого подъема (градиента) этой плоскости и двигаться в этом направлении, осуществляя опыты при соответствующих условиях, за пределы изученной области, то можно найти такое сочетание значений факторов, которое будет соответствовать началу снижения поверхности отклика (процесс идет по типу ингибирования) или постепенной стабилизации получаемых результатов (процесс идет по типу насыщения). Определение такого сочетания факторов, оптимального на данном этапе исследования, можно осуществить лишь экспериментальным путем, запланировав в направлении градиента, найденного по полученному уравнению, серию опытов.

Градиентом ($\text{grad } x$) скалярной функции $x = f(x_1, x_2, \dots)$ называют вектор, перпендикулярный линии (поверхности, гиперповерхности) равного вывода, указывающий направление наиболее быстрого изменения (роста или падения) скалярной функции:

$$\text{grad } x = \frac{\partial x}{\partial x_1} \vec{i} + \frac{\partial x}{\partial x_2} \vec{j} + \frac{\partial x}{\partial x_3} \vec{k} + \dots, \quad (3.39)$$

где $\frac{\partial x}{\partial x_1}$ - угловой коэффициент b_1 наклона поверхности отклика в рассматриваемой точке к соответствующей оси; \vec{i} , \vec{j} , \vec{k} - единичные векторы (орты) в положительном направлении координатных осей (их модуль равен единице).

Изменение факторов при расчете серии опытов программы оптимизации будет пропорционально составляющим вектора градиента, т.е.

$$\Delta_1:\Delta_2:\Delta_3 \dots = \frac{\partial x}{\partial x_1} \vec{i}:\frac{\partial x}{\partial x_2} \vec{j}:\frac{\partial x}{\partial x_3} \vec{k}:\dots, \quad (3.40)$$

где Δ_1 изменение i -й координаты последующего опыта по сравнению с координатой предыдущего.

Примеры и контрольные задачи

Пример 1. Построить план ПФЭ⁴ в безразмерном выражении и в натуральной размерности факторов по следующим параметрам плана:

$$C_{10} = 38^\circ C, C_{20} = 24\%, C_{30} = 0,6 \text{ м}, C_{40} = 500 \text{ об/мин.}$$

$$\lambda_1 = 6^\circ C, \lambda_2 = 4\%, \lambda_3 = 0,15 \text{ м}, \lambda_4 = 50 \text{ об/мин.}$$

Решение:

Число строк $1+16+2=19$, число столбцов $1+4+4=9$

i	X _{ij}				C _{ii}				\bar{x}_i ,
	X _{1i}	X _{2i}	X _{3i}	X _{4i}	C _{1i} , ⁰	C _{2i} ,	C _{3i} ,	C _{4i}	
1	-	-	-	-	32	20	0.45	450	46
2	-	+	-	-	32	28	0.45	450	62
3	+	-	-	-	44	20	0.45	450	66
4	+	+	-	-	44	28	0.45	450	64
5	-	-	+	-	32	20	0.75	450	54
6	-	+	+	-	32	28	0.75	450	70
7	+	-	+	-	44	20	0.75	450	74
8	+	+	+	-	44	28	0.75	450	90
9	-	-	-	+	32	20	0.45	550	70
10	-	+	-	+	32	28	0.45	550	86
11	+	-	-	+	44	20	0.45	550	90
12	+	+	-	+	44	28	0.45	550	106
13	-	-	+	+	32	20	0.75	550	78
14	-	+	+	+	32	28	0.75	550	94
15	+	-	+	+	44	20	0.75	550	98
16	+	+	+	+	44	28	0.75	550	114

C_{10}					38	24	0.6	500	80
λ_i					6	4	0.15	50	

$$C_1^+ = C_{10} + \lambda_1 = 38 + 6 = 48^\circ C.$$

$$C_1^- = C_{10} - \lambda_1 = 32^\circ C.$$

$$C_2^+ = C_{20} + \lambda_2 = 24 + 4 = 28\%.$$

$$C_2^- = C_{20} - \lambda_2 = 20\%.$$

$$C_3^+ = C_{30} + \lambda_3 = 0,6 + 0,15 = 0,75 \text{ м.}$$

$$C_3^- = C_{30} - \lambda_3 = 0,45 \text{ м.}$$

$$C_4^+ = C_{40} + \lambda_4 = 500 + 50 = 550 \text{ об/мин.}$$

$$C_4^- = C_{40} - \lambda_4 = 450 \text{ об/мин.}$$

Примечание. Эксперимент осуществлялся блоками по числу поверхностей каждого опыта $t=4$ (см. пример 8).

\bar{x}_i -средние результаты из 4-х повторностей опыта (см. пример 8).

Пример 2. По полученным в предыдущем примере результатам \bar{x}_i рассчитать коэффициенты линейного уравнения в безразмерном выражении (I) и в натуральной размерности (II) факторов.

Решение:

$$\text{I. } x = b_0 + \sum_{i=1}^N b_i x_i; \quad b_0 = \frac{\sum_{i=1}^N \bar{x}_i}{N}; \quad b_i = \frac{\sum_{i=1}^N x_{ij} \bar{x}_i}{N};$$

$$b_0 = \frac{1262}{16} = 78,87 = 78,9; \quad b_1 = \frac{702-560}{16} = \frac{142}{16} = 8,87 =$$

8,9;

$$b_2 = \frac{686-576}{16} = \frac{110}{16} = 6,87 = 6,9; \quad b_3 = \frac{672-590}{16} = 5,12 =$$

5,1;

$$b_4 = \frac{736-526}{16} = \frac{210}{16} = 13,31 = 13,3;$$

$$x = 78,9 + 8,9x_1 + 6,9x_2 + 5,1x_3 + 13,3x_4.$$

$$\text{II. } x_i = \frac{C_i - C_{i0}}{\lambda_i};$$

$$x_1 = \frac{C_1 - 38}{6} = 0,17C_1 - 6,3; \quad x_2 = \frac{C_2 - 24}{4} = 0,25C_2 - 6,0;$$

$$x_3 = \frac{C_3 - 0,6}{0,15} = 6,7C_3 - 4,0; \quad x_4 = \frac{C_4 - 500}{50} = 0,02C_4 - 10,0.$$

Подставим:

$$x = 78,9 + 8,9x_1 + 6,9x_2 + 5,1x_3 + 13,3x = 78,9 + 8,9(0,17C_1 - 6,3) + 6,9(0,25C_2 - 6,0) + 5,1(6,7C_3 - 4,0) + 13,3(0,02C_4 - 10,0) = 178,6 + 1,5C_1 + 1,72C_2 + 34,17C_3 + 0,27C_4;$$

Проверим, сравнив, например, предсказанные уравнениями и $\hat{x}_{0II} = \psi(c_{i0})$ результаты в центре экстремума:

(I): $\hat{x}_0 = 78,9 \text{ кг/м}^3 \cdot \text{ч}$ при $x_{i0} = 0$;

(II): $x = -178,6 + 1,51C_{10} + 1,72C_{20} + 34,01C_{30} + 0,27C_{40} = -178,6 + 1,51 \cdot 38 + 1,72 \cdot 24 + 34,01 \cdot 0,6 + 0,27 \cdot 500 = 78,5 \text{ кг/м}^3 \cdot \text{ч}$ – совпадение приемлемое.

Пример 3. Рассчитать коэффициенты уравнения процесса по результатам реализации плана ПФЭ³, представленным в таблице.

i	Основные столбцы			Вспомогательные столбцы				Выход процесса, мг/мл			
	x_{1i}	x_{2i}	x_{3i}	$x_{1i} \cdot x_{2i}$	$x_{2i} \cdot x_{3i}$	$x_{1i} \cdot x_{3i}$	$x_{1i} \cdot x_{2i} \cdot x_{3i}$	x_{i1}	x_{i2}	x_{i3}	\bar{x}_i
1	-	-	-	+	+	+	-	73	69	68	70,0
2	-	+	-	-	-	+	+	58	58	64	60,0
3	+	-	-	-	+	-	+	54	59	52	55,0
4	+	+	-	+	-	-	-	84	94	92	90,0
5	-	-	+	+	-	-	+	100	106	109	105,0
6	-	+	+	-	+	-	-	98	90	97	95
7	+	-	+	-	-	+	-	77	85	78	80,0
8	+	+	+	+	+	+	+	105	95	100	100,0
9	0	0	0					89	83	86	86,0

C_{i0}	C_{10} 45°C	C_{20} 18%	C_{30} 300 об/МИН
λ_i	λ_1 5°C	λ_2 8%	λ_3 10 об/МИН

Примечание: ПФЭ³ насчитывает 8 опытов. Дополнительный опыт в центре эксперимента обязательно планируют, но используют его результат лишь при статистическом анализе работоспособности уравнения.

Рассчитаем коэффициенты уравнения по средним результатам 8 опытов, пользуясь соответствующими формулами:

$$b_0 = \frac{\sum_{i=1}^N x_i}{N} = \frac{70+60+55+90+105+95+80+100}{8} = 81,87 = 81,9;$$

$$b_1 = \frac{\sum_{i=1}^N x_{1i}x_i}{N} = \frac{-70-60+55+90-105-95+80+100}{8} = -0,62 = -0,6;$$

$$b_2 = \frac{\sum_{i=1}^N x_{2i}x_i}{N} = \frac{-310+345}{8} = \frac{35}{8} = 4,37 = 4,4;$$

$$b_3 = \frac{\sum_{i=1}^N x_{3i}x_i}{N} = \frac{-275+380}{8} = \frac{105}{8} = 13,12 = 13,1;$$

$$b_{12} = \frac{\sum_{i=1}^N x_{1i}x_{2i}x_i}{N} = \frac{-290+365}{8} = \frac{75}{8} = 9,37 = 9,4;$$

$$b_{23} = \frac{\sum_{i=1}^N x_{2i}x_{3i}x_i}{N} = \frac{-335+320}{8} = \frac{-15}{8} = -1,87 = -1,9;$$

$$b_{13} = \frac{\sum_{i=1}^N x_{1i}x_{3i}x_i}{N} = \frac{-345+310}{8} = \frac{-35}{8} = -4,37 = -4,4;$$

$$b_{123} = \frac{\sum_{i=1}^N x_{1i}x_{2i}x_{3i}x_i}{N} = \frac{-335+320}{8} = \frac{-15}{8} = -1,87 = -1,9.$$

Следовательно, уравнение процесса будет иметь вид

$$x = 81,9 - 0,6x_1 + 4,4x_2 + 13,1x_3 + 9,4x_1x_2 - 1,9x_2x_3 - 4,4x_1x_3 - 1,9x_1x_2x_3.$$

Убедимся, что полученное уравнение получено без арифметических ошибок ($N=N'$, $\hat{x}_i = \bar{x}_i$), для чего рассчитаем по полученному уравнению предсказанный результат какого-либо

опыта (например, восьмого) при координатах $x_{18} = 1, x_{28} = 1, x_{38} = 1$

$$\hat{x}_8 = 81,9 - 0,6 + 4,4 + 13,1 + 9,4 - 1,9 - 4,4 - 1,9 = 100,0 = \bar{x}_8.$$

Если оценки коэффициентов уравнения получены без округления, то при тщательной реализации МНК невязки должна равняться нулю. Поскольку оценки коэффициентов при расчете округляются в данном уравнении до десятых, т.е. на каждой оценке можно потерять (или приобрести) максимально 0,05, то можно нуля не получить.

Если предположить почти невероятную ситуацию, когда каждый член уравнения при расчете \hat{x}_i будет из-за округления только терять (или приобретать) 0,05, то предельная ошибка округления предсказанного результата опыта может достигнуть величины $|\Delta x_{пред}| = 0,05N' = 0,05 \cdot 8 = 0,4$. Иными словами, если $|\hat{x}_i - \bar{x}_i| \leq 0,4$, то арифметических ошибок в определении оценок коэффициентов уравнения не имеется. В противном случае следует проверить расчет оценок коэффициентов уравнения.

Если бы по каким-то соображениям оценки коэффициентов рассчитывались с округлением до единицы, то $\hat{x}_8 = 82 - 1 + 4 + 13 + 9 - 2 - 4 - 2 = 99 \text{ мг/мл}$.

$$|\Delta x_{пред}| = 0,5 \cdot N' = 4 \text{ мг/мл}.$$

Поскольку $|\hat{x}_8 - \bar{x}_8| = |99 - 100| = 1 < |\Delta x_{пред}| = 4$, то и в этом случае арифметических ошибок не сделано (см. пример 4).

Пример 4. Получить уравнение по следующим результатам плана ПФЭ³.

i	x_{1i}	x_{2i}	x_{3i}	x, кг/м ³					S ² (x _{si}) (кг/м ³)
				x_{1i}	x_{2i}	x_{3i}	x_{4i}	\bar{y}_u	
1	-	-	-	75	55	60	70	65,0	83,33
2	-	+	-	90	85	95	70	85,0	116,67
3	+	-	-	55	65	150	70	60,4	83,33
4	+	+	-	125	115	105	95	110,0	166,67
5	-	-	+	70	60	75	75	70,0	133,33

6	-	+	+	115	100	95	110	105,0	83,33
7	+	-	+	75	80	90	75	80,0	50,0
8	+	+	+	155	160	145	140	150,0	83,33
	18%	40°	80 л/с	100	90	105	85	95,0	83,3
	3%	6°	20 л/с						

Решение:

1. Рассчитаем средние результаты:

$$\bar{x}_i = \frac{1}{m} \sum_{k=1}^m x_{ki};$$

$$\bar{x}_1 = \frac{1}{4} (75 + 55 + 60 + 70) = 65,0 \text{ кг/м}^3 \quad \text{и т.д. (см.}$$

таблицу).

2. Выделяющихся результатов среди поверхностей нет, поэтому анализ и наличие грубых ошибок не производим.

$$3. \quad b_0 = \frac{1}{N} \sum_{i=1}^N \bar{x}_i = \frac{725,0}{8} = 90,625 = 90,62;$$

$$b_1 = \frac{1}{N} \sum_{i=1}^N x_{1i} \bar{x}_i = \frac{-325,0 + 400,0}{8} = 9,36;$$

$$b_2 = \frac{1}{8} (450,0 - 275,0) = 21,88;$$

$$b_3 = \frac{1}{8} (-320,0 + 405,0) = 10,60;$$

$$b_{12} = \frac{1}{N} \sum_{i=1}^N x_{1i} x_{2i} \bar{x}_i = \frac{1}{8} (-330,0 + 395,0) = 8,12;$$

$$b_{13} = \frac{1}{8} (380,0 - 345,0) = 4,38;$$

$$b_{23} = \frac{1}{8} (380,0 - 345,0) = 4,38;$$

$$b_{123} = \frac{1}{N} \sum_{i=1}^N x_{1i} x_{2i} x_{3i} \bar{x}_i = \frac{1}{8} (365,0 - 360,0) = 0,62.$$

Получили:

$$x = 90,62 + 9,3x_1 + 21,88x_2 + 10,60x_3 + 8,12x_1x_2 + 4,38x_1x_3 + 4,38x_2x_3 + 0,62x_1x_2x_3.$$

Проверим: $\hat{x}_i = \bar{x}_i$ реализуется при $N = N'$, например, $\hat{x}_8 = 149,98 \approx \bar{x}_8$; $\Delta_8 = -0,02$ явно меньше $\Delta x_{пред}$.

Пример 4а.

Проверить значимость оценок коэффициентов полученного в примере 4 уравнения.

Решение:

1. Рассчитываем оценки дисперсии

$$S^2(x_{ki}) = \frac{\sum_{k=1}^m (x_{ki} - \bar{x}_i)^2}{m-1};$$

$$S^2(x_{k1}) = \frac{(75,0-65,0)^2 + (55,0-65,0)^2 + (60,0-65,0)^2 + (70,0-65,0)^2}{4-1} =$$

$$\frac{250,0}{3} = 83,33 \text{ (кг/м}^3\text{)}^2; f_1 = m - 1 = 3.$$

Подобным образом рассчитаны оценки $S^2(x_{ki})$ для восьми опытов плана и для опыта в центре эксперимента (см. таблицу к примеру 4).

2. Проверим однородность оценок $S^2(x_{ki})$:

$$G = \frac{S^2(x_{ki})_{max}}{\sum_{i=1}^N S^2(x_{ki})} = \frac{166,67}{883,32} = 0,1887,$$

$$G_{кр}(P; N + 1; f) = G_{кр}(0,95; 9; 3) = 0,3584 \text{ (приложение$$

2).

Поскольку $G < G_{кр}$, оценки однородны.

3. Средневзвешенная оценка дисперсии $S^2(x_k)$:

$$S^2(x_k) = \frac{\sum_{i=1}^{N+1} S^2(x_{ki}) f_i}{\sum_{i=1}^{N+1} f_i} = \frac{\sum_{i=1}^{N+1} S^2(x_{ki})}{N+1} = \frac{883,32}{9} =$$

$$98,146 \text{ (кг/м}^3\text{)}^2,$$

$$f = (N - 1)(m - 1) = 27.$$

4. Оценка дисперсии $S^2(\bar{x})$:

$$S^2(\bar{x}) = \frac{S^2(x_k)}{m} = \frac{98,146}{4} = 24,536 \text{ (кг/м}^3\text{)}^2; f = 27.$$

5. Оценки дисперсии $S^2(b)$:

$$S^2(b) = \frac{S^2(\bar{x})}{N} = \frac{24,536}{8} = 3,0670 \text{ (кг/м}^3\text{)}^2; f = 27.$$

6. Доверительная ошибка $\varepsilon(b)$:

$$\varepsilon(b) = t(P; f)S(b) = t(0,95; 27)\sqrt{3,067} = 2,06 \cdot 1,75 =$$

$$3,65 \text{ кг/м}^3;$$

$$t(0,95; 27) = 2.06 \text{ (приложение 7)}$$

7. Оказался незначительным лишь один коэффициент

$$b_{123} = 0,62 < 3,65.$$

Этот коэффициент следует исключить из уравнения.

Примечание:

1) при анализе однородности оценок дисперсии и расчете средневзвешенной оценки $S^2(x_k)$ и $S^2(\bar{x})$ учитывали оценку дисперсии дополнительного опыта в центре эксперимента (общее число опытов $N+1=9$);

2) поскольку оценки коэффициентов уравнения рассчитывались по восьми опытам, $S^2(b) = \frac{S^2(\bar{x})}{N}$.

Пример 4б.

Проверить адекватность и работоспособность полученного в примерах 4 и 4а уравнения.

Решение:

1. Оценка дисперсии неадекватности:

$$S_{ad}^2 = \frac{\sum_{i=1}^N (\hat{x}_i - \bar{x}_i)^2}{N - N'} = \frac{\sum_{i=1}^N (\pm b_{123})^2}{N - N'} = \frac{8 \cdot (\pm 0,62)^2}{8 - 7} = 3,075 \text{ (кг/м}^3\text{)}^2;$$

$$f_{ad} = N - N' = 1.$$

2. $F = \frac{S^2(\bar{x})}{S_{ad}^2} = \frac{24,5365}{3,075} = 7,98;$

$F_T(0,95; 27; 1) = \sim 250 > 7,98$ - уравнение адекватно описывает экспериментальные данные.

3. Можно ли полученное уравнение считать хорошей математической моделью процесса? Работает ли оно в центре эксперимента?

$|\hat{x}_0 - \bar{x}_0| = |90,62 - 95| = 4,38 \text{ кг/м}^3$ - невязка в центре эксперимента.

Доверительная ошибка невязки:

$$\varepsilon(\hat{x}_0 - \bar{x}_0) = t(P; f)S(b_0 - \bar{x}_0) = t(P; f)\sqrt{S^2(b_0) + S^2(\bar{x}_0)} = t(P; f)S(x_k)\sqrt{\frac{1}{mN} + \frac{1}{m_0}};$$

Здесь $S^2(b_0) = S^2(b) = \frac{S^2(\bar{x})}{N} = \frac{S^2(x_k)}{mN}$; $S^2(\bar{x}_0) = \frac{S^2(x_k)}{m_0}$;
 $m_0 = 4$;

$$\varepsilon(\hat{x}_0 - \bar{x}_0) = 2,06\sqrt{98,15\left(\frac{1}{8 \cdot 4} + \frac{1}{4}\right)} = 10,82 \text{ кг/м}^3, t(0,95; 27) = 2,06.$$

Уравнение можно считать хорошей математической моделью, так как оно работоспособно, $\varepsilon(\hat{x}_0 - \bar{x}_0) > |\hat{x}_0 - \bar{x}_0|$, и адекватно, $F_T > F$.

Пример 4в.

Рассчитать программу оптимизации по полученному в предыдущем примере уравнению, несколько упростив его (I) и, для тренировки, по линейной части уравнения (II):

$$I. \quad x = 90,6 + 9,4x_1 + 21,9x_2 + 10,6x_3 + 8,1x_1x_2 + 4,4x_1x_3 + 4,4x_2x_3.$$

Решение:

I.1. Линеаризованное уравнение:

$$\Delta = \sum_{i=1}^3 \frac{\partial x}{\partial x_i} \Delta x_i = (9,4 + 8,1x_2 + 4,4x_3)\Delta x_1 + (21,9 + 8,1x_1 + 4,4x_3)\Delta x_2 + (10,6 + 4,4x_1 + 4,4x_2)\Delta x_3.$$

I.2. За начало оптимизации примем условия опыта 8, давшего самый высокий результат ($x_{18} = +1; x_{28} = +1; x_{38} = +1$),

$$\Delta x_{8-l} = (9,4 + 8,1x_{28} + 4,4x_{38})\Delta x_{1(8-l)} + (21,9 + 8,1x_{18} + 4,4x_{38})\Delta x_{2(8-l)} + (10,6 + 4,4x_{18} + 4,4x_{28})\Delta x_{3(8-l)} = 21,9\Delta x_{1(8-l)} + 34,4\Delta x_{2(8-l)} + 19,4\Delta x_{3(8-l)}.$$

Список принятых обозначений

i – номер опыта;

x_i - критерий оптимальности;

x_s - параметр, случайная величина;

s - результаты повторностей одного и того же опыта;

Δ_s - величина абсолютной погрешности;

$\overline{\Delta}_s$ - средняя арифметическая величина абсолютного отклонения;

\overline{x}_s' - средний результат всех опытов эксперимента;

m - число повторностей;

Px - величина плотности распределения вероятности;

σ - среднеквадратичное отклонение;

σ^2 - дисперсия случайной величины;

$S^2(x_s)$ - среднеарифметический результат опыта;

S - экспериментальная оценка среднеквадратичного отклонения;

$S(x_s)$ – оценка среднеквадратичного отклонения единичного результата;

$S^2(\bar{x})$ – оценка среднеквадратичного отклонения среднего результата;

f - число степеней свободы;

k - влияние какого-либо фактора на выход процесса;

Δ_i - величина невязки;

\hat{x}_i - предсказанное значение выхода процесса в i -ом опыте;

\bar{x}_i - полученное в i -ом опыте значение выхода процесса;

$p(\Delta_i)$ - плотность вероятности появления любой i -ой невязки;

$\prod_{i=1}^N p(\Delta_i)$ - функция правдоподобия;

Φ - минимизируемая функция;

$F_T(P; f_a; f)$ – критерий Фишера;

F – адекватность уравнения экспериментальным данным;

S_a^2 - дисперсия неадекватности (неточности) предсказания экспериментальных данных уравнением;

$S^2(\bar{x})$ - средневзвешенная оценка дисперсии воспроизводимости среднего результата;

N – число экспериментальных точек,

N' - число коэффициентов аппроксимирующей зависимости;

x_{ij} - безразмерную величину i -го фактора в j -ом опыте;

C_{ij} – величина i – го фактора в j – ом опыте;

C_{i0} – центр эксперимента;

λ_i - интервал варьирования (изменения) фактора от 1-го уровня фактора до последующего $(1+1)$ -го уровня $(1=1 \div L)$;

λ_i - интервалы варьирования;

N - число опытов планов ПФЭ 2^n ;

x_i^+ - значение верхнего уровня факторов в безразмерном выражении;

x_i^- - значение нижнего уровня факторов в безразмерном выражении;

$\varepsilon(b_i)$ – доверительная ошибка коэффициентов;

$t(P; f)$ – критерий Стьюдента.

УЧЕБНОЕ ИЗДАНИЕ

**ОСНОВЫ МАТЕМАТИЧЕСКОГО АНАЛИЗА
ТЕХНОЛОГИЧЕСКИХ ПРОЦЕССОВ**

УЧЕБНОЕ ПОСОБИЕ

Для студентов вузов

ЛР № 020524 от 02.06.97

Подписано в печать 13.03.09. Формат 60 × 84^{1/16}

Бумага типографская. Гарнитура Times

Уч.-изд.л. 10. Тираж 500 экз.

Заказ № 45

Кемеровский технологический институт пищевой
промышленности

650056, г. Кемерово, б-р Строителей, 47

ПЛД № 44-09 от 10.10.99

Отпечатано в лаборатории множительной техники
Кемеровского технологического института пищевой
промышленности

650010, г. Кемерово, ул. Красноармейская, 52