

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ  
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ  
«ДОНСКОЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

Л.В. Борисова, В.П. Димитров, Е.М. Зубрилина

# ОСНОВЫ ТЕОРИИ ЭКСПЕРИМЕНТА. ПОСТРОЕНИЕ МАТЕМАТИЧЕСКИХ МОДЕЛЕЙ

Учебное пособие

Ростов-на-Дону  
ДГТУ  
2021

УДК 001.891

Б82

*Рецензент*

доктор технических наук, профессор, заведующий кафедрой

«Общетехнические дисциплины» *А.Г. Пастухов*

(Белгородский государственный аграрный университет им. В.Я. Горина)

**Борисова, Людмила Викторовна.**

Б82 Основы теории эксперимента. Построение математических моделей: учеб. пособие / Л.В. Борисова, В.П. Димитров, Е.М. Зубрилина; Донской гос. техн. ун-т. – Ростов-на-Дону, ООО «ДГТУ-Принт», 2021. – 84 с.

Рассматриваются условия применения математической теории планирования эксперимента в научных и инженерных исследованиях и основные положения математического моделирования сельскохозяйственных процессов. Представлены алгоритмы построения планов экспериментов, методы статистической оценки данных, полученных в ходе эксперимента.

Приведены основные понятия теории эксперимента, задачи дисперсионного, корреляционного, регрессионного анализа и методы их решения, планирование и обработка результатов многофакторного эксперимента.

Предназначено для специалистов, бакалавров и магистрантов, изучающих дисциплину «Основы научных исследований», «Основы научных исследований в отрасли», «Основы научных исследований в АПК». Пособие может быть полезным также аспирантам и научным сотрудникам, которые занимаются планированием экспериментальных исследований сложных, многофакторных систем.

УДК 001.891

Печатается по решению редакционно-издательского совета  
Донского государственного технического университета

© Борисова Л.В., Димитров В.П.,  
Зубрилина Е.М.  
© ДГТУ, 2021

## Введение

*Планирование эксперимента* – область знания, связанная с построением и оптимизацией математических моделей.

*Математическая модель* – это совокупность математических объектов (чисел, символов, множеств и т. д.) и связей между ними, отражающая важнейшие для исследователя свойства изучаемого объекта.

Существует два способа получения математической модели объекта: теоретический (аналитический) и экспериментальный. В первом случае исследователь строит модель на основе известных ему представлений об объекте моделирования с использованием законов и правил математики, логики, физики и других наук. При втором способе математическую модель объекта получают с помощью эксперимента. Обычно это – уравнение (система уравнений), связывающее выходные и входные параметры объекта, например, потери зерна за зерноуборочным комбайном как функция подачи хлебной массы в комбайн.

К математическому моделированию в технике прибегают в тех случаях, когда опыта и интуиции проектировщика недостаточно для принятия решения, а также при решении задач, требующих больших объемов вычислений.

В технике широко применяется и другой вид моделирования – физическое макетирование объектов в уменьшенном или упрощенном виде. У физического и математического моделирования есть свои преимущества и недостатки. Преимуществами математического моделирования являются отсутствие затрат на изготовление физических макетов, возможность проведения практически неограниченного числа экспериментов, в том числе на критических режимах, которые привели бы к разрушению физического макета. Вместе с тем в ряде случаев затраты на создание адекватной математической модели могут превосходить затраты на макетирование, а достоверность результатов математического моделирования часто зависит от принимаемых допущений и неизвестных коэффициентов математической модели.

Математическое моделирование для анализа физических процессов в сплошных средах с заданными краевыми условиями выполняется с помощью дифференциальных уравнений в частных производных. Система уравнений, как правило, известна (уравнения Ламе для механики упругих сред; уравнения теплопроводности для термодинамики; уравнения Навье – Стокса для гидравлики и т. д.), но точное решение удается получить лишь для простых частных случаев, поэтому обычно требуется построить приближенную дискретную модель. Примерами применения таких моделей являются расчеты на прочность ответственных деталей конструкций, исследование аэродинамических свойств летательных аппаратов, расчет тепловых режимов работы деталей и конструкций, проектирование гидротехнических сооружений. Среди математических методов моделирования широкую известность получили *методы конечных разностей, конечных элементов и граничных элементов*. В настоящее время существует большое число программных средств, реализующих эти методы на ЭВМ, среди них имеются как специализированные (рассчитанные на определенный класс конструкций и вид нагрузок), так и универсальные.

Получаемая при дискретизации пространства, описываемого системой дифференциальных уравнений в частных производных, аппроксимирующая система алгебраических уравнений имеет высокий порядок даже для простых технических систем: число неизвестных составляет  $10^3$ – $10^6$  и более. Поэтому при моделировании достаточно сложных технических объектов переходят к более укрупненному моделированию – обычно используя систему обыкновенных дифференциальных уравнений.

Основой математического моделирования сложных технических систем, как правило, является теория массового обслуживания, автоматического управления, графов, системный анализ. Объектами моделирования с использованием этих средств являются, например, системы автоматического регулирования технологическими процессами, гибкие автоматизированные производства, другие сложные системы.

# 1. ОСНОВНЫЕ ПОЛОЖЕНИЯ СИСТЕМНОГО АНАЛИЗА И ИССЛЕДОВАНИЯ ОПЕРАЦИЙ

Для анализа и синтеза систем можно предложить последовательность действий:

- 1) формулировка целевой функции системы;
- 2) определение места системы в иерархии систем и ограничений на ее функционирование, параметры элементов;
- 3) определение входных и выходных потоков системы;
- 4) формулировка критерия эффективности функционирования системы (математическая запись целевой функции);
- 5) определение структуры системы и ее элементов;
- 6) определение связей между элементами системы;
- 7) построение математической модели системы;
- 8) проверка адекватности модели;
- 9) использование математической модели (для оптимизации целевой функции, определения чувствительности целевой функции к изменению параметров системы и др.).

## 1.1. Целевая функция

Условие выбора лучшего (оптимального) варианта решения какой-либо задачи называют *критерием оптимальности*, или *целевой функцией*.

В качестве целевой функции могут приниматься экономические, технические, экологические и другие показатели. Обычно в технике предпочтение отдается экономическим показателям как наиболее общим, поскольку в них могут быть учтены все (или почти все) другие показатели.

Из экономических показателей одним из самых общих является прибыль  $\Pi = (Ц - З) \cdot A$  (где  $Ц$  – цена единицы продукции;  $З$  – затраты на производство единицы продукции;  $A$  – объем выпуска продукции).

При оптимизации параметров целевая функция принимает вид:

$$\Pi \rightarrow \max.$$

Сформулировать целевую функцию иногда сложно. Например, при проектировании или модернизации зерноуборочного комбайна можно сформулировать такие целевые функции: себестоимость, надежность, маневренность и др., а при ограничениях по ряду параметров: удельное давление на почву, дробление зерна, габариты комбайна и др.

Если при проектировании технических объектов можно выделить один параметр, которому отдается безусловное предпочтение и который наиболее полно характеризует свойства проектируемого объекта, то этот параметр можно принять за целевую функцию. В этом случае задача проектирования сводится к *однокритериальной* задаче оптимизации целевой функции при условии соблюдения ограничений. При этом одна часть параметров попадает под категорию ограничений, а другая часть параметров, на которые не накладываются ограничения, принимается такой, какой получилась при оптимизации целевой функции.

Задачи могут быть однокритериальными и многокритериальными. Так, например, комбайн характеризуется многими параметрами, определяющими его качество и ценность. Некоторые из этих параметров необходимо увеличивать, а другие уменьшать до возможных пределов. Ограничения и связи между отдельными параметрами комбайна приводят к необходимости идти на компромисс и выбирать для каждой характеристики не наилучшее возможное значение, а такое, при котором и другие важные характеристики будут иметь приемлемое значение. Поэтому при выборе окончательного варианта нельзя ограничиваться сравнением по одной какой-либо характеристике, а необходимо принимать во внимание всю их совокупность. Задачи проектирования, проводимые по нескольким критериям оптимизации, носят название *многокритериальных*, или *задач векторной оптимизации*.

Известные методы векторной оптимизации непосредственно или косвенно сводят решаемые задачи к задачам скалярной (одномерной) оптимизации, т. е. к функции:

$$F(x) = \Phi(F_i(x)), i = 1, 2, \dots, n,$$

где  $F_i(x)$  – частные критерии, тем или иным способом объединяющиеся в составной критерий  $\Phi$ , который затем максимизируется или минимизируется.

Решение объективно, если составной критерий объясняет физическую суть функционирования системы. Однако это не всегда возможно. Поэтому составной критерий чаще носит формальный (субъективный) характер.

Обобщенные критерии могут быть аддитивными, мультипликативными, минимаксными и др.

Если оптимизация ведется без учета статистического разброса характеристик, то соответствующий критерий оптимальности называют *детерминированным*, в противном случае – *стохастическим*. Последний более полно отражает свойства объекта проектирования.

*Принцип справедливой компенсации частных критериев* формулируется так: справедливым следует считать такой компромисс, при котором суммарный уровень снижения одних критериев не превышает суммарный уровень увеличения других критериев.

*Аддитивные критерии.* Целевая функция образуется сложением нормированных значений частных критериев. Оценивают не с натуральными значениями критериев, а с их нормированными значениями. Выбор нормирующего делителя логически обосновывается.

Принципы выбора нормирующего делителя:

- по директивным значениям параметров, которые указаны в техническом задании;
- по максимальным значениям критериев или по разности между максимальным и минимальным значениями.

В общем виде аддитивный критерий можно представить как

$$F(x) = \sum (C_i F_i(x)/F_{0i}) = \sum (C_i f_i(x)),$$

где  $C_i$  – весовой коэффициент частного критерия;  $f_i(x)$  – нормированное значение частного критерия;  $F_{0i}$  – нормирующий коэффициент.

Весовые коэффициенты учитывают значимость частных критериев, при этом  $\sum C_i = 1$ .

Основным недостатком аддитивного критерия является то, что при его использовании может наблюдаться взаимная компенсация частных критериев. Это значит, что существенное уменьшение одного из критериев (вплоть до нуля) может быть покрыто возрастанием другого критерия. Для ослабления этого недостатка следует вводить ограничения на предельные значения частных критериев. Другой недостаток аддитивного критерия – субъективность назначения весовых коэффициентов  $C_i$ .

*Мультипликативные критерии.* В ряде случаев целесообразно оперировать не с абсолютными, а с относительными изменениями частных критериев. Тогда *принцип справедливой относительной компенсации частных критериев* формулируется следующим образом: справедливым следует считать такой компромисс, при котором суммарный уровень относительного снижения одних критериев не превышает суммарный уровень относительного увеличения других критериев.

Условие оптимальности имеет вид:

$$\sum [\Delta F_i(x)/F_i(x)] = 0,$$

где  $\Delta F_i(x)$  – приращение величины  $i$ -го критерия;  $F_i(x)$  – первоначальная величина  $i$ -го критерия.

Так как  $\Delta F_i(x) \ll F_i(x)$ , то

$$\begin{aligned} \sum [\Delta F_i(x)/F_i(x)] &= \sum d[\ln F_i(x)] = \\ d[\sum (\ln F_i(x))] &= d[\ln \prod F_i(x)] = 0, \end{aligned}$$

где  $\prod$  – произведение по индексу  $i$ .

Принцип относительной справедливой компенсации приводит к обобщенному мультипликативному критерию:

$$F(x) = \ln[\prod F_i(x)] \rightarrow \min$$

или

$$F(x) = \prod F_i(x) \rightarrow \min.$$



Эта формула справедлива для равноценных частных критериев.

Для неравноценных частных критериев используется формула

$$F(x) = \prod (F_i(x))^{C_i},$$

где  $C_i$  – весовой коэффициент  $i$ -го частного критерия.

*Минимаксные критерии.* В теории многокритериальной оптимизации особое место занимает принцип компромисса, основанный на идее равномерности. На базе этого принципа работают минимаксные (максиминные) критерии.

Основываясь на идее равномерного компромисса, стараются найти такие переменные проектирования, при которых нормированные значения всех частных критериев становятся равными между собой, т. е.

$$F_i(x) = k, i = 1, 2, \dots, n,$$

или, с учетом весовых коэффициентов:

$$C_i f_i(x) = k.$$

Однако при больших значениях  $n$  из-за сложных взаимосвязей трудно добиться выполнения этих соотношений. В этом случае оказывается полезным применение *принципа минимакса (максимина)*. Суть его в том, что нужно выбрать такое решение  $X^0 \in X$  ( $X = x_1, \dots, x_n$ ), на котором реализуется максимум (минимум) из минимальных (максимальных) значений частных критериев:

$$F(X^0) = \max \min\{f_i(x)\} \rightarrow \max$$

или

$$F(X^0) = \min \max\{f_i(x)\}, \text{ т. е. } \max\{f_i(x)\} \rightarrow \min.$$

Такой принцип выбора  $X^0$  иногда носит название принципа «гарантированного результата». Он заимствован из теории игр, где является основным принципом.

Рассмотрим принцип максимина на следующем примере. Для уборки урожая предприятие может выбирать различные стратегии уборки, т. е. иметь различное количество комбайнов, менять соотношение раздельного и прямого комбайнирования и др. В зависимости от погодных условий и урожайности возможны различные потери урожая. В годы с высокой урожайностью нужно иметь большее количе-

ство комбайнов и тракторов, чтобы убрать урожай в срок, а в годы с низкой урожайностью – меньшее. В табл. 1.1 строками являются стратегии уборки, а столбцами – условия уборки или ситуации с вероятностями  $p_i (\sum p_i = 1)$ . Элементами таблицы будет прибыль  $\mathcal{E}_{ij}$  (где  $i$  – стратегия уборки;  $j$  – погодная ситуация). В центре каждой ячейки таблицы показано значение прибыли в данной ситуации и выбранной стратегии уборки, а в верхнем правом углу потери прибыли по отношению к лучшей стратегии в данной ситуации.

Выделим для каждой стратегии уборки элемент, где прибыль минимальна. Из этих минимальных прибылей выбираем стратегию, где минимальная прибыль максимальна – это стратегия 1, которая выбрана по максиминному критерию (максимум из минимальных значений).

Если в качестве критерия выбрать минимум максимальных потерь прибыли, то наилучшей будет стратегия 4. В этом случае вместо матрицы прибылей составляют матрицу потерь, которая, например, для первого столбца имеет значения 3, 1, 0, 2 для стратегий 1–4.

Таблица 1.1

Величина прибыли при различных стратегиях уборки

Стратегия уборки	Номер ситуации и ее вероятность			
	1 ( $p_1$ )	2( $p_2$ )	3( $p_3$ )	4( $p_4$ )
1	7 3	10 7	6 0	5 1
2	9 1	13 4	4 2	6 2
3	10 0	15 2	3 3	6 0
4	8 2	17 0	4 2	5 1

Кроме критерия максимина (минимакса) в теории игр предлагаются критерии, рассчитанные на индивидуальные склонности лица, принимающего решение:

– максимальный средний эффект при равновероятных условиях уборки:  $\sum \mathcal{E}_{ij} \rightarrow \max$  (в нашем примере – это равноценные стратегии 3 и 4, которые обеспечивают сумму прибылей для всех ситуаций в размере 34 единиц);

- максимум суммы произведения вероятности максимального эффекта на величину максимального эффекта и произведения вероятности минимального эффекта на величину минимального эффекта;
- минимум потерь с учетом вероятностей всех возможных условий уборки;
- максимум экономического эффекта (в нашем случае прибыли) с учетом всех возможных условий уборки.

$$\sum p_i \Delta_{ij} \rightarrow \max.$$

Этот критерий является наилучшим с экономической точки зрения.

На практике необходимо учитывать:

1. Выбор критерия оптимальности является сложной методологической проблемой и, как правило, может быть неоднозначным.
2. Аддитивный критерий выбирают тогда, когда существенное значение имеют абсолютные значения критериев.
3. Мультипликативный критерий целесообразно использовать тогда, когда существенную роль играет изменение абсолютных значений частных критериев, при вариации вектора переменных  $X$ .
4. Если перед разработчиком стоит задача достижения равенства нормированных значений конфликтных частных критериев, то оптимальное проектирование следует проводить по минимаксному критерию.

### *Экспертные оценки*

*1. Метод ранжирования.* Пусть экспертизу проводит группа из  $m$  экспертов. Каждый эксперт расставляет критерии  $F_i(x)$  в порядке убывания их важности. Ранг 1 получает оценку (преобразованный ранг)  $n$ , ранг 2 – оценку  $(n-1)$  и т. д., где  $n$  – число частных критериев.

Зная преобразованный ранг  $i$ -го критерия у  $k$ -го эксперта  $r_{ik}$ , определяют весовые коэффициенты:

$$C_i = \sum r_{ik} / \sum \sum r_{ik}.$$

С учетом авторитетности каждого эксперта  $m_k (\sum m_k = 1)$

$$C_i = \sum m_k r_{ik} / \sum \sum m_k r_{ik}.$$

*2. Метод приписывания баллов.* Эксперты оценивают важность частного критерия баллами по шкале 0 – 10, при этом разрешается

оценивать важность дробными величинами или приписывать одну и ту же величину нескольким критериям. Зная балл  $h_{ik}$   $i$ -го критерия и  $k$ -го эксперта, преобразованный ранг  $i$ -го критерия у  $k$ -го эксперта находят по формуле

$$r_{ik} = h_{ik} / \sum h_{ik} .$$

На достоверность экспертизы существенно влияют такие факторы, как число экспертов, уровень их компетентности, состав вопросов, самочувствие эксперта и др.

*Оптимальность по Парето.* Пусть  $X^p$  и  $X^q$  – два разных решения задачи минимизации. Если векторный критерий  $F(X^p)$  по отношению к векторному критерию  $F(X^q)$  удовлетворяет условию  $F(X^p) \leq F(X^q)$  для всех  $i=1, \dots, n$  и  $F_j(X^p) < F_j(X^q)$  хотя бы для одного  $j = 1, \dots, n$ , то говорят, что векторный критерий  $F_i(X^p)$  доминирует над  $F_i(X^q)$ , а область критериев называют областью согласия.

Точка  $X^0$  называется эффективной или решением, оптимальным по Парето, если не существует ни одной точки  $X$ , такой что  $F_i(X)$  меньше или равно  $F_i(X^0)$  для всех  $i = 1, \dots, n$  и  $F_i(X) < F_i(X^0)$  хотя бы для одного  $j = 1, \dots, n$ .

Эффективная точка может быть не единственной. Множество эффективных точек называется областью решений, оптимальных по Парето. Оптимальность по Парето означает, что нельзя дальше уменьшить значение одного из частных критериев, не увеличив при этом хотя бы одного из остальных.

## 1.2. Оптимизация целевой функции

Одной из классификаций задач оптимизации является деление их на задачи концептуальной, структурной и параметрической оптимизации.

Примером задачи концептуальной оптимизации является выбор типа привода для мотовила жатки комбайна: механический, гидравлический, электрический. Задачи этого типа обычно трудноформализуемы. Так, в приведенном примере, кроме поддающихся формализации стоимости привода, эксплуатационных затрат на ремонт и техническое обслуживание и др., нужно дополнительно учитывать, например,

надежность поставщиков комплектующих и другие трудноформализуемые факторы.

Пример структурной оптимизации: определение числа ступеней зубчатого редуктора с заданным передаточным отношением.

Параметрическая оптимизация предполагает наличие параметров объекта, которые входят в целевую функцию и ограничения; при определенных значениях этих параметров (называемых оптимальными) целевая функция принимает минимальное (или максимальное) значение.

Рассматривая объект проектирования на макроуровне, представим его в виде «черного ящика», на входе и выходе которого имеются:

$x_i$  – управляемые параметры объекта;

$y_i$  – выходные параметры объекта;

$z_i$  – неуправляемые параметры объекта;

$f_i$  – случайные воздействия на объект.

На управляемые параметры  $x$  могут накладываться ограничения, тогда параметры  $x_i$  должны принадлежать допустимой области  $D$ . Множество точек, образующих область  $D$ , называется выпуклым, если для любой пары точек отрезок прямой линии, соединяющий их, также полностью принадлежит этому множеству (рис. 1.1).

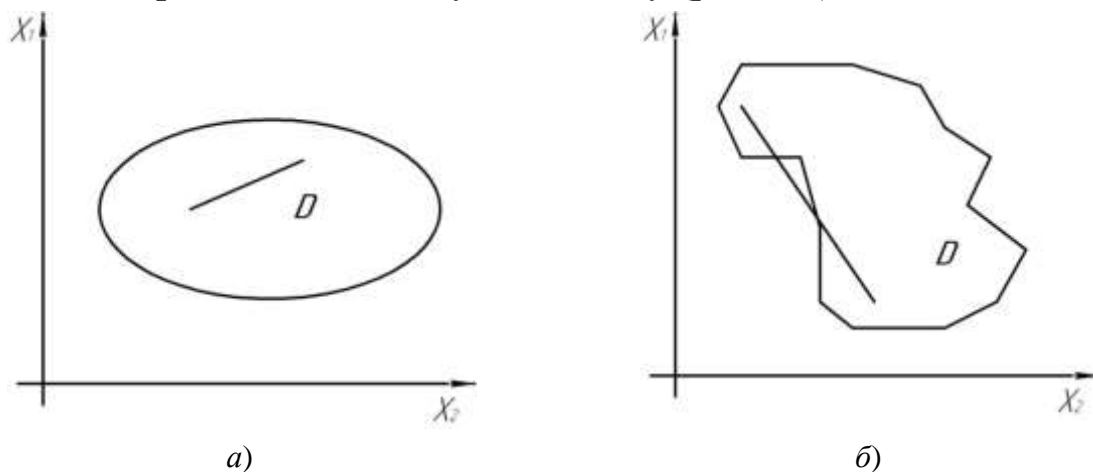


Рис. 1.1. Выпуклая (а) и невыпуклая (б) допустимые области переменных  $x_1$  и  $x_2$

Обычно в задачах оптимизации рассматривают задачу минимизации целевой функции, т. е.  $F(x) \rightarrow \min$ .

Задачу максимизации целевой функции легко можно свести к задаче минимизации, если изменить знак целевой функции на проти-

воположный, максимум функции  $F(x)$  будет при тех же значениях аргумента, что и минимум —  $F(x)$ .

Математически задачу параметрической оптимизации можно записать в общем виде:

$$F_j(x_1, x_n) \rightarrow \min, j = 1, k - \text{целевые функции};$$

$$g_p(x_1, x_n) \leq C_p, p = 1, m - \text{ограничения типа «неравенство»};$$

$$h_q(x_1, x_n) = B_q, q = 1, l - \text{ограничения типа «равенство»};$$

где  $x_i$  — оптимизируемые параметры ( $i = 1, 2, 3, \dots, n$ );  $C_p$  и  $B_q$  — коэффициенты.

Частным случаем ограничений типа  $g_p$  являются ограничения  $a_i \leq x_i \leq b_i$ .

В зависимости от количества и вида целевых функций и ограничений выделяют частные задачи оптимизации:

1) однокритериальной оптимизации ( $k = 1$ ) и многокритериальной (векторной) оптимизации;

2) однопараметрической (одномерной) оптимизации ( $n = 1$ ) и многопараметрической (многомерной) оптимизации ( $n > 1$ );

3) безусловной (отсутствуют ограничения  $g_p$  и  $h_q$ ) и условной (имеется хотя бы одно ограничение) оптимизации;

4) одноэкстремальной (целевая функция имеет один экстремум) и многоэкстремальной (целевая функция имеет несколько экстремумов) оптимизации;

5) задача параметрической оптимизации с дополнительными требованиями, чтобы управляемые параметры принимали только дискретные значения, называется задачей дискретной оптимизации. Если все  $x_i$  должны быть целыми числами, то задача называется задачей целочисленного программирования.

Рассмотрим кратко частные случаи задачи оптимизации с позиций сложности их решения.

Задачи *многокритериальной (векторной)* оптимизации, т. е. оптимизации одновременно нескольких целевых функций, значительно сложнее задач однокритериальной оптимизации. При этом зачастую не существует значений, при которых все целевые функции одновременно принимали бы оптимальное значение. Чаще наблюдается другая картина: при одних значениях параметров оптимальна первая целевая

функция, при других – вторая и т. д. В этих условиях проектировщику приходится идти на компромисс: допускать отклонение от оптимального значения одних целевых функций, чтобы получить приемлемое значение других целевых функций. Обычно этот компромисс достигается присвоением весовых коэффициентов значимости целевых функций и сведением всех целевых функций в одну. Таким образом, в задачах многокритериальной оптимизации решение обычно отсутствует и проектировщику приходится вносить элемент субъективности при назначении весовых коэффициентов или допустимых значений целевых функций.

Задачи многопараметрической оптимизации также сложнее одномерных задач, но сложность здесь другого рода: есть математические методы программирования, позволяющие получить решение многомерной задачи с требуемой точностью, но с увеличением числа переменных резко возрастают вычислительные затраты и, соответственно, время на решение задачи. Например, чтобы рассчитать целевую функцию двух переменных, каждая из которых может принять три различных значения, необходимо сделать  $3^2 = 9$  расчетов. Соответственно для тридцати переменных, каждая из которых может принимать 100 различных значений, получаем  $100^{30}$  расчетов – такое огромное количество расчетов не под силу даже современным супер-ЭВМ. Конечно, такое количество расчетов нужно сделать только при прямом переборе всех возможных вариантов. Существуют методы (градиентный, покоординатного спуска и др.), которые резко уменьшают число расчетов, но тем не менее проблема длительности решения многомерной задачи имеет место.

Задачи *условной оптимизации* являются более сложными, чем безусловной. Для задачи условной оптимизации в общем виде, т. е. при любом виде критерия оптимальности и ограничений, отсутствует метод «быстрого решения» даже однокритериальной задачи без использования той или иной формы перебора вариантов решения и постепенного приближения к окончательному решению.

Если целевая функция или ограничения нелинейны, то имеем *задачу нелинейного программирования*. Если функция и область  $D$  выпуклы – задачу *выпуклого программирования*.

Если критерии оптимальности и ограничения являются линейными функциями параметров  $x$ , то имеем задачу *линейного программирования*.

$$\begin{aligned} F(x) &= c_0 + c_1x_1 + \dots + c_nx_n \rightarrow \min \\ a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n R_1 b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n R_2 b_2 \\ &\dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n R_m b_m \\ x_i &\geq 0, \end{aligned}$$

где  $F(x)$  – критерий оптимальности, являющийся линейной функцией переменных  $x_i$ ;  $R_i$  – один из знаков:  $=, \geq, \leq$ ;  $a, b$  и  $c$  – коэффициенты.

Если критерий оптимальности  $F(x)$  – квадратическая функция, а ограничения – линейные функции, то имеем задачу *квадратического программирования*.

Задачей *сепарабельного программирования* называется такая задача, в которой критерии оптимальности и ограничения имеют вид сумм  $n$  функций только одной переменной.

### 1.3. Методы оптимизации

Целевая функция может иметь локальные минимумы (точки  $x_1$ ,  $x_2$  и др.) и глобальный минимум (точка  $x_3$ ) (рис. 1.2).

*Унимодальная функция* – функция, которая в допустимой области имеет один минимум (максимум). Функция на рис. 1.2 не унимодальная.

Большинство методов поиска оптимума основано на постепенном приближении к оптимуму от задаваемой исследователем или случайно выбираемой самой программой оптимизации начальной точки поиска. Пусть это будет точка  $a$  на рис. 1.2, и решается задача поиска минимума  $F(x)$ . Тогда при решении задачи оптимизации путем постепенного спуска локальный минимум будет в точке  $x_2$ .



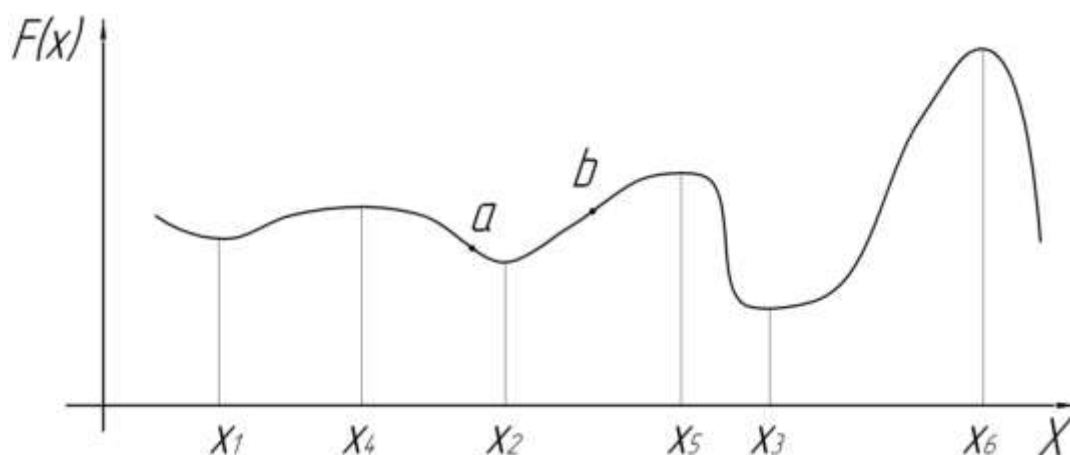


Рис. 1.2. Пример многоэкстремальной функции

Если начальной точкой поиска будет точка  $b$ , то получим то же решение – точка  $x_2$ . И вообще, если в качестве начальной точки поиска минимума будет задана любая точка из диапазона  $x_4 - x_5$ , то будет получено одно и то же решение –  $x_2$ , которое может быть ошибочно принято за минимум (глобальный же минимум находится в другой точке –  $x_3$ ).

Для того чтобы попасть в точку глобального минимума  $x_3$ , начальную точку поиска нужно задать из диапазона  $x_5 - x_6$ . Поскольку проектировщику заранее неизвестно ни число, ни положение точек экстремумов функции, существует опасность принять локальный экстремум за глобальный. Чтобы избежать этой ошибки, требуется повторять оптимизацию для разных начальных точек поиска с каким-то шагом вдоль оси  $X$ , при этом требуемая величина шага неизвестна (в нашем примере она должна быть меньше диапазона  $x_5 - x_6$ , так как при большом шаге можно "проскочить" глобальный оптимум; при малом шаге возрастает число расчетов).

В связи с этим важную роль играет предварительное представление проектировщика о поведении оптимизируемой функции, в том числе о ее унимодальности или многоэкстремальности.

Согласно теореме Вейерштрасса, задача параметрической оптимизации разрешима, если критерий оптимальности  $F(x)$  является непрерывной функцией, а допустимая область образует замкнутое ограниченное множество. Для вывода условий оптимальности, т. е. для получения решения, в большинстве случаев необходимо, чтобы функ-

ция  $F(x)$  была непрерывно дифференцируемой. Так, в случае безусловной минимизации унимодальной одномерной функции  $F(x)$  для того, чтобы точка являлась оптимальным решением, необходимо и достаточно, чтобы в этой точке первая производная минимизируемой функции равнялась нулю, а ее вторая производная была величиной положительной.

Обобщая это условие на многомерный случай для задачи безусловной минимизации многопараметрической унимодальной функции  $F(x)$ , можно сформулировать следующее условие оптимальности: чтобы точка  $x$  являлась экстремумом или точкой перегиба, необходимо, чтобы в этой точке градиент функции равнялся нулю (т. е. равнялись нулю все частные производные функции по всем переменным  $x_i$ ). Для определения характера точки экстремума необходимо дополнительное исследование.

Если на переменные  $x_i$  наложены ограничения, то задача усложняется, так как минимум может быть на границе области допустимых значений  $x$  (гиперпараллелепипеда), где условие равенства нулю всех частных производных функции может не выполняться. Это значит, что оптимальное решение может находиться как внутри области, так и на границе гиперпараллелепипеда. Такая ситуация приводит к тому, что нужно решить очень большое число задач безусловной оптимизации. Очевидно, что такой подход к решению задачи оптимизации требует больших вычислительных затрат. Поэтому в некоторых случаях целесообразно перейти от исходной задачи к задаче безусловной оптимизации.

*Методы покоординатного подъема (спуска).* Логическим развитием методики одномерного поиска было бы последовательное изменение каждого проектного параметра до тех пор, пока не будет достигнут максимум целевой функции. По завершении этой процедуры для всех переменных можно вернуться к первой и посмотреть, нельзя ли еще более усовершенствовать решение. Этот метод, называемый методом покоординатного подъема, не всегда позволяет найти оптимальное решение. Метод покоординатного подъема не применим, если линии

уровня имеют точки излома. Математическим эквивалентом обрыва на поверхности, образуемого целевой функцией, являются те ее места, где поставлены ограничения.

*Градиентные методы* оптимизации используют информацию о градиентах. Рассмотрим систему независимых единичных векторов  $e_1, e_2, \dots, e_n$ , направленных вдоль осей координат  $x_1, x_2, \dots, x_n$ , являющихся в то же время проектными параметрами. Вектор градиента произвольной целевой функции  $F(x_1, x_2, \dots, x_n)$  имеет вид:

$$\frac{\partial F}{\partial x_1} e_1 + \frac{\partial F}{\partial x_2} e_2 + \dots + \frac{\partial F}{\partial x_n} e_n,$$

где частные производные вычисляются в рассматриваемой точке.

Этот вектор направлен вверх, в сторону подъема; обратный ему вектор указывает направление спуска. Единичный вектор градиента имеет вид:

$$v_1 e_1 + v_2 e_2 + v_3 e_3 + \dots + v_n e_n,$$

где

$$v_j = \frac{\frac{\partial F}{\partial x_j}}{\sum_{i=1}^n \left[ \left( \frac{\partial F}{\partial x_i} \right)^2 \right]^{0,5}}.$$

Когда характер целевой функции хорошо известен, можно вычислить компоненты вектора градиента дифференцированием. Если же таким способом частные производные получить не удастся, то можно найти приближенные значения в окрестности рассматриваемой точки:

$$\frac{\partial F}{\partial x_i} = \frac{F(x_1, \dots, x_i + \Delta, \dots, x_n) - F(x_1, \dots, x_i, \dots, x_n)}{\Delta},$$

где  $\Delta$  – небольшое смещение в направлении  $x_i$ .

При определении минимума необходимо для каждой точки  $i$  поиска определить градиент и сделать шаг в направлении антиградиента (т. е. в направлении, обратном градиенту):

$$x_i^{(n+1)} = x_i^{(n)} - a \frac{\partial F}{\partial x_i},$$

где  $a > 0$  – длина рабочего шага;  $i = 1, \dots, n$ .

Выбор того или иного значения  $a$  зависит от априорных представлений о форме гиперповерхностей равного уровня целевой функции. Существенное ускорение процесса поиска оптимума при сохранении простоты определения шага может быть достигнуто в случае использования комбинированного метода: шаг принимается постоянным и достаточно большим при движении вдали от оптимума, а после входа в зону оптимума предусматривается уменьшением шага в 2, 4 раза и более.

Метод оптимизации, в основу которого положена идея движения по самой крутой тропе, называется методом наискорейшего спуска. Вектор градиента перпендикулярен линии уровня и указывает направление к новой точке в пространстве проектирования.

Градиентный метод можно использовать применительно к любой унимодальной функции.

В некоторых случаях задачу условной оптимизации можно свести к задаче безусловной оптимизации, заменяя переменную  $x$  другой переменной, такой, на которую не наложены ограничения. Например, задача параметрической оптимизации с ограничениями типа линейных неравенств

$$F_i(x_1, x_n) \rightarrow \min, a_i \leq x_i \leq b_i$$

сводится к задаче безусловной оптимизации

$$\Phi(z) \rightarrow \min$$

при помощи функций преобразования вида

$$x_i = a_i + (b_i - a_i) \sin^2 z_i (-\infty < z_i < \infty).$$

Задача оптимизации вида

$$F_i(x_1, x_n) \rightarrow \min \text{ при } x_i \geq 0$$

сводится к задаче безусловной оптимизации при использовании функций преобразования вида  $x_i = (z_i)^2$ .

Для ограничения вида  $-1 \leq x_i \leq 1$  можно использовать функцию преобразования  $x_i = \sin z_i$  и т. д.

Однако использование функций преобразования для перехода от задачи с ограничениями к задаче безусловной оптимизации усложняет процесс оптимизации:

- может возрасти степень нелинейности функции  $\Phi(z)$  по сравнению с функцией  $F(x)$ ;
- функция  $\Phi(z)$  может быть неограниченной для некоторых значений новых переменных;
- функция  $\Phi(z)$  может стать периодической по новым переменным.

Для сведения задачи условной минимизации к задаче безусловной минимизации широко применяются методы штрафных функций.

*Методы штрафных функций.* Привести общую задачу к задаче на безусловный минимум можно, построив вспомогательную функцию вида

$$M'' = M + \Delta M,$$

где  $\Delta M$  – штрафная добавка, такая, что ее величина за границами допустимой области резко возрастает.

Функцию  $M''$  называют штрафной, а методы, основанные на ее использовании, – методами штрафных функций.

Структура функции  $M''$  такова, что поверхности уровня (геометрическое место точек в пространстве проектирования, в которых величина этой функции одинакова) у этой функции в большей степени искривлены, чем у исходной функции. Когда функция в некоторых направлениях изменяется значительно медленнее, чем во всех остальных, то говорят, что она имеет «овраг». Наличие у целевой функции «оврага» представляет большие трудности при решении оптимизационных задач. Обычно, чем ближе условный минимум функции к безусловному минимуму, тем глубже «овраг» функции. Зачастую начальные значения выбирают заведомо «не очень хорошими», не обеспечивая достаточной близости условного и безусловного минимумов, а потом уточняют.

Используют различные виды штрафных добавок, например:

$$\Delta M = \sum_{i=1}^S \lambda_i \varphi_i(X)^2 + \sum_{i=1}^r \lambda_{i+5} \Psi_i(X)^2 n_i^2;$$

$$\Delta M = \sum_{i=1}^r \frac{\lambda_i}{\Psi_i(X)};$$

$$\Delta M = \sum_{i=1}^r \lambda_i (\Psi_i(X) - c_i)^2,$$

где  $n_i$  – коэффициенты, равные нулю, когда  $\Psi_i \leq 0$ , и единице, когда  $\Psi_i > 0$ ;  $\lambda_i, c_i$  – настроечные параметры.

Предложено несколько модификаций метода штрафных функций, например метод множителей Лагранжа.

Каждой сложной задаче проектирования присущи свои специфические особенности, затрудняющие применение традиционных алгоритмов оптимизации. Поэтому тщательный выбор подходящего алгоритма позволяет сэкономить и машинное время, и усилия, затрачиваемые на решение задачи. Выбирая алгоритм, необходимо:

1. Проанализировать топологические особенности поверхности, описываемой целевой функции. Так, если поверхность имеет гладкие складки, не рекомендуется применять градиентные методы.

Если есть основания считать поверхность мультимодальной, то правильной будет выбрать в пространстве проектирования несколько начальных точек и убедиться, что во всех случаях получается одно и то же решение. При обнаружении нескольких локальных оптимумов конструкцию следует разрабатывать с учетом лучшего из них.

2. Изучить характер проектных параметров: как правило, большинство параметров могут принимать любые значения.

3. Проанализировать особенности алгоритма. Не рекомендуется использовать методы, требующие частных производных в аналитическом виде, для целевых функций, полученных экспериментально. Кроме того, если структура производных, полученных аналитически, сложна, то само дифференцирование может резко увеличить стоимость подготовки задачи и вероятность ошибок программиста.

Так как всегда желательно испробовать для решения данной задачи несколько разных алгоритмов, то заслуживает внимания "пакетный" подход, позволяющий переходить с одного алгоритма на другой.

#### **1.4. Примеры однокритериальной однопараметрической задачи безусловной оптимизации**

Однокритериальная однопараметрическая задача имеет вид  $F(x) \rightarrow \min$ .

*Расчет оптимального числа комбайнов в хозяйстве.* Пусть требуется определить оптимальное число комбайнов  $N_{\text{опт}}$  в хозяйстве при следующих исходных данных: площадь хозяйства  $S(\text{га})$ , цена комбайна

$\Pi_k$ (руб.), урожайность зерновых  $U$ (ц/га), цена зерна  $\Pi_3$ (руб/ц), суточная выработка одного комбайна  $W$ (га/сут).

Данная задача принадлежит к классу оптимизационных, что видно из следующих рассуждений.

Известно, что после наступления полной спелости зерно теряет свою связь с колосом и под действием ветра, дождя, а также ударов планок мотовила осыпается на землю. Это явление известно под названием «биологические потери» урожая. Установлено, что в среднем за сутки биологические потери урожая составляют 1% от первоначального; есть культуры, например овес и горох, средние биологические потери которых доходят до 2% в сутки, есть устойчивые к осыпанию сорта пшеницы, у которых потери равны 0,5% в сутки.

Рассмотрим постановку задачи оптимизации. Если урожай не убирается (число комбайнов  $N_k = 0$ ), то потери урожая  $\Pi$  равны стоимости урожая. С увеличением числа комбайнов сроки уборки сокращаются и, следовательно, сокращаются потери урожая, но растут затраты на уборку  $Z$ . Эти зависимости отображены графически на рис. 1.3.

В качестве целевой функции  $F(N_k)$  можно взять максимум прибыли, получаемой хозяйством. Можно показать, что максимуму прибыли соответствует минимум суммы затрат на уборку  $Z$  и потерь урожая  $\Pi$ , т.е.

$$F(N_k) = Z + \Pi \rightarrow \min.$$

Конкретизируем целевую функцию, выразив ее через число комбайнов  $N_k$  в хозяйстве. Сначала выразим затраты на уборку через число комбайнов. Так как затраты на горючее, заработную плату и ремонт не зависят от числа комбайнов (они зависят от величины убираемой площади), то в затратах на уборку учитывается только переменная их часть, т. е. отчисления на амортизацию  $Z_a$  и затраты на хранение  $Z_{xp}$  комбайнов:

$$Z = Z_a + Z_{xp}.$$

В свою очередь,

$$Z_a = 0,01 \cdot a \cdot \Pi_k \cdot N_k,$$

где  $a$  — годовые амортизационные отчисления для комбайнов ( $a = 8-10\%$ );

$$Z_{\text{хр}} = Z_x \cdot N_k,$$

где  $Z_x$  — затраты на хранение одного комбайна в год, руб.

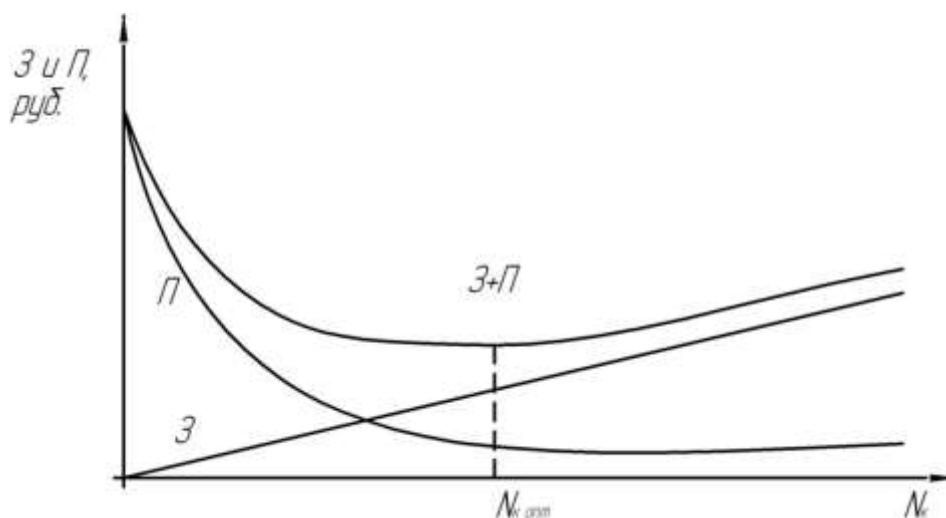


Рис. 1.3. Зависимости затрат на уборку  $Z$  и потерь урожая  $\Pi$  от числа комбайнов  $N_k$

Потери урожая определим по известной формуле:

$$\Pi = 0,5 \cdot K_{\Pi} \cdot U \cdot T S^2 \Pi_3,$$

где  $K_{\Pi}$  — коэффициент интенсивности потерь урожая в сутки,  $K_{\Pi} = (0,005-0,02)$ ; в среднем  $K_{\Pi} = 0,01$  (1% в сутки);  $T$  — продолжительность уборки, сут.

$$T = S / (W_k N_k),$$

где  $W_k N_k$  — суточная выработка всех комбайнов.

Выражение для  $\Pi$  принимает вид:

$$\Pi = 0,5 K_{\Pi} U S^2 \Pi_3 / (W_k N_k).$$

С учетом этого целевая функция имеет вид:

$$F(N_k) = 0,01 \cdot \Pi_k \cdot N_k + Z_x \cdot N_k + 0,5 \cdot K_{\Pi} \cdot U \cdot S^2 \cdot \Pi_3 / (W_k \cdot N_k) \rightarrow \min.$$

Для определения точки минимума функции  $F(N_k)$  нужно приравнять нулю ее первую производную по  $N_k$  и вычислить корни:

$$F'(N_k) = 0,01 \cdot a \cdot \Pi_k + Z_x - 0,5 \cdot K_{\Pi} \cdot U \cdot S^2 \cdot \Pi_3 / (W_k \cdot N_k^2) = 0,$$

откуда

$$N_{\text{копт}} = S \sqrt{\frac{0,5 \cdot K_{\Pi} \cdot U \cdot \Pi_3 \cdot 100}{W_k (0,01 \cdot a \cdot \Pi_k + Z_x)}}.$$



Предложенная методика определения оптимального числа комбайнов не учитывает ряд факторов: неодновременность созревания урожая на разных полях, разномарочность комбайнового парка, влияние погодных условий на уборку, колебания урожайности по годам и полям, применение одновременно двух способов уборки и др. Эти особенности могут быть учтены в более сложной модели, например имитационной.

*Задача определения оптимальных размеров бункера заданного объема.* Постановка задачи: спроектировать бункер из стального листа в виде конуса (без крышки) заданного объема  $V$  с минимальным расходом стали.

Наличие задачи оптимизации в этом случае можно предположить, если рассмотреть разные варианты бункеров одного и того же объема (рис. 1.4).

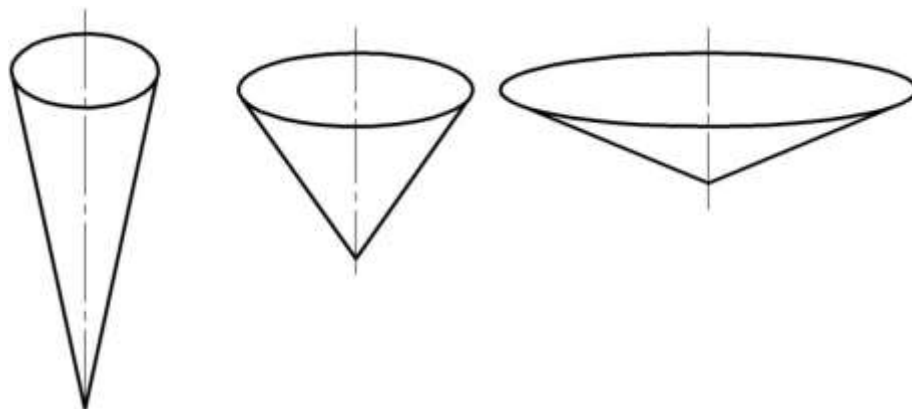


Рис.1.4. Варианты конусного бункера заданного объема

Так как расход металла на изготовление бункера при неизменной толщине стенок пропорционален площади его боковой поверхности, то в качестве критерия оптимальности принимаем  $S_{\text{бок}} \rightarrow \min$ .

Учитывая обозначения, принятые на рис 1.5, площадь боковой поверхности определяем по формуле

$$S_{\text{бок}} = \pi RL.$$

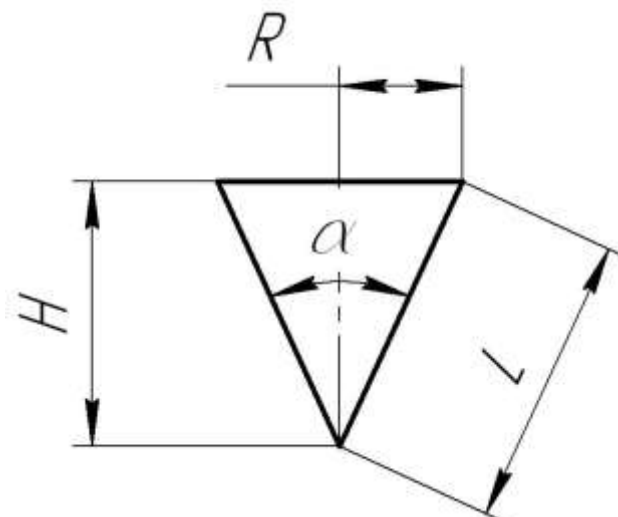


Рис. 1.5. Схема бункера

Выразим  $S_{\text{бок}}$  не через две ( $R$  и  $L$ ), а через одну неизвестную, например  $R$ , используя то, что объем бункера задан; тогда получим одномерную (однопараметрическую) задачу оптимизации.

По теореме Пифагора  $L = (R^2 + H^2)^{0.5}$ , тогда

$$S_{\text{бок}} = \pi R (R^2 + H^2)^{0.5}.$$

Из формулы  $V = \pi R^2 \cdot H/3$  выразим  $H$  через  $R$ :

$$H = 3V/(\pi R^2).$$

Произведя подстановку, получим

$$S_{\text{бок}} = \pi R \{ R^2 + [3V/(\pi R^2)]^2 \}^{0.5}.$$

Для поиска оптимального значения  $R^\circ$  необходимо определить производную  $S_{\text{бок}}$  и, приравняв ее нулю, найти корни уравнения, один из которых и будет искомым значением  $R^\circ$ .

После этого можно найти  $H^\circ$  и угол  $\alpha = \arctg (R^\circ + H^\circ)$ .

## 1.5. Статистическое моделирование (метод статистических испытаний)

Этот вид моделирования применяется для процессов, в которых происходят случайные события. Метод статистических испытаний, или метод Монте-Карло, – это численный метод решения задач с использованием моделирования случайных величин. До появления ЭВМ этот метод не имел широкого применения, так как моделировать случайные

величины вручную очень трудоемкая работа. Особенность метода – простая структура вычислительного алгоритма. Обычно составляется программа для осуществления одного случайного испытания. Затем это испытание повторяется  $N$  раз, причем каждый опыт не зависит от всех остальных, а результаты всех опытов усредняются.

Погрешность вычислений по методу Монте-Карло, как правило, пропорциональна величине  $(D/N)^{0.5}$  (где  $D$  – некоторая постоянная;  $N$  – число испытаний). Ясно, что для уменьшения погрешности в 10 раз (т. е. для получения в ответе еще одного верного десятичного знака), нужно увеличить  $N$  в 100 раз. Считается, что метод Монте-Карло эффективен при решении тех задач, в которых устраивает результат со сравнительно небольшой точностью (5–10%). Удастся значительно увеличить точность решения многих задач, выбрав способ расчета, которому соответствует малое значение  $D$ .

Метод Монте-Карло позволяет моделировать любой процесс, зависящий от случайных факторов. Наиболее часто метод используют для расчета систем массового обслуживания, качества и надежности изделий, а также вычисления интегралов, решения оптимизационных задач.

Для применения метода Монте-Карло нужно уметь разыгрывать случайные величины в заданной области. Для этого используют таблицы случайных чисел, генераторы случайных чисел и методы получения псевдослучайных чисел.

Числа, получаемые по какой-либо формуле генератором ЭВМ, называются псевдослучайными. На получение каждого числа затрачивается всего несколько операций. Поэтому скорость генерирования случайных чисел имеет тот же порядок, что и скорость работы ЭВМ. Кроме того, нужно лишь один раз проверить «качество» такого генератора, затем его можно много раз безбоязненно использовать при решении однотипных задач.

Недостатки метода: ограниченность количества псевдослучайных чисел, так как последовательность чисел, вычисляемых на ЭВМ, обязательно периодическая; необходимость проведения большого чис-

ла испытаний для получения решения, достаточно близкого к оптимальному, т. е. наличие медленной сходимости к экстремуму.

*Оценка надежности простейших технических.* Рассмотрим упрощенную схему подсистемы передвижения комбайна (рис. 1.6).

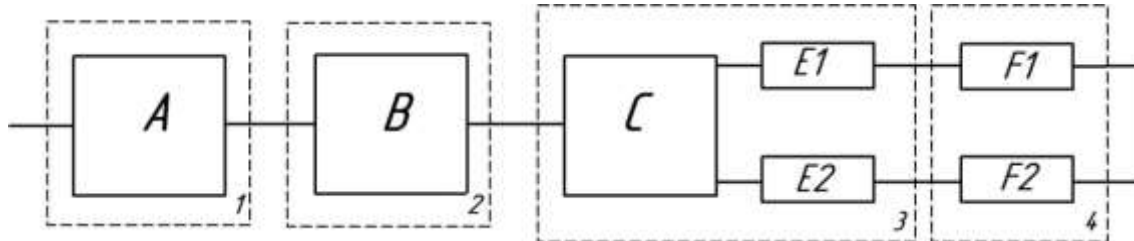


Рис. 1.6. Подсистема передвижения комбайна

Блок 1 – двигатель; содержит один условный элемент A и не срабатывает при отказе этого элемента.

Блок 2 – вариатор хода; содержит один условный элемент B и не срабатывает при отказе этого элемента.

Блок 3 – мост ведущих колес; содержит три условных элемента: C – коробка передач; E<sub>1</sub> E<sub>2</sub> – бортовые редукторы; не срабатывает при отказе любого из этих элементов.

Блок 4 – колеса; содержит два условных элемента F<sub>1</sub>, и F<sub>2</sub> и не срабатывает при отказе любого из элементов.

Используя метод Монте-Карло, оценить вероятность  $P$  безотказной работы системы, зная вероятности безотказной работы элементов:  $P(A) = 0,95$ ;  $P(B) = 0,7$ ;  $P(C) = 0,8$ ;  $P(E_1) = P(E_2) = 0,75$ ;  $P(F_1) = P(F_2) = 0,95$ .

В качестве оценки искомой надежности примем:

$$P = N_1/N,$$

где  $N$  – общее число испытаний;  $N_1$  – число испытаний с благоприятным исходом испытаний (в которых система работает безотказно).

Результаты испытаний заносятся в табл. 1.2.

Таблица 1.2

## Расчетные данные для оценки надежности системы

Но- мер исп	Блок	A	B	C	E <sub>1</sub>	E <sub>2</sub>	F <sub>1</sub>	F <sub>2</sub>	A	B	C	E <sub>1</sub>	E <sub>2</sub>	F <sub>1</sub>	F <sub>2</sub>	Блок	Си- стема
1	1	0,83							+							+	-
	2		0,65							+						+	
	3			0,91	0,24	0,35					-	+	+			-	
	4						0,8	0,92						+	+	+	

Принцип заполнения таблицы покажем на примере первого испытания. На ЭВМ с помощью стандартного статистического пакета программ получаем семь случайных чисел (табл. 1.2), моделирующих состояние элементов системы A, B, C, E<sub>1</sub>, E<sub>2</sub>, F<sub>1</sub>, F<sub>2</sub>. Считаем, что если случайное число больше или равно вероятности события, то событие не наступило. По этому правилу разыгрываются события, состоящие в безотказной работе каждого из элементов системы.

Поскольку  $P(A) = 0,95$  и  $0,83 < 0,95$ , событие A наступило, т. е. элемент A, а значит и первый блок, в этом испытании сработали безотказно. В клетках таблицы ставим соответствующие знаки.

Поскольку  $P(B) = 0,7$  и  $0,65 < 0,7$ , событие B наступило, т. е. элемент B, а значит и второй блок, сработали безотказно в этом испытании. В клетках таблицы ставим соответствующие знаки.

Поскольку  $0,91 > P(C)$ ,  $0,24 < P(E_1)$ ,  $0,35 < P(E_2)$ , элементы E<sub>1</sub> E<sub>2</sub> сработали безотказно в этом испытании, а элемент C — отказал, а значит, отказал и третий блок. В клетках таблицы ставим соответствующие знаки.

Поскольку  $0,8 < P(F_1)$ ,  $0,92 < P(F_2)$ , элементы F<sub>1</sub> и F<sub>2</sub> сработают безотказно, а значит, сработает четвертый блок.

Так как третий блок отказал, то отказала и система в целом в этом испытании. В клетках таблицы ставим соответствующие знаки.

Аналогично разыгрываются остальные события (испытания).

## 2. ИНЖЕНЕРНЫЕ ЗАДАЧИ, РЕШАЕМЫЕ ЭКСПЕРИМЕНТАЛЬНО

В инженерной деятельности в области сельскохозяйственных машин решаются задачи:

- создания рациональных схем и конструкций машин;
- управления технологическими процессами;
- определения оптимальных условий технологических процессов.

До 90% информации, необходимой для решения этих задач, получают из экспериментальных исследований. При этом эксперименты должны быть построены и проведены таким образом, чтобы по их результатам можно было получить количественную оценку влияния различных факторов на интересующий исследователя показатель работы машины; получить подходящую математическую модель; определить оптимальные условия протекания данного процесса и т. д., при минимальном числе опытов и минимуме затрат трудовых и материальных ресурсов. Такие результаты могут быть получены только при тщательном планировании эксперимента.

### 2.1. Эксперимент и наблюдения

*Эксперимент* – это такой метод познания, при котором исследователь искусственно вызывает явления или изменяет условия так, чтобы лучше выяснить сущность явления, происхождение, причинность и взаимосвязь предметов и явлений.

*Наблюдение* – это количественная или качественная регистрация интересующих исследователя сторон развития явления, констатация наличия того или иного его состояния, признака или свойства.

Эксперимент, при котором исследователь может задавать значения факторов и поддерживать их с достаточной точностью на определенном уровне по определенной схеме (плану), называется *активным*.

Эксперимент, при котором исследователь не задает, а только регистрирует значения факторов, называется *пассивным*.

Считалось, что единственно правильной является методология однофакторного эксперимента. При этом предполагалось, что поддерживая с достаточной точностью постоянными независимые переменные (факторы) и изменяя последовательно первый, второй и т. д. факторы, исследователь может отыскать интересующую его зависимость. Переменные, которые исследователь не мог поддерживать на определенном уровне, оставляли без внимания как несущественные (не всегда имея на это основания). Однако такой способ экспериментальных исследований может быть использован в том случае, когда число факторов достаточно мало и есть возможность изменять их значения или же фиксировать на одном уровне по своему желанию.

Математическая теория эксперимента рекомендует одновременно изменять все исследуемые факторы согласно определенному *плану эксперимента*. Этим достигается резкое сокращение объема экспериментов и получение более достоверных зависимостей, учитывающих взаимодействие факторов.

Поскольку планирование эксперимента относится к оптимальному управлению, то к нему могут быть применены известные кибернетические принципы. Объект исследования при этом рассматривается как кибернетическая система в виде «черного ящика». Последний представляет собой систему связей, недоступную для наблюдения, так как о содержании и механизме процесса нам ничего не известно, или известно лишь частично. Известны только входы «ящика» – переменные, участвующие в процессе, и выходы – результаты процесса, называемые также откликом системы (рис. 2.1).



Рис. 2.1. Схема черного ящика

Выходные переменные  $Y_i$ , характеризуют интересующие исследователя параметры. Входные переменные  $X_i$ , которые известны и которыми можно управлять, называются управляемыми факторами. Кроме управляемых факторов  $X$  на «черный ящик» оказывают воздействие и случайные величины  $W_i$ , которыми исследователь управлять не может. Они образуют шум в системе и называются возмущающими факторами.

Например, при исследовании влияния частоты вращения молотильного барабана и зазоров на входе и выходе молотильного аппарата, при которых суммарные потери от необмолоченного, непросеянного и раздробленного зерна будут самыми малыми, частота вращения и зазоры на входе и выходе аппарата являются управляемыми факторами  $X_1$ ,  $X_2$  и  $X_3$ , а потери зерна – исследуемым параметром –  $Y$ . Другие факторы, такие как длина соломы, относительное содержание в ней зерна, диаметр барабана и т. д., являются возмущающими факторами  $W_i$ .

Изменяя входные переменные, можно наблюдать, как будет изменяться выход, а затем, статистически обрабатывая результаты замеров входных и выходных величин, получить математическую модель процесса и оптимальные значения входных величин.

*Статистические данные* – это множество численных значений одной или нескольких случайных величин, которыми описывают соответствующие признаки данной совокупности однородных элементов, называемой *генеральной совокупностью*. Например, растения данной культуры, выращиваемой при определенных условиях, образуют генеральную совокупность. Любой их размер является признаком генеральной совокупности. Однотипные детали, изготовленные при определенных условиях, образуют также генеральную совокупность.

Генеральная совокупность содержит конечное или бесконечное число элементов. Отдельные признаки генеральной совокупности могут быть описаны одномерными случайными величинами. Для одновременного описания двух или более признаков необходимо использовать систему случайных величин.



Изучение и анализ всех элементов генеральной совокупности практически невозможны. Поэтому при изучении ее признаков обычно используют одно подмножество случайно подобранных элементов генеральной совокупности. Полученное таким образом подмножество называется *выборкой*. Изучение какого-либо элемента выборки равносильно проведению одного опыта. Число элементов выборки  $n$  называют объемом выборки. Выборки с объемом  $n < 30$  условно называются малыми, а с объемом  $n > 30$  – большими. Чтобы судить достаточно уверенно об изучаемом признаке генеральной совокупности, выборка должна быть *репрезентативной* (представительной), т. е. достаточно большого объема. Полученные в выборке значения изучаемого признака генеральной совокупности должны быть *независимыми* друг от друга.

Увеличение объема выборки, с одной стороны, ведет к удорожанию экспериментального исследования, а с другой – к увеличению достоверности полученных результатов и сделанных на их основе выводов. Поэтому объем выборки необходимо согласовывать с планируемой достоверностью результатов.

Репрезентативность выборки можно обеспечить случайным отбором ее элементов из генеральной совокупности. Если же значения изучаемого признака получают в виде непрерывной записи (осциллограммы), то для получения независимых друг от друга отсчетов значения  $X$  величину интервалов  $t_0$  (рис. 2.2) для их замера определяют следующим образом.

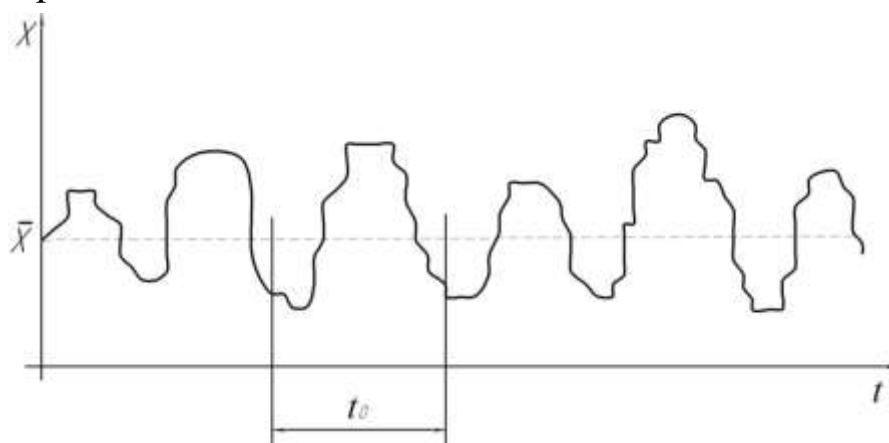


Рис. 2.2. Осциллограмма процесса

На осциллограмме проводят горизонтальную прямую, приблизительно соответствующую среднему значению изучаемой величины, и отмечают точки пересечения прямой с осциллограммой. Опытами установлено, что среднее число пересечений прямой в единицу времени  $f_0 = f/t$  (где  $f$  – число пересечений за время  $t$ ) стабилизируется за время, соответствующее  $f = 40\text{--}70$ . Тогда интервал  $t_0$ , т. е. время, через которое должен производиться отсчет, определяется формулой  $t_0 \geq 2 / f_0$ .

## 2.2. Измерения

Измерения могут быть классифицированы следующим образом:

- *по количеству однократных измерений*: разовые или однократные; многократные: равноточные, неравноточные, дискретные, непрерывные;
- *по отношению к измеряемой величине*: прямые; косвенные;
- *по связи измерений*: синхронные (по времени или пути); асинхронные.

Чем глубже техническое исследование, чем большие проблемы оно решает, тем совершеннее и точнее должны быть способы измерений и измерительная аппаратура. Большую часть измерительной аппаратуры выпускает промышленность, но во многих случаях для специфических измерений исследователь должен сам конструировать приборы, используя в готовом виде отдельные узлы заводских приборов.

*Точность измерений* есть степень соответствия результата измерений действительному значению измеряемой величины. Понятие точности связано с понятием ошибки: чем выше точность, тем меньше ошибка. Самые точные приборы не могут показать действительного значения измеряемой величины, их показания содержат ошибку.

*Абсолютная ошибка* – это разность между действительным значением измеряемой величины  $x$  и ее измеренным значением  $a$ . Практически под абсолютной ошибкой  $\Delta$  понимают разность между результатом измерения при помощи более точных приборов (образцо-

вых)  $a_{\text{обр}}$  и значением этой же величины, полученным прибором, применяемым в исследовании:

$$\Delta = x - a \approx a_{\text{обр}} - a.$$

Часто пользуются относительной ошибкой

$$\Delta_1 = \pm \Delta / x \approx \pm \Delta / a_{\text{обр}}$$

или  $\Delta_1 = \pm \Delta / x \cdot 100\% \approx \Delta / a_{\text{обр}} \cdot 100\%$ .

Ошибки измерений могут возникать из-за несовершенства измерительного прибора, влияния условий его эксплуатации, опытности наблюдателя и т. п.

Ошибки классифицируют как систематические, случайные и грубые промахи.

Систематическими называют такие ошибки, которые возникают из-за известных причин, действующих по определенным законам и, как правило, в определенном направлении. Их можно количественно определить и вносить в измерения соответствующие поправки. Систематические ошибки может определить сам исследователь, калибруя прибор.

При калибровке показания приборов, выбранных для измерений, сравнивают с показаниями приборов более высокой точности и устанавливают погрешности прибора. С этой же целью строят калибровочные кривые.

*Случайными ошибками* называют такие, причины которых неизвестны и которые невозможно учесть заранее. Случайные ошибки характеризуют точность измерений.

Такие ошибки можно выразить несколькими способами. Часто пользуются понятием предельной ошибки, под которой подразумевают наибольшую случайную ошибку при пользовании исправным прибором и при условии устранения систематических ошибок, т. е.

$$x = a \pm \Delta_{\text{п}},$$

где  $\Delta_{\text{п}}$  – поправка прибора (предельная ошибка).

Предельная случайная ошибка может быть определена из паспорта прибора, отметки о погрешности прибора на его шкале, калибровкой прибора, или, в крайнем случае, максимальная ошибка может быть принята равной половине наименьшего деления на шкале прибора.

При определении величины случайных ошибок, кроме предельной ошибки можно пользоваться статистической ошибкой, полученной неоднократными измерениями и обработкой результатов методами математической статистики.

Последовательность определения случайных ошибок: прибором измеряют несколько раз (например,  $n$ ) практически постоянную величину и находят ее среднее арифметическое:

$$a = (a_1 + a_2 + \dots + a_n)/n,$$

где  $a_i$  – результат любого измерения.

Затем вычисляют дисперсию измеренной величины:

$$\sigma^2 = \sum (a_i - a)^2 / (n - 1)$$

и среднее квадратическое отклонение или стандарт:

$$\sigma = \sigma^{0,5}.$$

Тогда наибольшая возможная статистическая ошибка с вероятностью 99,73% будет  $\Delta_n = \pm 3\sigma$  или относительная величина

$$\frac{\Delta_n}{a} = \pm 3\sigma/a; \quad \frac{\Delta_n}{a} \cdot 100\% = \pm \frac{3\sigma}{a} \cdot 100\%.$$

Такая ошибка будет иметь место только в том случае, если распределение ошибок подчиняется нормальному закону распределения случайной величины. При других законах распределения вероятная ошибка будет другой, обычно меньшей. Например, для равномерного распределения (весьма далекого от нормального), если его принять за нормальное (предельный случай), то и тогда вероятная ошибка будет лишь на 20% (т. е. 0,2) меньше ее точного значения. Для других видов распределения ошибка не превышает  $0,1\sigma$ .

Во многих экспериментах невозможно выделить систематическую и случайную составляющие ошибки измерений, так как нельзя прокалибровать измерительную аппаратуру, т. е. имеет место неопределенность. При этом считают, что неопределенность обусловлена лишь случайными ошибками и учитывается стандартными статистическими методами.

Например, при калибровке электрического тахометра производили замер частоты вращения синхронного электродвигателя с  $n = 1000 \text{ мин}^{-1}$  и получили показания:  $n_1 = 995$ ,  $n_2 = 1005$ ,  $n_3 = 1010$ ,  $n_4 = -990 \text{ мин}^{-1}$ .

Тогда

$$\begin{aligned}
n &= (995+1005+1010+990)/4=1000 \text{ мин}^{-1}; \\
\sigma &= [(995 - 1000)^2 + (1005-1000)^2 + (1010-1000)^2 + \\
&\quad + (990 - 1000)^2 ]/(4-1)= 9,1 \text{ мин}^{-1}; \\
\Delta_n &= \pm 3\sigma = \pm 39,1 = 27,3 \text{ мин}^{-1}; \\
\frac{\Delta_n}{n} \cdot 100\% &= \left( \frac{27,3}{1000} \right) \cdot 100\% = 2,73\%.
\end{aligned}$$

Таким образом, при  $\Delta_n = \pm 3\sigma$  в 99,73% замеров точность измерения частоты вращения составляет  $n \pm 0,0273 n$ .

*Измерение переменных величин.* В ходе экспериментальных исследований, как правило, измеряют переменные величины, т. е. такие, значения которых под воздействием случайных причин непрерывно изменяются относительно средней арифметической. При измерении таких величин дисперсия измерений включает в себя как меру изменчивости переменной величины, так и случайную ошибку измерений. Чтобы правильно определить меру изменчивости переменной величины, из результатов измерений необходимо выделить ошибку измерений. Для этого прибором многократно измеряют какую-либо постоянную величину, по своему значению приближающуюся к предполагаемому среднему арифметическому значению переменной. Получив стандарт прибора  $\sigma_n$  при измерении постоянной величины, измеряют в опыте и вычисляют стандарт переменной величины  $\sigma_o$  по формуле

$$\sigma_o = (\sigma_d - \sigma_n)^{0,5},$$

где  $\sigma_d$  – стандарт, содержащий оба рассеяния, вычисленный по результатам измерения переменной величины.

### 2.3. Ошибка и неопределенность эксперимента в целом

Рассмотренные виды ошибок имеют место при измерении какой-либо одной величины измерительной системой, включающей чувствительный элемент, индикаторную часть и наблюдателя. Однако в большинстве случаев в ходе эксперимента несколькими приборами измеряется несколько величин и для получения конечного результата эти измерения необходимо обработать, используя математические операции: сложение, умножение и т. п. Может оказаться, что, получив в процессе эксперимента довольно точные значения измеряемых величин, в конце

последовательности вычислений получим такие ошибки, которые фактически сведут на нет цель эксперимента.

Поэтому необходимо оценивать точность опыта в целом на основании вычисления предельной ошибки опыта.

Правила вычисления предельной относительной ошибки опыта:

1. Ошибка суммы заключена между наибольшей и наименьшей из относительных ошибок слагаемых. Обычно учитывается или наибольшая ошибка, или средняя арифметическая величина.

2. Ошибка произведения или частного от деления равна сумме относительных ошибок сомножителей или, соответственно, делимого и делителя.

3. Ошибка  $n$ -й степени какого-либо основания в  $n$  раз больше относительной ошибки основания.

4. Ошибка синуса или косинуса равна произведению абсолютного значения котангенса (или соответственно тангенса) измеренного угла на предельную абсолютную ошибку угла, выраженную в радианах.

5. Ошибка тангенса или котангенса равна частному от деления двойной абсолютной ошибки угла (в радианах) на синус двойного угла.

Вычисленная общая ошибка опыта учитывает и неопределенность опыта.

Например, при вычислении производительности агрегата (га/смену) используется формула

$$W_{\text{см}} = 0,1BT_p = 0,1B \left( \frac{S}{T_s} \right) T_p,$$

где  $B$  – захват агрегата, м;  $\frac{S}{T_s} = V$  – скорость прохождения пути  $S$  за время  $T_s$ , км/ч;  $T_p$  – чистое рабочее время за смену, ч.

В соответствии со вторым правилом

$$\frac{\Delta_{\Pi}(W_{\text{см}})}{W_{\text{см}}} = \pm \left( \frac{\Delta_{\Pi}(B)}{B} + \frac{\Delta_{\Pi}(S)}{S} + \frac{\Delta_{\Pi}(T_s)}{T_s} + \frac{\Delta_{\Pi}(T_p)}{T_p} \right).$$

С учетом использованной аппаратуры имеем: при измерении  $B$  и  $S$  20-метровой рулеткой

$$\frac{\Delta_{\pi}(B)}{B} = \frac{\Delta_{\pi}(S)}{S} = (0,02 - 0,03)\% \approx 0,025\%,$$

при измерении  $T_s$  и  $T_p$  секундомером

$$\frac{\Delta_{\pi}(T_s)}{T_s} = \frac{\Delta_{\pi}(T_p)}{T_p} = (0,04 - 0,07)\% \approx 0,055\%.$$

Тогда

$$\frac{\Delta_{\pi}(W_{\text{см}})}{W_{\text{см}}} = \pm(0,025 + 0,025 + 0,055 + 0,055)\% = \pm 0,16\%.$$

Оценивая эксперимент в целом, можно заранее подобрать аппаратуру, обеспечивающую заданную точность опыта.

## 2.4. Рабочая гипотеза

Определив цель и выбрав предмет исследования, составляют рабочую гипотезу исследования. *Рабочая гипотеза* – это научное предположение о развитии явления. Это предположение еще не доказано, но в некоторой степени вероятно.

Рабочая гипотеза устанавливает факторы, т. е. измеримые переменные величины, воздействующие на объект исследования. При этом рекомендуется вначале записать возможно более полный перечень факторов, которые предположительно могут воздействовать на развитие изучаемого явления. Затем установить наиболее действенные из них и отбросить все остальные (т. е. осуществить ранжирование факторов). Одновременно установить возможные границы изменения факторов. При неправильном определении факторов исследование может быть безрезультатным.

Рабочая гипотеза может давать более или менее полное предположительное объяснение всего процесса. При этом в гипотезе выделяют из всех уже известных наиболее важные и решающие причинные связи и взаимодействия и намечают возможные связи, взаимодействия, вероятное направление и ход развития процесса.

Если анализ литературных источников и пассивные наблюдения за процессом не позволяют установить факторы, влияющие на процесс, то проводят поисковые опыты.

Для этого составляют частную методику и назначают число поисковых опытов на основании следующих предположений:

1. Если проверяют направление развития процесса, то достаточно двух опытов – в начале и в конце процесса.

Например, для определения направления изменения тягового сопротивления плуга от глубины вспашки измеряют его значения при минимальной и максимальной глубине пахоты.

2. Если устанавливают, какие факторы обуславливают развитие явления, то наименьшее число поисковых опытов равно удвоенному предполагаемому числу факторов. Для того чтобы выяснить, влияет ли данный фактор на развитие явления, необходимо поставить как минимум два опыта – в начале (когда данный фактор явно отсутствует) и в конце (когда данный фактор явно присутствует).

Например, для определения влияния затупления лемеха на изменение средней глубины вспашки проводят два опыта: один с новым острым лемехом, другой – с тупым.

3. При проверке вариантов рабочей гипотезы выбирают главный фактор и по небольшой серии из 3–5 опытов определяют ход изменения основных закономерностей.

Например, при определении зависимости тягового сопротивления плуга от скорости движения агрегата ставят три опыта: при минимальной, максимальной и промежуточной (обычно средней) скорости. При этом устанавливают характер изменения тягового сопротивления, по виду которого можно обосновать количество измерений и шаг изменения скорости агрегата.

## **2.5. Планирование эксперимента**

Современный уровень развития теории эксперимента позволяет применять специальные методы для обработки и анализа результатов, а также выбора рациональной схемы проведения эксперимента. Общая схема эксперимента может быть представлена в виде определенной последовательности: постановка цели эксперимента; выбор критерия эффективности (зависимой переменной или отклика); отбор значимых и выбор варьируемых факторов (независимых переменных); выбор типа эксперимента; установление необходимого числа наблюдений; выбор измерительной аппаратуры; практическая реализация эксперимента; обработка результатов и построение математической модели: оценка



адекватности модели; интерпретация результатов экспериментов. Некоторые из этих вопросов рассматривались в предыдущих разделах. Остановимся кратко на вопросах планирования эксперимента.

При планировании и реализации эксперимента необходимо учитывать ряд важных требований. Первое требование заключается в *воспроизводимости* полученных результатов. Это означает, что при повторении одного и того же эксперимента через различные интервалы времени или в других местах должны получаться одинаковые или близкие в статистическом смысле результаты. Для обеспечения этого необходимо учитывать достаточное число факторов, а также исключить влияние систематически действующих, изменяющихся во времени и пространстве неуправляемых факторов.

Чтобы построенная по результатам экспериментов модель была *адекватной*, она должна представлять опытные данные с ошибкой, не превышающей ошибки опыта.

Третье требование к эксперименту касается его *управляемости*. Эксперимент необходимо строить так, чтобы обеспечивалось требуемое активное вмешательство экспериментатора в изменение числа и уровней управляемых факторов. Управлять факторами необходимо, чтобы обеспечить варьирование и смешивание факторов для выявления как основных эффектов, так и эффектов их взаимодействия.

### 2.5.1. Количество опытов. Выбор интервалов

Предполагаемый характер функциональных зависимостей фактически и определяет количество опытов (рис. 2.3).

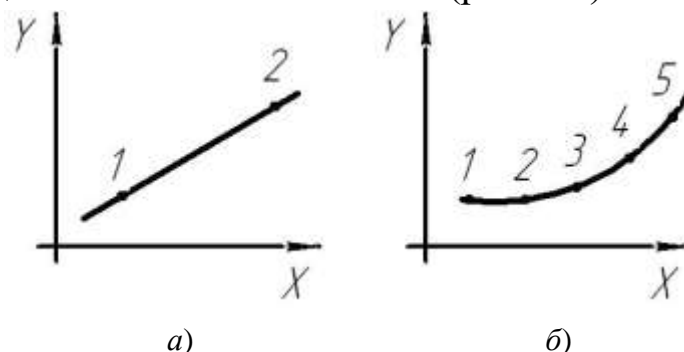


Рис. 2.3. Схема для выбора количества опытов

Положение прямой определено двумя точками (рис. 2.3, *а*). Если исследователь уверен в том, что данная зависимость прямая, он назна-

чает два опыта с промежутком между ними, близким к пределам изменения варьируемого фактора, чтобы увеличить точность положения прямой. Дуга вполне определяется тремя точками, которые располагают так: две близко к пределам и одну в середине. Если кривые более сложные, то применяют правило: рассматривая сложную кривую как комбинацию прямых и простых кривых, описать каждый перегиб кривой тремя опытами; каждый участок, близкий к прямолинейному, двумя опытами, и близко к назначенным пределам поставить два концевых опыта. Если требуется установить не только общие закономерности, но и возможно более точно численные значения функций, каждый перегиб кривой должен быть обоснован минимум пятью опытами.

Особенно тщательно должны обосновываться резкие перегибы кривых или их скачкообразные изменения, так как эти перегибы и скачки могут характеризовать переход количественных изменений в качественные.

Если кривая представляет собой некоторую степенную зависимость (рис. 2.3, б), то необходимые для ее описания пять точек должны быть распределены так, чтобы большая их часть располагалась по круто восходящей ветви. В общем случае исходят из того, чтобы обеспечить в любой части кривой равную точность. Рекомендуется выбирать интервал варьирования фактора в 3–5 раз большим ошибки его измерения.

*Необходимое число повторных опытов.* При проведении экспериментальных исследований всегда имеют место случайные ошибки опыта. Для уменьшения значения таких ошибок целесообразно повторять опыты. Из теории ошибок известно, что количество повторных опытов зависит от стандарта измерений и заданной доверительной вероятности.

Под *доверительной вероятностью* подразумевается вероятность получения одинаковых результатов при новых измерениях величины или вероятность получения тех же результатов при повторении опытов в аналогичных условиях.

Чем больше относительные колебания результатов и чем больше надежность опытов желательно получить, тем больше должно быть повторных опытов.

В литературе по методам проведения экспериментальных исследований приводятся специальные таблицы зависимости между необходимым числом повторных опытов, заданной надежностью и относительной ошибкой (табл. 2.1).

Для того чтобы найти по этой таблице необходимое количество повторных опытов, необходимо задаться доверительной вероятностью  $\alpha$  и ошибкой  $\Delta_n$ , взятой в долях стандарта  $\sigma$ . Чтобы добиться желаемой достоверности, почти всегда выгоднее увеличить точность прибора, а не количество измерений.

Таблица 2.1

Необходимое количество повторных опытов  
при доверительной вероятности  $\alpha$

Ошибка $\Delta_n$ в долях $\sigma$	Доверительная вероятность $\alpha$							
	0,5	0,6	0,7	0,8	0,9	0,95	0,99	0,999
3	1	1	1	1	2	3	4	5
2	1	1	1	2	3	4	5	7
1	2	2	3	4	5	7	11	17
0,5	3	4	6	9	13	18	31	50
0,4	4	6	8	12	19	27	46	74
0,3	6	9	13	20	32	46	78	127
0,2	13	19	29	43	70	99	171	277
0,1	47	72	169	266	273	387	668	1089

Практикой установлено, что при исследовании закономерностей в самом общем виде (например, определение характера кривых) достаточна доверительная вероятность  $\alpha = 0,7$ . Для измерений, связанных с конструкциями машин, достаточна  $\alpha = 0,9$ . При определении значений величин, являющихся основой для дальнейшего расчета, необходима  $\alpha = 0,99$ .

Если поисковые опыты не проводились и нет данных, по которым можно было бы установить значение среднего квадратического

отклонения, то можно принять, что предельная ошибка приближенно равна наибольшей возможной статистической, т. е.

$$\Delta_{\pi} = \pm 3\sigma.$$

Например, пусть необходимо измерить анемометром скорость воздушного потока над решетом сепаратора зернового вороха комбайна с точностью  $\Delta_{\pi} = \pm 3\sigma \pm 0,2$  м/с. Систематическая ошибка анемометра неизвестна. Тогда одно измерение анемометра будет иметь доверительную вероятность, равную 0,8 (пересечение данной строки  $\Delta_{\pi} = 3\sigma$  с соответствующими столбцами), два измерения – 0,9, три – 0,95. Следовательно, если требуется, чтобы в 95 случаях из 100 ошибка измерения не превышала 0,2 м/с, нужно сделать три замера и найти их среднее значение.

Еще пример. Исследуется зависимость частоты вращения молотильного барабана от величины подачи хлебной массы. Запись частоты вращения барабана ведется от электротаксметра на бумажную ленту. Значения частоты вращения барабана были сняты с участка ленты, соответствующего времени работы барабана 5 с. На ленте оказалось возможным произвести 30 замеров значений частоты вращения в соответствии с ранее рассмотренным правилом получения выборки с осциллограмм. Среднее арифметическое значение частоты вращения оказалось  $n = 1000$  мин<sup>-1</sup>, стандарт  $\sigma = 50$  мин<sup>-1</sup>. Какой должна быть продолжительность записи, чтобы ошибка средней частоты вращения не превышала 1% при доверительной вероятности  $\alpha = 0,99$ ?

Прежде всего нужно определить рассеяние точек, вызываемое ошибкой тахометра. При калибровке тахометра на частоте 1000 мин<sup>-1</sup> получен стандарт  $\sigma = 9,1$  мин<sup>-1</sup>. Тогда искомое рассеяние будет равно

$$\sigma_o = (\sigma - \sigma_{\tau})^{0,5} = (50 - 9,1)^{0,5} = 49 \text{ мин}^{-1}.$$

Допускаемая абсолютная ошибка  $\Delta$  средней частоты вращения равна  $1000 \cdot 0,01 = 10$  мин<sup>-1</sup>.

Вычисляем эту ошибку в долях стандарта:

$$\Delta = \frac{10}{49\sigma} = 0,204\sigma \approx 0,2\sigma.$$

Для  $\alpha = 0,99$  и  $\Delta = 0,2\sigma$  находим необходимое количество замеров на осциллограмме, равное 171 (см. табл. 2.1). Принимая, что за 1 с на ленте имеется  $30/5 = 6$  замеров, для получения 171 замера необходима запись в течение  $171/6 \approx 29$  с. Следовательно, для обеспечения ошибки измерения в 1% с достоверностью  $\alpha = 0,99$  необходимо вести запись не менее 29 с.

Зная стандарт, можно установить необходимое количество повторных опытов. Для определения стандарта необходимо, как видно из приведенных примеров, иметь данные хотя бы предварительных экспериментов. Однако в большинстве случаев трудоемкость и стоимость предварительных опытов сравнимы с основными. Тогда для определения числа повторных опытов используют метод последовательного анализа.

*Последовательный анализ* позволяет установить необходимое количество повторных измерений при заданных надежности и ошибке в долях стандарта, если сам стандарт не известен. Единственное обязательное условие при этом – устанавливая по каждому измерению его результаты перед последующим измерением. Такой анализ применим, если для измерения использовались инструменты и приборы, быстро дающие результаты. Такой анализ не применим при записях данных, требующих расшифровки.

Сущность последовательного анализа состоит в следующем. После каждого измерения (начиная с третьего) вычисляют среднее и стандарт. При достижении удовлетворительного результата (надежность и вычисленная ошибка в долях стандарта должны давать по табл. 2.1 такое количество измерений, которое уже произведено) измерения следует прекратить.

Например, требуется измерить скорость воздушного потока над верхним решетом очистки комбайна в ее средней части с ошибкой 1% при доверительной вероятности 0,9. Так как сведений о потоке воздуха нет, установить предварительно количество повторных опытов невозможно. Пользуясь методом последовательного анализа, производим первые три измерения и получаем значения скорости: 7,4; 7,6 и 7,5 м/с.

Средняя скорость по этим измерениям будет  $V \approx 7,5$  м/с. Стандарт  $\sigma = 0,1$  м/с, 1%-я ошибка от средней  $7,5/100 = 0,075$  м/с, относительная ошибка составляет

$$\Delta = \left( \frac{0,075}{0,1} \right) \sigma = 0,75\sigma.$$

Устанавливаем, что при полученной относительной ошибке и трех измерениях обеспечивается достоверная вероятность результата измерений менее 0,7.

Делаем четвертое измерение, которое дает результат 7,55 м/с. При этом получаем  $V = 7,51$  м/с,  $\sigma = 0,0856$  м/с, 1%-я ошибка  $7,51/100 = 0,0751$  м/с, т. е. относительная ошибка составляет  $0,0751/0,0856 = 0,877\sigma$ . По табл. 2.1 находим, что доверительная вероятность – 0,82, что также недостаточно.

Производим пятый замер, который дает значение 7,45 м/с. Находим  $V=7,5$  м/с;  $\sigma= 0,079$  м/с; 1%-я ошибка 0,075 м/с, что составляет  $0,075/0,079 = 0,95\sigma$ . При пяти замерах и такой ошибке обеспечивается доверительная вероятность 0,92, что выше заданной точности, поэтому замеры прекращаем.

### 2.5.2. Порядок проведения эксперимента

*Однофакторные эксперименты.* Если исследуется влияние одного фактора на изучаемый объект (одна независимая переменная), то такой эксперимент называется однофакторным.

*Типы планов эксперимента.* После выбора (из соображений обеспечения требуемой точности) интервалов между точками необходимо выбрать порядок проведения эксперимента.

Существует два вида планов проведения эксперимента.

1. Берут вначале верхнее или нижнее предельное значение независимой переменной величины (так называемые уровни) и изменяют его ступенчато до тех пор, пока не будет достигнуто другое предельное значение. Такой план называется *последовательным*. Например, исследуется зависимость потерь зерна  $d$  за соломотрясом в молотилке зерноуборочного комбайна от величины подачи хлебной массы, т. е.

$d = f(q)$ . Были установлены пределы изменения независимой переменной:  $q_{\min} = 1$  кг/с, :  $q_{\max} = 5$  кг/с хлебной массы.

При последовательном плане устанавливают последовательно подачи  $q = 1, 2, 3, 4, 5$  кг/с и определяют соответствующие этим подачам потери  $d$ , проводя для каждого значения  $q$  столько повторных опытов, сколько необходимо для получения заданной достоверности результатов.

2. Однако в процессе эксперимента (особенно, если он длится несколько дней) различные неучитываемые факторы могут изменяться и оказывать существенное влияние на исследуемый параметр. Так, в рассматриваемом примере в ходе эксперимента могут изменяться свойства хлебной массы, свойства поверхности клавиш соломотряса из-за налипания полеры, показания измерительной аппаратуры и др. В результате этого каждый последующий переход к более высокой подаче будет сопровождаться ошибкой (результат будет систематически только завышаться или только занижаться), а общий результат эксперимента будет иметь систематическую ошибку, которую трудно обнаружить. Эту систематичность ошибки можно предотвратить, если выбор точек (уровней) производить случайным образом. При этом полученные результаты замеров для каждой точки будут также иметь разброс, но они будут группироваться вокруг точных значений. Такой план эксперимента называется случайным или *рандомизированным*.

Для нашего примера подача  $q$  должна чередоваться следующим случайным образом: 3; 5; 1; 4; 2 кг/с.

Порядок чередования точек устанавливают с помощью таблиц случайных чисел или другими известными способами.

В связи с указанным недостатком последовательных планов их применение целесообразно лишь в случаях:

- когда продолжительность, стоимость или сложность эксперимента такова, что рандомизация становится нецелесообразной. Например, при исследованиях атомного реактора, для которого тепловое равновесие достигается в течение нескольких суток;

- если известно, что эксперимент является невоспроизводимым. Например, при исследовании образцов материалов на прочность.

*Рандомизированные блоки.* При проведении исследований имеют

место случаи, когда на исследуемый процесс помимо исследуемого фактора оказывают влияние и ряд других внешних факторов. Например, определяется сменная выработка  $S$  комбайна в зависимости от ширины захвата  $B$  жатки (урожайность хлебов постоянна). При этом предполагается использовать жатки различной ширины захвата: № 1 – 4.

Производительность комбайна зависит не только от ширины жатки  $B$ , но и от квалификации оператора, его темперамента и т. п., т. е. имеет место некоторый внешний фактор – «механизатор». Поэтому для рандомизации эксперимента по данному фактору необходимо случайным образом выбрать по числу жаток четырех рабочих А, В, С и D, каждый из которых будет работать полную смену на комбайне с жаткой данного захвата. Эксперимент должен длиться четыре дня (например: понедельник, вторник, среда и четверг).

К концу срока испытаний рабочие могут устать, потерять интерес к работе, что приведет к снижению производительности. Или же, наоборот, оператор, получив больший практический опыт, будет работать производительнее. Так, если все механизаторы будут последовательно испытывать комбайн с жатками № 1 – 4, то к концу срока производительность с жатками, например, № 2 и 3 может возрасти или уменьшиться по указанным причинам.

Поэтому необходимо рандомизировать эксперимент по последовательности испытания номеров жаток. Но при этом может оказаться, что некоторым рабочим выпало испытывать одинаковые номера жаток в один и тот же день, и на производительность окажут влияние вышеуказанные причины (кроме этого потребуется дополнительное количество жаток).

Поэтому необходимо рандомизировать эксперимент таким образом, чтобы в данный день каждый номер жатки использовался только один раз. Такой полностью рандомизированный план эксперимента приведен в табл. 2.2.

Таблица 2.2

#### План проведения эксперимента



Механизатор	Дни недели			
	Пн.	Вт.	Ср.	Чт.
A	1	2	3	4
B	3	4	1	2
C	2	1	4	3
D	4	3	2	1

В табл. 2.2 приведен так называемый *латинский квадрат* (4x4). После завершения испытаний результаты по каждому номеру жаток усредняются.

Рассмотренный эксперимент можно еще усовершенствовать. Если каждый механизатор работал на одной и той же машине и только менял каждый день жатки, то вследствие различий между молотилками может появиться систематическая ошибка. Если обозначить молотилки через № M1, M2, M3, M4, то эксперимент можно спланировать так, чтобы каждый механизатор работал на каждой молотилке только по одному дню (табл. 2.3).

Таблица 2.3

#### План проведения эксперимента

Механизатор	Дни недели			
	Пн.	Вт.	Ср.	Чт.
A	1M1	2M2	3M4	4M3
B	3M2	4M1	1M3	2M4
C	2M3	1M4	4M2	3M1
D	4M4	3M3	2M1	1M2

В табл. 2.3 приведен так называемый *греко-латинский квадрат* (4x4).

*Многофакторные эксперименты.* При воздействии на изучаемый объект нескольких факторов влияние на исследуемый параметр оказывает не только каждый фактор в отдельности, но и различные комбинации этих факторов, взаимодействующих между собой.

Эксперименты, проводимые для исследования влияния нескольких факторов, называются *многофакторными*.

В многофакторных экспериментах возможно использование плана одного из двух типов: *классического* или *факторного*.

Классический план применяется во всех областях исследований. Факторный план часто бывает короче, он всегда точнее (при одинаковом с классическим числе опытов), но, по ряду причин, применяется реже.

*Классический план.* Пусть некоторая зависимая переменная  $R$  является функцией нескольких переменных  $X, Y, Z$  и т. д. Классический план эксперимента состоит в том, что все независимые переменные, кроме одной, полагают постоянными, и эта одна переменная изменяется на всем интервале ее значений. Затем изменяют другую переменную, полагая все остальные постоянными, т. е. последовательно находят зависимости  $R = f(X), R = f(Y), R = f(Z)$ . По существу классический многофакторный эксперимент представляет собой последовательность однофакторных экспериментов.

Результат многофакторного эксперимента трудно представить графически, если число применяемых факторов больше двух. Поэтому конечная цель эксперимента — эмпирическая зависимость  $R = f(X, Y, Z, \dots)$ .

В зависимости от сложности функции  $R$  план эксперимента может иметь различную степень сложности. Например, план двухфакторного эксперимента, в котором каждый фактор берется на пяти уровнях, схематически можно представить в следующем виде (рис. 2.4):

		Уровни переменной $Y$							Уровни переменной $Y$				
		1	2	3	4	5			1	2	3	4	5
Уровни переменной $X$	1			+			Уровни переменной $X$	1	+	+	+	+	+
	2			+				2	+		+		+
	3	+	+	+	+	+		3	+	+	+	+	+
	4			+				4	+		+		+
	5			+				5	+	+	+	+	+
а)							б)						

Рис. 2.4. Планы двухфакторного эксперимента

Знаком «+» обозначены комбинации условий, при которых должен производиться эксперимент.

Если заранее известно, что исследуемые зависимости могут быть описаны такими простыми функциями:  $R = AY^n + BX^m$ ,  $R = AY^nX^m$ ,  $R = AYB^{nx}$  и т. д., то можно ограничиться планом, когда обе переменные  $X$  и  $Y$  поочередно берутся на одном уровне (рис. 2.4, а).

В случаях использования более сложных функций, например таких как  $R = AX \sin\left(\frac{BY}{X}\right)$ ,  $R = AX^{by}$  и др., необходимо рассмотреть несколько уровней переменных  $X$  и  $Y$  (см. рис. 2.4, б). Может появиться необходимость заполнить весь квадрат и пронести эксперимент для всех 25 комбинаций условий.

План классического эксперимента необязательно должен быть сбалансированным. Это означает, что можно выбрать десять уровней переменной  $X$  и только три уровня переменной  $Y$ , если известно заранее, что зависимость  $R$  от  $X$  является более важной или более сложной.

После проведения экспериментов строят графические или аналитические зависимости  $R = f(X)$ ,  $R = f(Y)$ .

## 2.6. Предварительный анализ экспериментальных данных

Предварительный анализ экспериментальных данных включает определение функциональных связей, графический анализ данных, сглаживание графиков и другие операции.

*Определение функциональных связей. Графический анализ данных.* Академик В.П. Горячкин говорил, что «закономерность ряда чисел воспринимается нашим сознанием только в грубых, общих чертах. Тот же ряд чисел, представленный в виде графика, сразу открывает развитие явления».

Общие правила построения графиков таковы. По оси абсцисс откладывают значение аргумента, а по оси ординат – значение функции.

Начало отсчета необязательно располагать в пределах листа. Желательно, чтобы масштаб графика соответствовал абсолютной ошибке измерений, тогда ошибка отсчета по графику не будет превышать ошибки измерений. Рекомендуются масштабы шкал:  $1:1 \cdot 10^n$ ;  $1:2 \cdot 10^n$ ;  $1:5 \cdot 10^n$ , при этом  $n = 0, 1, 2, 3, \dots$ .

Если отрезок длиной  $m$ , изображающий единицу величины  $U$ , назвать модулем шкалы, и если известны пределы  $a$  и  $b$  изменения ве-

личины  $U$ , то длина шкалы будет определяться зависимостью:  $L = m(b - a)$ . Расчетное значение длины шкалы округляют в сторону увеличения. Только после вычерчивания шкал опытные данные наносят на график.

*Сглаживание графиков.* При построении графиков необходимо установить, не являются ли изломы линий следствием естественных закономерностей, связанных с переходом количественных изменений в качественные, и, если потребуется, повторить все опыты, приняв меры, исключающие влияние ошибок наблюдений.

Физическим основанием сглаживания кривых является плавность изменения функции при плавном изменении аргумента; выравненные кривые должны наиболее близко отображать общую закономерность развития явления. Это означает, что они не обязательно должны быть средними и что усреднение лишь один из приемов сглаживания и выравнивания.

В первую очередь сглаживают данные непосредственных измерений. Нельзя сглаживать результаты косвенных измерений без предварительного выравнивания данных непосредственных измерений, так как могут появиться большие ошибки. Например, при испытании двигателей измеряют крутящий момент и частоту вращения вала, а затем вычисляют мощность и определяют зависимость  $N = f(n)$ . Если предварительно не сгладить результаты этих измерений, а лишь выровнять кривую вычисленной мощности, то все данные характеристики двигателя могут оказаться несопоставимыми.

При сглаживании устанавливают, нет ли отдельных резких односторонних отклонений от плавной кривой. Если такие односторонние отклонения (одно-два на всю серию опытов) имеются и находят объяснение в изменении условий измерения, их опускают, а вместо них при помощи *интерполяции* находят точки, близко расположенные к естественной плавной кривой развития.

Далее надо выяснить, насколько разбросаны точки опытов в обе стороны от воображаемой плавной кривой, соединяющей их, и необходимо ли сглаживание. Если разброс точек таков, что соединить их плавной кривой невозможно, то кривую нужно сгладить (рис. 2.5).

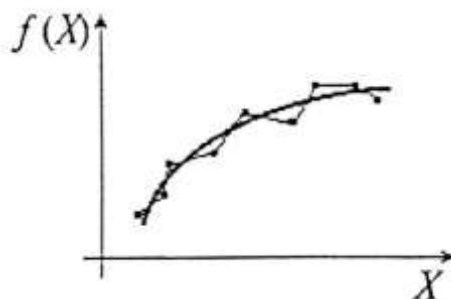


Рис. 2.5. Сглаживание кривой

Наиболее точные результаты получают при использовании для сглаживания методов наименьших квадратов и наибольшего правдоподобия. Для получения достаточно точного характера развития исследуемого процесса применяют так называемое графическое сглаживание. Для этого используют карандаш и прозрачные гибкие шаблоны (лекала, угольники, линейки). Основное правило графического сглаживания: проведенная кривая должна располагаться как можно ближе ко всем опытным точкам.

Для этого должны быть выполнены следующие условия:

- сумма отрезков нормалей, опущенных из опытных точек на кривую (см. рис. 2.5), должна равняться нулю (отрезок нормали над кривой берется со знаком *плюс*, а под кривой – со знаком *минус*);
- если соединить все опытные точки прямыми и провести плавную кривую, то сумма площадей, отсеченных кривой по обе ее стороны, должна быть равна нулю.
- в обоих случаях также необходимо, чтобы сумма абсолютных величин отрезков нормалей или площадей была минимальной.

Вначале тонкой линией проводят первую кривую. Если она не удовлетворяет проверочным требованиям, ее не стирают, а наносят вторую линию. Если понадобится, вычеркивают третью. Затем толстой линией обводят лучший вариант, а остальные стирают.

### 3. ПОСТРОЕНИЕ МОДЕЛЕЙ ПО РЕЗУЛЬТАТАМ КОРРЕЛЯЦИОННО-РЕГРЕССИОННОГО АНАЛИЗА

Одним из эффективных методов установления взаимосвязей является корреляционно-регрессионный анализ.

Широкое использование линейных связей и связей, приводящихся к линейным, объясняется несколькими причинами: во-первых, линейные связи просты и требуют относительно меньшего объема вычислений, а методика их установления более глубоко разработана; во-вторых, существуют теоретические предпосылки более частого использования линейных форм связи, которым свойственно нормальное распределение, а оно наиболее часто встречается на практике; в-третьих, корреляционную зависимость часто заменяют линейной формой связи потому, что при сравнительно малых диапазонах изменения входного параметра любую кривую в первом приближении можно аппроксимировать кусочно-линейной связью.

Корреляционный и регрессионный анализы позволяют решать различные, но взаимосвязанные между собой задачи: коэффициент корреляции оценивает силу связи, уравнение регрессии – ее форму, с привлечением оценки достоверности определяется реальность существования связи.

Сущность корреляционно-регрессионного метода заключается в нахождении эмпирического уравнения (формулы), характеризующего связь результативного параметра с определенным входным фактором (или несколькими факторами в случае множественной корреляции).

Для обработки исходной информации, установления эмпирической формулы связи можно принять такую схему:

- выбор результативного  $Y$  и входного  $X$  параметров;
- вычисление средних арифметических значений и дисперсий результативного и входного параметров;
- выбор формы связи;
- оценка силы связи, расчет коэффициента корреляции;
- оценка значимости коэффициента корреляции;

- расчет теоретического уравнения регрессии;
- оценка значимости коэффициентов уравнения регрессии;
- определение адекватности уравнения регрессии и доверительных границ.

Средние арифметические значения входного и выходного параметров вычисляются соответственно по формулам:

$$\bar{Y} = \frac{\sum_{i=1}^n Y_i}{n}, \quad \bar{X} = \frac{\sum_{i=1}^n X_i}{n},$$

где  $X_i$  и  $Y_i$  – значение входного и выходного параметров соответственно;  $n$  – число пар исходных данных в массиве.

Дисперсии рассматриваемых факторов определяются по формулам:

$$S_X^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n - 1}; \quad S_Y^2 = \frac{\sum_{i=1}^n (Y_i - \bar{Y})^2}{n - 1}.$$

### 3.1. Выбор формы связи

Начальный этап выбора формы исследуемой связи заключается в аналитическом или графическом анализе исходной информации (например, построение поля корреляции). Если значения выходного параметра изменяются приблизительно равномерно с изменением входного фактора, то, возможно, существует линейная связь, если неравномерно – нелинейная.

При решении многих практических задач чаще используются линейные связи и связи, которые могут быть преобразованы в линейный вид.

*Гиперболическая связь* вида  $Y = a_0 + a_1/X$  линеаризуется заменой входного параметра переменной:

$$X' = 1/X,$$

тогда

$$Y = a_0 + aX'.$$

*Показательная связь* вида  $Y = a_0 \cdot e^{a_1 X}$  линеаризуется заменой выходного фактора переменной:

$$Y = e^{Y'},$$

после логарифмирования

$$Y' = \ln a_0 + a_1 X = a_0' + a_1 X .$$

*Степенная связь* вида  $Y = a_0 x^{a_1}$  после логарифмирования линеаризуется заменой:

$$\ln Y = Y' , \quad \ln a_0 = a_0' , \quad \ln X = X' ,$$

тогда

$$Y' = a_0' + a_1 X' .$$

*Логарифмическая связь* вида  $Y = a_0 + a_1 \ln X$  линеаризуется заменой

$$X' = \ln X ,$$

тогда  $Y = a_0 + a_1 X' .$

*Комбинированная связь* вида  $Y = 1/(a_0 + a_1 e^{-x})$  линеаризуется заменой выходного фактора переменной

$$Y' = 1/Y ,$$

а входного фактора – переменной

$$X' = e^{-x} ,$$

тогда  $Y' = a_0 + a_1 X' .$

### 3.2. Оценка силы связи между факторами

В общем случае силу связи между исследуемыми факторами можно оценить, используя корреляционное отношение.

Для проверки значимости предполагаемой линейной формы связи (или приведенных к ней) вычисляется коэффициент корреляции  $r$ . Коэффициент корреляции может принимать положительные и отрицательные значения.

Если связь между выходным параметром и фактором-аргументом может быть точно представлена прямой линией, коэффициент корреляции  $r = \pm 1$ .

При  $r = 1$  – связь между  $Y$  и  $X$  прямо пропорциональная, при  $r = -1$  – обратно пропорциональная.

Если значение коэффициента корреляции  $r = 0$ , говорят о некоррелированности исследуемых факторов. Для нормального закона



распределения случайных величин при  $r = 0$  можно говорить о независимости этих величин. Для других распределений зависимость может иметь место.

Принято считать:

при  $r = 0,3$  – слабая связь;

$r = 0,3 - 0,7$  – средняя связь;

$r = 0,7$  – сильная;

$r = 0,9$  – весьма сильная связь.

Коэффициент корреляции определяется по формуле

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{n S_X S_Y},$$

где  $n$  – число пар исследуемых факторов;  $S_X$ ,  $S_Y$  – среднеквадратическое отклонение входного и выходного факторов соответственно.

Для удобства расчета коэффициента корреляции можно использовать формулу

$$r = \frac{n \sum_{i=1}^n X_i Y_i - \sum_{i=1}^n X_i \sum_{i=1}^n Y_i}{\sqrt{n \sum_{i=1}^n X_i^2 - (\sum_{i=1}^n X_i)^2} \sqrt{n \sum_{i=1}^n Y_i^2 - (\sum_{i=1}^n Y_i)^2}}.$$

*Проверка значимости коэффициента корреляции.* Коэффициент корреляции  $r$ , определяемый по выборочным данным, может не совпадать с действительным значением, соответствующим генеральной совокупности (в силу малой представительности выборки). Ошибка выборочного коэффициента корреляции определяется по формуле

$$S_r = \frac{1 - r^2}{\sqrt{n - 1}}.$$

При больших выборках ( $n = 100$ ) можно предположить, что коэффициент корреляции распределен нормально. Тогда справедливо следующее выражение:

$$P\{r - X_r S_r \leq r \leq r + X_r S_r\},$$

где  $P$  – вероятность.

При  $P = 0,9$ ,  $X_r = 1,653$ ;  $P = 0,95$ ,  $X_r = 1,96$ ;  $P = 0,99$ ,  $X_r = 2,576$ .

При малых  $n$  гипотеза о нормальном распределении коэффициента корреляции, как правило, не подтверждается. В этом случае используют  $t$ -критерий Стьюдента:

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}.$$

Для числа степеней свободы  $n = n - 2$  и уровня значимости по таблицам находят теоретическое значение  $t$  – критерия. Если фактическое значение критерия Стьюдента будет больше теоретического значения, то предположение о нулевом значении коэффициента корреляции не подтверждается. В противном случае считается, что величина  $r$  близка к нулю.

### 3.3. Расчет коэффициентов уравнения регрессии

В основе построения эмпирической зависимости (уравнения регрессии) лежит метод наименьших квадратов. Он является одним из наиболее употребительных методов приближенной оценки неизвестных параметров уравнения. Проанализировав форму связи между исследуемыми переменными, искомое эмпирическое уравнение можно привести к виду

$$Y = a_0 + a_1 X.$$

Затем из полученных экспериментальных данных устанавливается приближенное равенство:

$$Y_i \cong a_0 + a_1 X_i, \quad (i = 1, 2, 3, \dots, n).$$

Коэффициенты  $a_0$  и  $a_1$  определяются из условия, чтобы сумма квадратов разности между левой и правой частями приближенных равенств была минимальной:

$$\sum_{i=1}^n [Y_i - (a_0 + a_1 X_i)]^2 = \min.$$

Для нахождения экстремума находим частные производные и приравниваем к нулю:

$$\frac{\partial}{\partial a_0} \left\{ \sum_{i=1}^n [Y_i - (a_0 + a_1 X_i)]^2 \right\} = 0;$$

$$\frac{\partial}{\partial a_1} \left\{ \sum_{i=1}^n [Y_i - (a_0 + a_1 X_i)]^2 \right\} = 0.$$

Решив полученную систему уравнений, определяют неизвестные коэффициенты модели:

$$\begin{cases} a_0 n + a_1 \sum_{i=1}^n X_i = \sum_{i=1}^n Y_i ; \\ a_0 \sum_{i=1}^n X_i + a_1 \sum_{i=1}^n X_i^2 = \sum_{i=1}^n X_i Y_i . \end{cases}$$

*Значимость коэффициентов уравнений регрессии.* Для оценки параметров, входящих в уравнение регрессии, при решении практических задач можно ограничиться построением доверительных интервалов.

Задаваясь уровнем значимости  $\alpha$  (или доверительной вероятностью  $1 - \alpha$ ) при известном числе степеней свободы, доверительные границы для параметров  $a_0$  и  $a_1$  определяются по формулам:

$$\begin{aligned} a_{01} &= a_0 - t_T \cdot C_{0a_0}; & a_{11} &= a_1 - t_T \cdot C_{0a_1}; \\ a_{02} &= a_0 + t_T \cdot C_{0a_0}; & a_{12} &= a_1 + t_T \cdot C_{0a_1}, \end{aligned}$$

где  $t_T$  – теоретическое значение  $t$  – критерия Стьюдента;  $C_{0a_0}$  – случайная ошибка параметра  $a_0$ ;  $C_{0a_1}$  – случайная ошибка параметра  $a_1$ .

Случайная ошибка параметра  $a_0$  определяется по формуле

$$C_{0a_0} = S_{\text{ост}} \sqrt{\frac{\sum_{i=1}^n X_i^2}{n \left[ \sum_{i=1}^n X_i^2 - \frac{(\sum_{i=1}^n X_i)^2}{n} \right]}},$$

где  $S_{\text{ост}}$  – остаточная дисперсия.

Остаточная дисперсия  $S_{\text{ост}}^2$  определяется так:

$$S_{\text{ост}} = \frac{1}{n-2} \sum_{i=1}^n [Y_i - (a_0 + a_1 X_i)]^2.$$

Случайная ошибка параметра  $a_1$  определяется по формуле

$$C_{0a_1} = \frac{S_{\text{ост}}}{\sqrt{\sum_{i=1}^n X_i^2 - \frac{(\sum_{i=1}^n X_i)^2}{n}}}.$$

Если фактическое значение  $t$  – критерия Стьюдента

$$t_{a_i} = \sqrt{\frac{(n-2) \sum_{i=1}^n (X_i - \bar{X})^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2}}$$

будет меньше или равно теоретическому, то параметры  $a_0$  и  $a_1$  незначимы.

*Доверительные границы уравнения регрессии.* Для определения доверительных границ уравнения регрессии необходимо для каждого значения аргумента определить ширину доверительного интервала функции.

Наименьшее значение ширины доверительного интервала теоретического уравнения регрессии определяется:

$$\Delta X_{min} = \pm \frac{t_T S_{ост}}{\sqrt{n}}.$$

Доверительные границы уравнения регрессии представляют собой интервалы, определяемые из выражения

$$Y \pm t_{a, 2, n-2} S_{ост} \sqrt{\frac{1}{n} + \frac{(X_0 - \bar{X})^2}{\sum_{i=1}^n (X_i - \bar{X})^2}}.$$

*Линейная зависимость  $Y = b_0 + b_1 X$ .* Для этой функции коэффициенты  $b_1$  и  $b_0$  определяются по формулам:

$$b_1 = \frac{\sum_{i=1}^n X_i \sum_{i=1}^n Y_i - n \sum_{i=1}^n X_i Y_i}{(\sum_{i=1}^n X_i)^2 - n \sum_{i=1}^n X_i^2}; \quad b_0 = \frac{1}{n} \left( \sum_{i=1}^n Y_i - b_1 \sum_{i=1}^n X_i \right).$$

*Экспоненциальная зависимость  $Y = b_0 \cdot e^{b_1 X}$ .* Для этой функции коэффициенты  $b_1$  и  $b_0$  определяются по формулам:

$$b_1 = \frac{\sum_{i=1}^n X_i \sum_{i=1}^n \ln Y_i - n \sum_{i=1}^n X_i \ln Y_i}{(\sum_{i=1}^n X_i)^2 - n \sum_{i=1}^n X_i^2};$$

$$b_0 = \exp \left[ \frac{1}{n} \left( \sum_{i=1}^n \ln Y_i - b_1 \sum_{i=1}^n X_i \right) \right].$$

*Параболическая зависимость  $Y = b_0 + b_1 X + b_2 X^2$ .* Для этой функции коэффициенты  $b_0$ ,  $b_1$ ,  $b_2$  определяются при решении системы из трех уравнений:

$$\left\{ \begin{array}{l} nb_0 + b_1 \sum_{i=1}^n X_i + b_2 \sum_{i=1}^n X_i^2 = \sum_{i=1}^n Y_i ; \\ b_0 \sum_{i=1}^n X_i + b_1 \sum_{i=1}^n X_i^2 + b_2 \sum_{i=1}^n X_i^3 = \sum_{i=1}^n X_i Y_i ; \\ b_0 \sum_{i=1}^n X_i^2 + b_1 \sum_{i=1}^n X_i^3 + b_2 \sum_{i=1}^n X_i^4 = \sum_{i=1}^n X_i^2 Y_i . \end{array} \right.$$

### Задачи для упражнений

*Задача 1.* Потери зерна на соломотрясе составляют 50 - 80% от общих потерь за молотилкой. Для автоматического регулирования необходимо выбрать параметр, с учетом которого можно определить загрузку соломотряса. Одним из параметров является толщина хлебной массы в наклонной камере. Получены экспериментальные данные зависимости потерь зерна на соломотрясе от толщины слоя хлебной массы под плавающим транспортером (табл. 3.1).

Таблица 3.1

#### Экспериментальные данные потерь

Толщина слоя, см	0,8	1,0	1,2	1,3	1,4	1,6	1,8
Потери зерна на соломотрясе, %	0,4	0,5	0,7	0,8	1,2	1,6	3,2

- Оценить силу связи между указанными факторами.
- Установить зависимость между потерями зерна и толщиной хлебной массы под плавающим транспортером.

*Задача 2.* Одним из важных приемов повышения урожайности овощных культур является калибрование семян. Разделение партии семян на фракции с одинаковыми физико-механическими и биологическими свойствами возможно на основе данных о связи между этими свойствами. В результате исследований зависимости всхожести семян моркови от их толщины получены данные (табл. 3.2).

Таблица 3.2

## Экспериментальные данные всхожести

Толщина семян, мм	0,5	0,6	0,7	0,8	0,9	1	1,1	1,2	1,3
Всхожесть, %	40	60	70	77	72	60	57	48	20

- Оценить силу связи между указанными факторами.
- Установить зависимость между исследуемыми факторами.

*Задача 3.* Качество обмолота зерна зависит от различных параметров валка. В результате исследований получены экспериментальные данные суммарных потерь при обмолоте в зависимости от ширины валка, подбираемого комбайном на скорости  $V_1 = 1,1$  и  $V_2 = 2$  м/с (табл. 3.3). Ширина молотилки комбайна – 1,2 м, производительность – 4 кг/с, соотношение зерна к соломе – 1:1.

Таблица 3.3

## Зависимость суммарных потерь от ширины валка

Скорость комбайна, м/с		Ширина валка S, м
$V_1 = 1,1$	$V_2 = 2$	
Потери зерна, %		
4,4	2	0,8
2,8	1	0,9
2,1	0,2	1
1,6	0,4	1,1
2	1,2	1,2
4	2,8	1,3

- Установить зависимость суммарных потерь зерна от ширины валка при различных скоростях движения комбайна.

*Задача 4.* В результате сравнительных испытаний жаток получены данные о величине потерь зерна  $\Pi$  в зависимости от скорости движения  $V$  жатки (табл. 3.4).

Таблица 3.4

## Статистические данные потерь

V, км/ч	6	6,4	6,7	9,3	9,5	10	13	13,5	14
$\Pi$ , %	1	0,7	1,1	1	1,5	1,2	1,6	1,9	1,7

- Определить параметры уравнения связи между потерями зерна и скоростью движения жатки.

**Задача 5.** Установлено, что наивысшей урожайности культур достигают при оптимальной плотности почвы, имеющей различное значение для различных типов почв. При проходах сельскохозяйственной техники по полю в следах движителей почва уплотняется. Значения недобора урожая  $H_i$  различных культур в зависимости от плотности почвы  $P_{\Pi}$  в следах движителей приведены в табл. 3.5.

Таблица 3.5

Недобор урожая  $H_i$  в следах движителей тракторов

Ячмень							
$P_{\Pi}$ , г/см <sup>3</sup>	0,9	1,1	1,15	1,17	1,2	1,24	1,27
$H_i$ , %	0	3	11,5	15,2	27	34,2	37,8
Озимая пшеница							
$P_{\Pi}$ , г/см <sup>3</sup>	1,3	1,36	1,4	1,42	1,45	1,5	1,52
$H_i$ , %	0	2	4	3	5	8	7
Овес							
$P_{\Pi}$ , г/см <sup>3</sup>	1,4	1,42	1,45	1,5	1,52	1,55	
$H_i$ , %	0	3	12	24	22	35	

– Определить параметры уравнения связи между недобором урожайности для различных культур и плотностью почвы в следах движителей.

**Задача 6.** При проектировании животноводческих комплексов необходимо учитывать размеры санитарно-защитных зон вокруг них (комплексов). Граница распространения вредностей есть расстояние от комплекса, на котором загрязнения не превышают предельно допустимых концентраций (ПДК) (табл. 3.6).

Таблица 3.6

Границы санитарно-защитной зоны

Комплекс крупного рогатого скота							
Поголовье, тыс. гол.	2	3	4	5	6	7	8
Расстояние от комплекса, км	0,8	1,2	1,5	1,8	2	2,5	2,5
Свиноводческий комплекс							
Поголовье, тыс. гол.	20	30	40	50	60	70	
Расстояние от комплекса, км	1,2	1,6	2,4	2,7	3,2	3,6	

– Определить границу санитарно-защитной зоны для комплекса крупного рогатого скота с поголовьем 9 и 10 тыс. голов.

– Определить границу санитарно-защитной зоны для свиноводческого комплекса с поголовьем 80 и 100 тыс. голов.

*Задача 7.* Влажность травы, закладываемой на сенаж, должна быть в пределах 40–50%. Отклонение влажности массы от этого интервала ведет к резкому снижению качества корма. Разработан влагомер, который на основе зависимости электрического сопротивления проявленной массы от влажности позволяет в полевых условиях за 7–10 минут определить искомую влажность  $W$  кормов. Для его тарировки необходимо установить зависимость между указанными параметрами. На основе исследований получены данные для измельченной люцерны (табл. 3.7).

Таблица 3.7

Значения сопротивлений  $R$  влагомера

W, %	37			45			55			65			75		
Повторные опыты	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
R, кОм	48	45	46	27	25	24	15	16	13	13	12	10	8	7	5

– Описать функциональной зависимостью экспериментальные данные, полученные при определении влажности измельченной люцерны. Проверить гипотезу однородности дисперсий.

*Задача 8.* Один из возможных способов уборки риса – очесывание метелок на корню прямостоящего риса (с использованием щеточного очесывающего устройства). В результате испытаний очесывающего устройства получены экспериментальные данные зависимости общих потерь  $\Pi$  от приведенной подачи  $Q_{\Pi}$  (табл. 3.8).

Таблица 3.8

Значение общих потерь

$Q_{\Pi}$ , кг/с/м	0,9	1,2	1,4	1,8	2	2,2	2,4	2,7	3,4	3,7
$\Pi$ , %	10	7,6	6	4,5	4	3,5	3,6	3,6	6	7,5

– Определить зависимость потерь от приведенной подачи.

– Определить значение приведенной подачи, при которой потери минимальны.



*Задача 9.* На машиноиспытательной станции зерноуборочных комбайнов получены данные о производительности  $G$  машин и соответствующих ей потерях  $\Pi$  (табл. 3.9).

Таблица 3.9

Общие потери за комбайном

$G$ , кг/с	2	4	5	6,3	7	8	9	10	3	6	4
$\Pi$ , %	0,3	0,7	1	1,5	2	3,0	3,9	5	0,5	1,4	0,7

– Установить зависимость между указанными параметрами.

*Задача 10.* Для расчета конструктивных показателей пресс-подборщика и определения энергии на прессование необходимо знать усилие прессования, которое учитывает как сопротивление материала сжатию, так и его трение о стенки прессовальной камеры. Экспериментальные значения объемного модуля упругости  $K$  в зависимости от плотности различных материалов приведены в табл. 3.10.

Таблица 3.10

Экспериментальные значения модуля

Материал	Плотность, кг/м <sup>3</sup>				
	100	150	200	250	300
	Упругость, МПа				
Солома пшеничная влажностью 11,5 %	1,2	2	2,9	4,2	5,5
Солома овсяная влажностью 16,5 %	1	2	3,5	6	8,1
Сено естественных трав влажностью 24,3 %	2,1	3,5	5	7,2	10,1

– Определить эмпирическую зависимость объемного модуля упругости от плотности прессуемых материалов.

*Задача 11.* На машиностроительной станции получены данные о зависимости скорости  $V$  комбайна от урожайности  $U$  при различных соотношениях зерна к соломе и различной засоренности (табл. 3.11–3.14).

Таблица 3.11

Экспериментальные данные скорости при *большой* засоренности  
и соотношении зерна к соломе  $З:С$   
 $1:0,8 - 1:1$

Урожайность U, ц/га	Влажность хлебостоя, %		
	9-12	13-16	17-20
	Скорость, км/ч		
20	8,3	7,4	6,8
35	4,2	3,7	3,5
50	3,4	2,9	2,5
65	2,8	2,4	1,8

Таблица 3.12

Экспериментальные данные скорости при *малой* засоренности  
и соотношении зерна к соломе  $З:С$   
 $1:0,8 - 1:1$

Урожайность U, ц/га	Влажность хлебостоя, %		
	9-12	13-16	17-20
	Скорость, км/ч		
20	9,4	8,2	7,1
35	5,2	4,6	3,9
50	4,2	3,9	3,2
65	3,8	3,0	2,5

Таблица 3.13

Экспериментальные данные скорости при *большой* засоренности  
и соотношении зерна к соломе  $З:С$   
 $1:1,2 - 1:1,5$

Урожайность U, ц/га	Влажность хлебостоя, %		
	9-12	13-16	17-20
	Скорость, км/ч		
20	8,2	7,3	6,7
35	4,1	2,6	3,4
50	3,3	2,8	2,3
65	2,9	2,2	1,7

Таблица 3.14

Экспериментальные данные скорости при *малой* засоренности  
и соотношении зерна к соломе 3:С  
1:1,2 - 1:1,5

Урожайность У, ц/га	Влажность хлебостоя, %		
	9-12	13-16	17-20
	Скорость, км/ч		
20	8,5	7,6	6,9
35	4,5	3,9	3,7
50	3,7	3,1	2,8
65	3,0	2,4	2,18,3

– Установить регрессионную зависимость между скоростью комбайна и урожайностью при различных значениях засоренности и соотношения зерна к соломе.

### 3.4. Использование пакета Excel для построения регрессионных уравнений

*Последовательность действий при решении задач:*

1. Загрузите Excel.
2. Занесите в столбец известные значения  $Y$  (дробное значение вводить с запятыми):
  - выделите ячейку (адрес ячейки появится в поле имени);
  - введите с клавиатуры необходимый текст или значение (вводимое значение появится в строке формул).
3. Занесите в столбец известные значения  $X$ .
4. Выделите блок свободных ячеек размером в 5 строк и 5 столбцов: удерживая левую кнопку мыши, нужно отметить необходимый блок ячеек.
5. Введите функцию ЛИНЕЙН для массива известных значений  $Y$  и  $X$ :
  - в строке формул наберите, например,  
=ЛИНЕЙН (A7:A25;B7:D25; истина; истина)  
A7:A25 – адреса ячеек значений  $Y$   
B7:D25 – адреса ячеек значений  $X$   
(адреса ячеек вводите в английском регистре);

– удерживая нажатыми клавиши Ctrl и Shift нажмите клавишу Enter, тогда набранная формула будет воспринята как формула для массива;

– если функция была введена как функция для массива, то она автоматически будет взята в фигурные скобки;

$\{=\text{ЛИНЕЙН}(A7:A25;B7:D25; \text{истина}; \text{истина})\}$

– если все сделано правильно, то в выделенном блоке ячеек появится таблица результатов расчета.

6. В ячейку F7 введите формулу:  $= m_1x_1 + m_2x_2 + \dots + b$ , подставив вместо  $m_1, m_2, \dots, b$  их значения из таблицы результатов расчета, а вместо  $x_1, x_2, \dots$  - адреса ячеек;

– нажмите клавишу Enter (в ячейке появится расчетное значение Y).

7. Поместите курсор мыши в правый нижний угол ячейки F7, при этом изменится вид курсора.

8. Удерживая нажатой левую кнопку мыши, двигайте курсор по столбцу F вниз. Получите столбец расчетных значений Y.

Уравнение линии в общем случае имеет вид

$$y = m_1x_1 + m_2x_2 + \dots + b ,$$

для прямой линии

$$y = mx + b ,$$

где  $m$  – это коэффициенты, соответствующие каждой независимой переменной  $x$ ;  $b$  – постоянная.

Функция ЛИНЕЙН возвращает массив  $\{m_n, m_{n-1}, \dots, m_1, b\}$ . ЛИНЕЙН может также возвращать дополнительные статистики.

ЛИНЕЙН (известные\_значения\_y, известные\_значения\_x, конст, статистика)

*Известные\_значения\_y* – это множество значений  $y$ , которые уже известны для соотношения  $y = mx + b$ .

Если массив *известные\_значения\_y* имеет один столбец, то каждый столбец массива *известные\_значения\_x* интерпретируется как отдельная переменная.

Если массив *известные\_значения\_y* имеет одну строку, то каждая строка массива *известные\_значения\_x* интерпретируется как отдельная переменная.

*Известные\_значения\_x* – это необязательное множество значений  $x$ , которые уже известны для соотношения

$$y = tx + b .$$

Массив *известные\_значения\_x* может содержать одно или несколько множеств переменных. Если используется только одна переменная, то *известные\_значения\_y* и *известные\_значения\_x* могут быть массивами любой формы, при условии, что они имеют одинаковую размерность. Если используется более одной переменной, то *известные\_значения\_y* должны быть вектором (интервалом высотой в одну строку или шириной в один столбец).

Если *известные\_значения\_x* опущены, то предполагается, что это массив  $\{1, 2, 3, \dots\}$  такого же размера, как и *известные\_значения\_y*.

*Конст* – это логическое значение, которое указывает, требуется ли, чтобы константа  $b$  была равна 0.

Если *конст* имеет значение ИСТИНА или опущено, то  $b$  вычисляется обычным образом.

Если *конст\_* имеет значение ЛОЖЬ, то  $b$  полагается равным 0 и значения  $t$  подбираются так, чтобы выполнялось соотношение  $y = tx$ .

*Статистика* – это переменная логического типа, которая указывает, требуется ли вернуть дополнительную статистику по регрессии.

Если *статистика\_* имеет значение ИСТИНА, то функция ЛИНЕЙН возвращает дополнительную регрессионную статистику, т.е. возвращаемый массив будет иметь вид

$$\{m_n, m_{n-1}, \dots, m_1, b; se_n, se_{n-1}, \dots, se_1, se_b; r^2, se_y; F, df; ss_{reg}, ss_{resid}\}.$$

Если *статистика* имеет значение ЛОЖЬ или опущена, то функция ЛИНЕЙН возвращает только коэффициенты  $t$  и постоянную  $b$ .

Дополнительные статистики:

$se_1, se_2, \dots, se_n$  – стандартные значения ошибок для коэффициентов  $m_1, m_2, \dots, m_n$ ;

$se_b$  – стандартное значение ошибки для постоянной  $b$  ( $se_b = \#Н/Д$ , если *конст* имеет значение ЛОЖЬ);

$r^2$  – коэффициент детерминированности;

$se_y$  – стандартная ошибка для оценки  $y$ ;

$F$  – F-статистика используется для определения того, является ли наблюдаемая зависимость между зависимой и независимой переменными случайной или нет;

$df$  – степени свободы. Степени свободы полезны для нахождения F-критических значений в статистической таблице. Для определения уровня надежности модели нужно сравнивать значения в таблице с F-статистикой, возвращаемой функцией ЛИНЕЙН;

$SS_{reg}$  – регрессионная сумма квадратов;

$SS_{resid}$  – остаточная сумма квадратов.

В табл. 3.15 показано, в каком порядке возвращаются дополнительные статистики.

Таблица 3.15

Порядок расположения статистик

$m_n$	$m_{n-1}$	...	$m_2$	$m_1$	$b$
$se_n$	$se_{n-1}$	...	$se_2$	$se_1$	$se_b$
$r^2$	$se_v$				
$F$	$df$				
$SS_{reg}$	$SS_{resid}$				

Проводя регрессионный анализ, Excel «вычисляет» для каждой точки квадрат разности между прогнозируемым и фактическими значениями  $y$ . Сумма этих квадратов разностей называется *остаточной суммой квадратов*. Затем Excel «подсчитывает» сумму квадратов разностей между фактическими и средними значениями  $y$ , которая называется *общей суммой квадратов* (регрессионная сумма квадратов плюс остаточная сумма квадратов). Чем меньше остаточная сумма квадратов по сравнению с общей суммой квадратов, тем больше значение коэффициента детерминированности  $r^2$ , который показывает, насколько полученное уравнение объясняет взаимосвязь между переменными.

#### 4. ПОЛНЫЙ ФАКТОРНЫЙ ЭКСПЕРИМЕНТ

Полный факторный эксперимент (ПФЭ) – это первое звено в цепи планов, последовательное создание и совершенствование которых привело к разработке математических методов моделирования сложных процессов. В полном факторном эксперименте реализуются все возможные, неповторяющиеся комбинации уровней  $n$  независимых факторов. Планирование и реализация ПФЭ включает этапы: *выбор параметра оптимизации, факторов и уровней их варьирования; кодирование факторов; составление плана-матрицы эксперимента; рандомизация опытов; реализация плана эксперимента; проверка однородности дисперсий параллельных опытов, их воспроизводимости; расчет коэффициентов уравнения модели, их ошибок и значимости, а также проверка адекватности модели.*

По результатам двухфакторного эксперимента ( $n = 2$ ) можно построить математическую модель процесса, в которой помимо линейных членов будет член, учитывающий эффект парного межфакторного взаимодействия:

$$y = b_0 + b_1x_1 + b_2x_2 + b_{12}x_1x_2 .$$

План ПФЭ типа  $2^3$  ( $n = 3$ ) позволяет построить модель вида

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_{12}x_1x_2 + b_{13}x_1x_3 + b_{23}x_2x_3 + b_{123}x_1x_2x_3 .$$

Коэффициенты  $b_{12}$ ,  $b_{13}$  и  $b_{23}$  характеризуют эффект парного взаимодействия факторов  $x_1$ ,  $x_2$  и  $x_3$ , а коэффициент  $b_{123}$  – эффект тройного межфакторного взаимодействия.

План ПФЭ типа  $2^4$  ( $n = 4$ ) дает возможность рассчитать 16 коэффициентов модели:

$$\begin{aligned} y = & b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4 + b_{12}x_1x_2 + \\ & + b_{13}x_1x_3 + b_{14}x_1x_4 + b_{23}x_2x_3 + b_{24}x_2x_4 + \\ & + b_{34}x_3x_4 + b_{123}x_1x_2x_3 + b_{124}x_1x_2x_4 + b_{234}x_2x_3x_4 + \\ & + b_{134}x_1x_3x_4 + b_{1234}x_1x_2x_3x_4 . \end{aligned}$$

Чтобы определить, какое влияние на функцию  $y$  оказывает каждый фактор  $x_i$ , проводят эксперимент, в котором меняют значения  $x_i$ , измеряя при этом  $y$ .

Значения факторов  $x_i$  задают на двух уровнях ( $x_{i \min}$  и  $x_{i \max}$ ), или на трех ( $x_{i \min}$ ,  $x_{i \text{ ср}}$ ,  $x_{i \max}$ ), или по другой схеме, при этом в пол-

ном факторном эксперименте проводят опыты для всех сочетаний факторов друг с другом, что при двух уровнях варьирования факторов дает общее число  $2^n$  опытов, при трех уровнях –  $3^n$ , где  $n$  – число факторов  $x_i$  ( $i=1, 2, \dots, n$ ).

Приведенные модели являются линейными относительно факторов. Это значит, что если зафиксировать все факторы, кроме одного, то зависимость  $y$  от  $x_i$  будет линейной и график  $y = f(x_i)$  будет прямой линией, возрастающей при  $b_i > 0$  и убывающей при  $b_i < 0$ .

Известно, что для построения прямой достаточно задать две ее точки, поэтому для определения зависимостей рассмотренного типа достаточно двухуровневого эксперимента, т.е. определения отклика  $y$  в двух точках –  $x_{i \min}$  и  $x_{i \max}$ .

Линейные модели и двухуровневый эксперимент можно применять, если заранее известно, что функция отклика в исследуемом диапазоне изменения факторов близка к линейной. Однако часто эта функция может быть нелинейной.

Трехуровневый эксперимент позволяет избежать ошибки при определении вида модели, точнее оценить коэффициенты линейной модели и является обязательным при построении нелинейной модели, т.е. модели, в которой факторы  $x_i$  имеют степень, отличную от единицы.

#### 4.1. Составление плана ПФЭ типа $2^n$

План типа  $2^n$  является самым простым планом, у которого ортогональны любые два столбца независимых переменных. При этом исследуемые факторы изменяются на двух уровнях: верхнем  $Z_{\max}$  и нижнем  $Z_{\min}$ . При планировании эксперимента преобразуют размерные управляемые независимые факторы  $Z_i$  в безразмерные, нормированные  $X_i$  по формуле

$$X_i = \frac{Z_i - Z_{i0}}{\Delta Z_i},$$

где  $X_i$  – кодовое значение  $i$ -го фактора;  $Z_i$  – натуральное значение  $i$ -го фактора;  $Z_{i0}$  – центр эксперимента или начальный (нулевой) уровень  $i$ -го фактора;  $\Delta Z_i$  – интервал варьирования  $i$ -го фактора.



Интервал варьирования определяется как

$$\Delta Z_i = \frac{Z_{imax} - Z_{imin}}{2},$$

а центр эксперимента по формуле

$$Z_{i0} = \frac{Z_{imax} + Z_{imin}}{2}.$$

План эксперимента можно составлять разными способами, в частности способом, изображенным в табл. 4.1 и 4.2. Значения первого фактора  $x_1$  чередуются ( $x_{min}$ ,  $x_{max}$ ,  $x_{min}$  и т. д.) в каждом опыте, значения  $x_2$  – через два опыта,  $x_3$  – через четыре опыта,  $x_4$  – через восемь опытов и т.д. (знак *плюс* соответствует максимальному значению фактора, знак *минус* – минимальному).

Таблица 4.1

План эксперимента типа  $2^2$

$u$	$x_1$	$x_2$
1	-	-
2	+	-
3	-	+
4	+	+

Таблица 4.2

План эксперимента типа  $2^3$

$u$	$x_1$	$x_2$	$x_3$
1	-	-	-
2	+	-	-
3	-	+	-
4	+	+	-
5	-	-	+
6	+	-	+
7	-	+	+
8	+	+	+

Если требуется получить модель с учетом межфакторных взаимодействий, то в матрицу добавляют столбцы с соответствующими членами модели, при этом знак столбца для каждого опыта определя-

ется по знаку произведения входящих в этот член факторов, например знак столбца  $x_1x_2x_3$  в плане эксперимента (табл. 4.3) для строки 3 равен  $(-1)(+1)(-1) = 1$ .

Таблица 4.3

Матрица эксперимента с учетом межфакторных взаимодействий

$u$	$x_1$	$x_2$	$x_3$	$x_1x_2$	$x_1x_3$	$x_2x_3$	$x_1x_2x_3$
1	-	-	-	+	+	+	-
2	+	-	-	-	-	+	+
3	-	+	-	-	+	-	+
4	+	+	-	+	-	-	-
5	-	-	+	+	-	-	+
6	+	-	+	-	+	-	-
7	-	+	+	-	-	+	-
8	+	+	+	+	+	+	+

**Пример.** При оптимизации работы молотильного аппарата при уборке кукурузы в качестве управляемых независимых факторов выбраны:  $Z_1$  – зазор на входе в молотильный аппарат и  $Z_2$  – зазор на выходе из молотильного аппарата, мм. Возможные интервалы изменения факторов составляют для  $30 < Z_1 < 60$ ;  $15 < Z_2 < 35$ .

Известно наиболее подходящее сочетание факторов:  $Z_{10} = 40$  мм,  $Z_{20} = 20$  мм. Интервал варьирования факторов в эксперименте принять шириной 20% от интервала изменения соответствующих факторов.

*Определить* интервал варьирования, нижние и верхние уровни факторов, составить матрицу эксперимента, графически изобразить план.

*Решение.*

Интервал варьирования для 1 – го фактора равен  $0,2 \cdot 30 = 6$  мм, для второго фактора равен  $0,2 \cdot 20 = 4$  мм.

Центр эксперимента  $Z_{10} = 40$  мм,  $Z_{20} = 20$  мм.

Верхние уровни факторов  $Z_{1max} = 46$  мм,  $Z_{2max} = 24$  мм.

Нижние уровни факторов  $Z_{1min} = 34$  мм,  $Z_{2min} = 16$  мм.

Область определения (обозначена штриховкой) может быть изображена на плоскости в соответствующей размерности факторов (рис. 4.1, а) и при безразмерном выражении величин факторов (рис. 4.1, б).

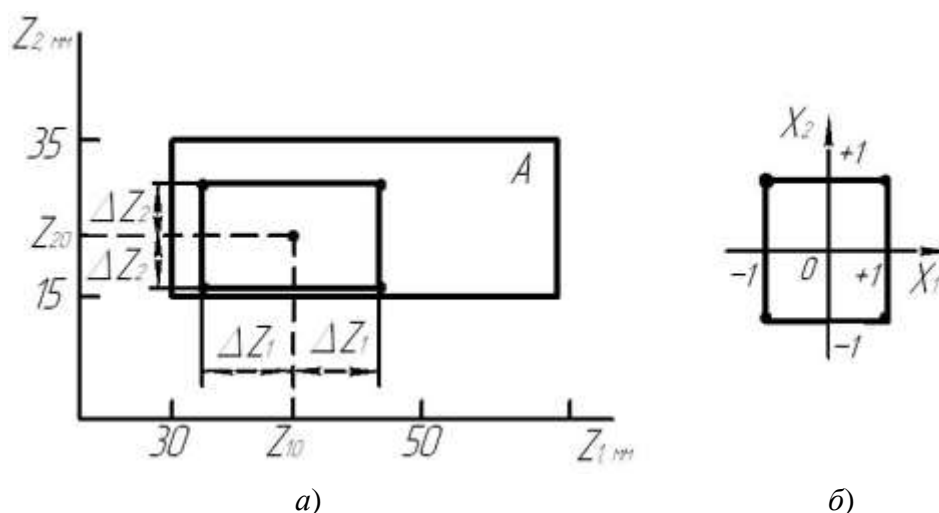


Рис. 4.1. Графическая интерпретация плана ПФЭ  $2^2$   
(А – область допустимых значений факторов)

План эксперимента приведен в табл. 4.4.

Таблица 4.4

Матрица эксперимента

$u$	$z_{1u}$	$z_{2u}$
1	34	16
2	46	16
3	34	24
4	46	24
$Z_{i0}$	40	20
$\Delta Z_i$	6	4

### Задачи для упражнений

**Задача 1.** При оптимизации работы молотильного аппарата при уборке зерновых культур с целью снижения потерь зерна в качестве управляемых факторов выбраны: частота вращения молотильного барабана -  $Z_1$ , мин<sup>-1</sup>; зазор на входе -  $Z_2$ , мм; зазор на выходе -  $Z_3$ , мм.

Пределы изменения факторов следующие:

$$Z_{1min} = 800; Z_{1max} = 1200; Z_{2min} = 16;$$

$$Z_{2max} = 24; Z_{3min} = 6; Z_{3max} = 10.$$

– Произвести кодирование факторов. Составить план эксперимента. Графически изобразить область исследования при безразмерном выражении факторов.

– Произвести кодирование факторов. Составить план эксперимента с учетом парных взаимодействий.

*Задача 2.* При исследовании энергоемкости процесса измельчения листостебельной массы в дробилке агрегата травяной муки выявлены четыре наиболее значимых фактора:  $Z_1$  – зазор между концами молотков и стенкой камеры дробилки, мм;  $Z_2$  – окружная скорость молотков, м/с;  $Z_3$  – количество молотков на роторе;  $Z_4$  – подача материала в дробилку, кг/ч. Центром эксперимента могут быть значения  $Z_{10} = 25$ ;  $Z_{20} = 100$ ;  $Z_{30} = 32$ ;  $Z_{40} = 550$ . Интервал варьирования факторов принять 20 – 25% от нулевого уровня факторов.

– Произвести кодирование факторов. Составить план эксперимента с учетом межфакторных взаимодействий.

## 4.2. Рандомизация опытов

Поскольку изменение отклика  $Y$  имеет случайный характер, то в каждой точке факторного пространства необходимо проводить  $m$  повторных опытов и результаты наблюдений усреднять. Перед реализацией плана на объекте необходимо рандомизировать варианты варьирования факторов, т.е. используя таблицу равномерно распределенных случайных чисел (прил. 1), определить последовательность реализации вариантов варьирования плана в  $Nm$  опытах (где  $N$  – число строк матрицы ПФЭ, а  $m$  – число повторных опытов).

**Пример.** Произвести рандомизацию опытов ПФЭ типа  $2^2$  с тремя повторными опытами.

**Решение.** Из условия задачи следует, что  $N = 4$ ,  $m = 3$ , общее количество опытов равно 12. Матрица планирования эксперимента с

рандомизированной последовательностью опытов приведена в табл. 4.5.

Таблица 4.5

Матрица планирования эксперимента  
с рандомизированной последовательностью опытов

$u$	$x_1$	$x_2$	Последовательность проведения опытов		
			$y_{1u}$	$y_{2u}$	$y_{3u}$
1	-	-	10	4	6
2	+	-	9	2	7
3	-	+	1	8	11
4	+	+	5	3	12

### Задачи для упражнений

Рандомизировать опыты:

- 1) для ПФЭ типа  $2^3$  с тремя повторными опытами;
- 2) ПФЭ типа  $2^2$  с пятью повторными опытами;
- 3) ПФЭ типа  $2^4$  с тремя повторными опытами;
- 4) ПФЭ типа  $2^3$  с пятью повторными опытами;
- 5) ПФЭ типа  $2^4$  с двумя повторными опытами.

### 4.3. Проверка воспроизводимости эксперимента

Проверка воспроизводимости эксперимента есть ни что иное, как проверка однородности выборочных дисперсий  $S_u^2$ .

Оценки дисперсий находят по формуле

$$S_u^2 = \frac{1}{m-1} \sum_{i=1}^m (y_{ui} - \bar{y}_u)^2,$$

где  $\bar{y}_u = \frac{\sum_{i=1}^m y_{ui}}{m}$ .

Так как значения дисперсий получают по выборкам одинакового объема  $m$ , то число степеней свободы для них одинаково и составляет  $v_{\text{вос}} = m - 1$ .

В этом случае для проверки гипотезы об однородности дисперсий можно использовать критерий Кохрена, статистика которого имеет вид

$$G = \frac{\max\{S_u^2\}}{\sum_m^n S_u^2\{y\}}.$$

Если расчетное значение критерия  $G_p$  окажется меньше табличного значения  $G_{\text{табл}}$  для  $v_{1\text{вос}} = m - 1$  и  $v_{1\text{вос}} = n$  при заданном уровне значимости, то гипотеза об однородности выборочных дисперсий принимается. В этом случае оценка генеральной дисперсии имеет вид

$$S_{\text{вос}}\{y\} = \frac{1}{n} \sum_{u=1}^n S_u^2\{y\}$$

с числом степеней свободы  $v = n(m - 1)$ .

В противном случае, если  $G_p > G_{\text{табл}}$ , то гипотеза об однородности дисперсий отвергается.

Некоторые значения критерия Кохрена приведены в табл. 4.6

Таблица 4.6

Значения критерия Кохрена при уровне значимости 0,05

Степень свободы	$v_2$		
$v_1$	4	8	16
1	0,9065	0,6798	0,4789
2	0,7679	0,5157	0,3358
3	0,6841	0,4377	0,2768
4	0,6287	0,3910	0,2429

**Пример.** В результате реализации плана эксперимента типа  $2^3$  получены экспериментальные данные (табл. 4.7).

Таблица 4.7

Данные модельного примера ( $m = 3$ )

$u$	$y_{u1}$	$y_{u2}$	$y_{u3}$	$\bar{y}_u$	$S_u^2$
1	73	69	68	70	7
2	58	58	64	60	12
3	54	59	52	55	13
4	84	94	92	90	28
5	100	106	109	105	21
6	98	90	97	95	19
7	77	85	78	80	19
8	105	95	100	100	25

Значения  $y_u$  и  $S_u^2$  приведены в табл. 4.7.

Расчетное значение критерия Кохрена  $G_p = 0,2$ ; табличное -  $G_{\text{табл}} = 0,5157$  при  $v_1 = 3 - 1 = 2$ ;  $v_2 = 8$  и  $\alpha = 0,05$ . Видно, что  $G_p < G_{\text{табл}}$ , поэтому гипотеза об однородности не отвергается.

Оценка дисперсии воспроизводимости определяется как

$$S_{\text{вос}}^2\{y\} = \frac{7 + 12 + 13 + 28 + 21 + 19 + 19 + 25}{8} = \frac{144}{8} = 18,$$

$$с \nu = 8 \cdot (3 - 1) = 16.$$

### Задачи для упражнений

*Задача 1.* Проверить воспроизводимость эксперимента, данные которого приведены в табл. 4.8.

Таблица 4.8

Экспериментальные данные к задаче 1

$u$	$y_{u1}$	$y_{u2}$	$y_{u3}$	$y_{u4}$	$\bar{y}_u$	$S_u^2$
1	2,4	2,5	2,4	2,6		
2	3,3	3,3	3,3	3,5		
3	2,9	2,8	2,6	2,8		
4	1,8	2,0	1,9	1,8		

*Задача 2.* Проверить однородность дисперсий замеров массовой концентрации аммиака ( $\text{мг/м}^3$ ) в животноводческом помещении с целью выявления систематических погрешностей универсального газоанализатора УГ – 2.

Эксперимент проводился на трех уровнях:

1-й уровень	7,5	9,0	8,0	8,0	9,5
2-й уровень	28,5	28,0	25,0	30,0	27,0
3-й уровень	54,0	60,0	55,5	58,0	60,5

### 4.4. Построение и анализ математической модели

Пользуясь методом ПФЭ, можно получить описание изучаемого процесса в виде различных полиномиальных моделей.

Формулы для определения оценок коэффициентов линейного уравнения имеют вид:

$$b_i = \frac{\sum_{u=1}^n x_{iu} \bar{y}_u}{n}, i = \overline{1, n};$$

$$b_0 = \frac{\sum_{u=1}^n \bar{y}_u}{n}.$$

Оценки эффектов межфакторных взаимодействий определяются по формулам:

$$b_{ij} = \frac{\sum_{u=1}^n x_{iu} x_{ju} \bar{y}_u}{n}, \quad i \neq j; \quad i, j = \overline{1, n};$$

$$b_{ijk} = \frac{\sum_{u=1}^n x_{iu} x_{ju} x_{ku} \bar{y}_u}{n}, \quad i \neq j \neq k; \quad i, j, k = \overline{1, n};$$

$$b_{ijkl} = \frac{\sum_{u=1}^n x_{iu} x_{ju} x_{ku} x_{lu} \bar{y}_u}{n}, \quad i \neq j \neq k \neq l; \quad i, j, k, l = \overline{1, n}.$$

После определения оценок  $b$  коэффициентов уравнения модели необходимо проверить гипотезу об их значимости, другими словами, проверить, значимо ли полученная оценка отличается от нуля. Для этого используют критерий Стьюдента, расчетное значение которого определяется как

$$t_p = \frac{|b|}{S^2\{b\}},$$

где  $S^2\{b\} = \frac{s_{\text{вос}}^2\{y\}}{nm}$ ;  $n$  – число точек факторного пространства, в которых проводился эксперимент;  $m$  – число повторных опытов в этих точках.

Если найденная величина критерия  $t_p$  превышает табличное  $t_{\text{табл}}$  (см. прил. 2) для числа степеней свободы  $\nu = n(m - 1)$  при заданном  $\alpha$ , то оценку  $b$  коэффициента модели признают значимой. В противном случае (если  $t_p \leq t_{\text{табл}}$ ) оценка  $b$  считается незначимой, т.е. равной нулю. В этом случае математическую модель объекта составляют в виде уравнения связи отклика  $y$  и факторов  $x$ , включающего только значимые оценки коэффициентов.

Для проверки гипотезы адекватности модели необходимо рассчитать теоретические значения  $\hat{y}_u$ . Рассеяние результатов наблюдений вблизи уравнения регрессии характеризуется дисперсией адекватности:

$$S_{\text{ад}}^2 = \frac{m}{n - a} \sum_{u=1}^n (\bar{y}_u - \hat{y}_u)^2,$$

где  $a$  – число членов модели.

Дисперсия адекватности определяется числом степеней свободы  $\nu_{\text{ад}} = n - a$ .

Проверяют гипотезу адекватности, используя критерий Фишера (F - критерий):



$$F_p = \frac{S_{ад}^2}{S_{вос}^2\{y\}}.$$

Если расчетное значение критерия  $F_p$  меньше табличного  $F_{табл}$  для соответствующих степеней свободы

$$v_{1ад} = n - a; \quad v_{2ад} = v_{вос} = n(m - 1)$$

при заданном уровне значимости, то гипотезу об адекватности не отвергают. В противном случае модель признается неадекватной.

**Пример.** Построить математическую модель по данным предыдущего примера:

$$S_{вос}^2\{y\} = 18; \quad n = 8; \quad m = 3; \quad t_{табл}(0,95; 16) = 2,12.$$

Расчет коэффициентов модели:

$$b_1 = \frac{\sum_{u=1}^n x_{1u} \bar{y}_u}{n} = \frac{-70 - 60 + 55 + 90 - 105 - 95 + 80 + 100}{8} = -0,625;$$

$$b_{12} = \frac{\sum_{u=1}^n x_{1u} x_{2u} \bar{y}_u}{n} = \frac{70 - 60 - 55 + 90 + 105 - 95 - 80 + 100}{8} = 9,375$$

и т.д.

Уравнение модели имеет вид

$$y = 81,875 - 0,625x_1 + 4,375x_2 + 13,125x_3 + 9,375x_1x_2 - \\ - 1,875x_2x_3 - 4,375x_1x_3 - 1,875x_1x_2x_3.$$

Далее, определяется стандартное отклонение для коэффициента  $b$ :

$$b^2\{b\} = \frac{18}{8 \cdot 3} = 0,75, \quad S\{b\} = 0,87.$$

Находим значение

$$t_p(b_1) = \frac{|-0,625|}{0,87} \cong 0,72; \quad t_p(b_2) = \frac{4,375}{0,87} \cong 5,02.$$

При  $\alpha = 0,05$  и  $v = 16$ ,  $t_{табл} = 2,12$ .

Уравнение модели будет иметь вид:

$$y = 81,875 + 4,375x_2 + 13,125x_3 + 9,375x_1x_2 - \\ - 1,875x_2x_3 - 4,375x_1x_3 - 1,875x_1x_2x_3.$$

Для проверки гипотезы адекватность полученной модели необходимо рассчитать значения  $\hat{y}_u$  по данной модели, например, для 1-го опыта  $x_{21} = -1$ ;  $x_{31} = -1$  и т.д.

$$\hat{y}_1 = 81,875 - 4,375 - 13,125 + 9,375 -$$

$$-1,875 - 4,375 + 1,875;$$

$$\hat{y}_1 = 69,375 \text{ и т.д.}$$

Результаты расчетов сведены в табл. 4.9.

Таблица 4.9

План ПФЭ типа  $2^3$

$u$	Матрица плана							Результаты	
	$x_{1u}$	$x_{2u}$	$x_{3u}$	$x_{1u}x_{2u}$	$x_{2u}x_{3u}$	$x_{2u}x_{3u}$	$x_{1u}x_{2u}x_{3u}$	опытов $\bar{y}_u$	расчетов
1	-	-	-	+	+	+	-	70	69,375
2	-	+	-	-	-	+	+	60	59,375
3	+	-	-	-	+	-	+	55	55,625
4	+	+	-	+	-	-	-	90	90,625
5	-	-	+	+	-	-	+	105	105,375
6	-	+	+	-	+	-	-	95	94,375
7	+	-	+	-	-	+	-	80	80,625
8	+	+	+	+	+	+	+	100	100,625

Определяем дисперсию адекватности

$$S_{\text{ад}}^2 = \frac{3}{8-7} \cdot 3,12 = 9,36.$$

Определяем расчетное значение  $F$  – критерия (при этом необходимо в числитель выражения подставлять большее из двух значений дисперсии).

$$F_p = \frac{S_{\text{вос}}^2\{y\}}{S_{\text{ад}}^2} = \frac{18}{9,36} = 1,92.$$

При  $v_1 = 8 \cdot (3 - 1)$  и  $v_2 = 8 - 7 = 1$ , и  $\alpha = 0,05$  находим табличное значение  $F_{\text{табл}} = 246$ .

Так как  $F_p = 1,92 \ll F_{\text{табл}} = 246$ , гипотезу адекватности модели не отвергаем.

### Задачи для упражнений

**Задача 1.** Построить эмпирическую зависимость степени измельчения соломы от скорости молотков (измельчитель ИРТ - 80) и влажности соломы. Экспериментальные данные приведены в табл. 4.10. Скорость молотков меняется от 40 до 70 м/с, влажность соломы –

от 8,9 до 36,6%. Степень измельчения соломы характеризуется процентным содержанием частиц длиной до 50 мм.

Таблица 4.10

Экспериментальные данные к задаче 1

$u$	$x_{1u}$	$x_{2u}$	$x_{1u}x_{2u}$	$y_{u1}$	$y_{u2}$	$y_{u3}$	$\bar{y}_u$	$S_u^2$	$\hat{y}$
1	-	-	+	50	54	48			
2	+	-	-	64	63	65			
3	-	+	-	22	18	21			
4	+	+	+	86	85	86			

Проверить гипотезы о значимости коэффициентов уравнения и адекватности модели.

*Задача 2.* Работоспособность ременно-планчатых транспортеров валковых жаток (ЖВН – 6; ЖРБ – 4,2) определяется тяговой способностью, долговечностью и ползучестью (вытяжкой) отдельных лент транспортера.

Анализ исследований по тяговой способности плоских транспортных лент и приводных ремней показал, что основными факторами, определяющими тяговую способность лент, являются:

– *конструктивное оформление ленты:*

«+ 1» - лента с привулканизированными прокладками;

«- 1» - серийная лента;

– *жесткостные свойства ленты:*

«+ 1» - лента предварительно обкатана;

«- 1» - лента в состоянии поставки (не обкатана);

– *способ сшивки лент:*

«+ 1» - сшивка «гребешком» с прорезиненной прокладкой;

«- 1» - сшивка с помощью металлических накладок;

– *удельное начальное натяжение:*

«+ 1» - 1,2 МПа;

«- 1» - 0,8 МПа.

В качестве функции отклика было выбрано значение критического крутящего момента  $M_k$ . Матрица планирования четырехфакторного эксперимента и результаты опытов приведены в табл. 4.11.

Таблица 4.11

## Исходные данные к задаче 2

№ опыта	$x_1$	$x_2$	$x_3$	$x_4$	$y_{u1}$	$y_{u2}$	$\bar{y}_u$	$S_u^2$	$\hat{y}_1$
1	+	+	+	+	2,58	2,62			
2	-	+	+	+	2,53	2,57			
3	+	-	+	+	2,24	2,26			
4	-	-	+	+	2,14	2,16			
5	+	+	-	+	2,50	2,50			
6	-	+	-	+	2,16	2,14			
7	+	-	-	+	1,96	1,94			
8	-	-	-	+	2,02	2,08			
9	+	+	+	-	4,84	4,86			
10	-	+	+	-	4,77	4,73			
11	+	-	+	-	3,80	3,70			
12	-	-	+	-	3,90	3,50			
13	+	+	-	-	4,00	3,60			
14	-	+	-	-	3,54	3,56			
15	+	-	-	-	2,70	2,50			
16	-	-	-	-	2,96	3,05			

– Построить уравнение модели и произвести проверку гипотезы адекватности модели.

## 4.5. Пример многофакторного планирования эксперимента

### 4.5.1. Постановка задачи

Решим задачу определения основных факторов и степени их влияния на эффективность пневмосепарации зернового вороха зерноуборочного комбайна. Под декой молотильного барабана комбайна установлено устройство, например труба с отверстиями, из которого под давлением выходят струи воздуха. Зерновой ворох, просыпающийся сквозь деку, продувается струями, при этом скорость воздушного потока подобрана такой, что мелкие примеси (полова) подхватываются и уносятся струей к концу верхнего решета, а зерно падает на стрясную доску. Это дает возможность уменьшить количество примесей, попа-

дающих на очистку, и тем самым повысить производительность комбайна.

В качестве критериев оптимизации выбраны эффективность пневмосепарации  $E$  и вынос полноценного зерна  $\Pi$ . Факторы, влияющие на выходные величины, и пределы варьирования их значений представлены ниже.

Имя фактора	Обозначение
Скорость выхода струй, м/с.....	X1
Диаметр отверстий, мм.....	X2
Расстояние между отверстиями, мм.....	X3
Подача, кг/с.....	X4
Толщина продуваемого слоя, мм.....	X5

При решении многофакторных задач получается математическая модель процесса, которая увязывает воедино все учтенные факторы. Геометрический образ функции отклика называется поверхностью отклика в факторном пространстве. Функцию отклика можно аппроксимировать полиномом вида

$$Y = b_0 + \sum_{i=1}^n b_i x_i + \sum_{i=1}^n b_{ij} x_i x_j + \sum_{i=1}^n b_{ii} x_i^2 + \dots, \quad (4.1)$$

где  $Y$  – значение критерия оптимизации;  $b_0, b_i, b_{ij}, b_{ii}$  – коэффициенты, по величине которых можно судить о степени влияния соответствующих факторов ( $b_{ij}$  – эффекты взаимодействий;  $b_{ii}$  – эффекты при квадратичных членах).

Целью эксперимента является определение численных значений коэффициентов уравнения регрессии.

Для отыскания оптимальных условий протекания процесса находят значения факторов  $x_1, x_2, \dots, x_n$ , соответствующих экстремуму целевой функции.

При планировании эксперимента важным является выбор критерия оптимизации – параметра, по которому оценивается исследуемый процесс. Задача заключается в минимизации или максимизации критерия оптимизации. Необходимо выбрать такой критерий, который являлся бы совокупной и исчерпывающей характеристикой объекта исследования, имел ясный физический смысл и количественную оценку.

При планировании эксперимента ставилась задача определения математической модели в виде уравнения второго порядка, адекватно описывающего процесс пневмосепарации. Для определения коэффициентов уравнения был реализован экспериментальный план Хартли на гиперкубе (близкий по своим статистическим характеристикам к Д-оптимальному плану).

Для описания процесса пневмосепарации уравнением второго порядка необходимо применить такой план эксперимента, в котором каждый фактор можно варьировать не менее чем на трех уровнях при условии получения отдельных оценок всех коэффициентов уравнения по результатам эксперимента.

При центральном композиционном планировании второго порядка общее число точек эксперимента  $N$  при  $n$  факторах определяется зависимостью

$$N = 2^n + 2^n + n_0, \quad (4.2)$$

где  $n_0$  – число точек в центре эксперимента.

В рассматриваемом случае ядром плана служит полуреплика факторного эксперимента  $2^{n-1}$ . Тогда выражение (4.2) примет вид

$$N = 2^{n-1} + 2n + n_0,$$

где  $2^{n-1}$  – полуреплика многофакторного эксперимента (ядро плана);  $n$  – число факторов;  $2n$  – число звездных точек;  $n_0 = 1$  – нулевая точка.

При пяти варьируемых параметрах имеем

$$N = 2^{5-1} + 2 \cdot 5 + 1 = 27 \text{ опытов.}$$

Матрица планирования эксперимента для  $N = 27$  приведена в табл. 4.12.

Таблица 4.12

План – матрица эксперимента

$u$	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$
1	-1	-1	-1	-1	-1
2	-1	-1	-1	-1	-1

3	-1	-1	-1	-1	-1
4	-1	-1	-1	-1	-1
5	-1	-1	-1	-1	-1
6	-1	-1	-1	-1	-1
7	-1	-1	-1	-1	-1
8	-1	-1	-1	-1	-1
9	-1	-1	-1	-1	-1
10	-1	-1	-1	-1	-1
11	-1	-1	-1	-1	-1
12	-1	-1	-1	-1	-1
13	-1	-1	-1	-1	-1
14	-1	-1	-1	-1	-1
15	-1	-1	-1	-1	-1
16	-1	-1	-1	-1	-1
17	0	0	0	0	0
18	-1	0	0	0	0
19	-1	0	0	0	0
20	0	-1	0	0	0
21	0	-1	0	0	0
22	0	0	-1	0	0
23	0	0	-1	0	0
24	0	0	0	-1	0
25	0	0	0	-1	0
26	0	0	0	0	-1
27	0	0	0	0	-1

При идентификации изучаемого объекта можно выдвигать различные требования к свойствам математической модели. В условиях активного эксперимента выполнение этих требований приводит к необходимости выбора планов эксперимента, обладающих различными критериями оптимальности.

Все многообразие критериев разделяют на две группы. К первой группе относятся критерии, связанные с точностью оценок параметров. Ко второй – критерии и свойства планов, связанные с ошибкой в оценке модели. Для исследования многофакторных процессов широко применяются планы второго порядка, близкие по статистическим характеристикам к Д-оптимальным.

Основным преимуществом Д-оптимальных планов является то, что они минимизируют обобщенную дисперсию или объем эллипсоида рассеяния оценок параметров. Эффективность Д-оптимальных планов обуславливается оптимальным расположением точек в пространстве факторов.

В эксперименте использование Д-оптимальных планов предполагает варьирование факторов (для описания поверхности отклика полиномом второго порядка) только на трех уровнях вместо пяти уровней факторов у рототабельных, ортогональных и других. Это сокращает время проведения эксперимента и позволяет повысить точность результатов экспериментального исследования, поскольку Д-оптимальные планы и планы, близкие к ним, позволяют получить наиболее эффективные оценки коэффициентов уравнения модели. Также положительным свойством принятого плана является то, что ядро плана (полуреплика полного факторного эксперимента  $2^{5-1}$ ) образовано на основе генерирующего соотношения  $X_5 = X_1X_2X_3X_4$ , позволяющего оценить линейные коэффициенты  $b_i$  независимо друг от друга.

Планы, отвечающие критериям А – оптимальности, обеспечивают минимум средней дисперсии оценок коэффициентов.

Планы, отвечающие критериям Е – оптимальности, не допускают, чтобы отдельные оценки параметров имели слишком большие дисперсии и ковариации.

Свойство ортогональности плана позволяет получить оценки параметров модели взаимно независимыми.

К критериям второй группы относятся G – оптимальность и Q – оптимальность, которые соответственно обеспечивают минимум максимального значения дисперсии предсказания по модели и минимум средней дисперсии предсказания по модели. Кроме этого, важными свойствами планов являются рототабельность и композиционность. Рототабельность плана обеспечивает постоянство дисперсии предсказания на равных расстояниях от центра эксперимента. Композиционность плана позволяет проводить эксперимент в несколько этапов и



постепенно переходить от простых моделей к более сложным, используя предыдущие наблюдения.

Для исследования влияния варьируемых факторов на эффективность пневмосепарации был реализован симметричный центральный композиционный план эксперимента Хартли  $Ha_5$ , построенный на гиперкубе.

*Рандомизация опытов.* Для уменьшения влияния медленно изменяющихся воздействий на результаты эксперимента применяется принцип рандомизации, т. е. опыты проводятся не в том порядке, какой указан в плане, а в случайной последовательности. Рандомизировать опыты можно, используя таблицу равномерно распределенных случайных чисел или генерируя псевдослучайные числа. Например, алгоритм генерации псевдослучайных чисел, распределенных по равномерному закону в интервале  $[a, b]$ , имеет вид:

$$V_{i+1} = a + (b - a)V_i;$$

$$V_{i+1} = F\{11V_i + p\},$$

где  $F\{\}$  – символ дробной части числа; при  $i = 0$ ,  $V_i = V_0$  – стартовое значение (например,  $V_0 = 0,5$ ).

Общее число опытов равно  $Nm = 27 \cdot 3 = 81$ , где  $N = 27$  – число строк плана дробного факторного эксперимента (ДФЭ),  $m = 3$  – число повторных опытов.

При заданной доверительной вероятности 0,95 и точности  $e = 2,5S_0$  число повторных опытов определялось по формуле

$$m \geq |t(P)/e|^2 S_0^2,$$

где  $t(P)$  определяется из равенства  $2\Phi(t) = P$  по статистическим таблицам. Расчеты показали, что  $m = 2,9$ , отсюда принимаем  $m = 3$ .

Матрица планирования эксперимента с рандомизированной последовательностью опытов приведена в табл. 4.13.

*Кодирование факторов.* Планирование и обработка данных эксперимента осуществляется не в физических  $Z_i$ , а в кодированных переменных  $X_i$ .

Основным (нулевым) уровнем фактора называется значение, относительно которого фактор изменяется в процессе эксперимента.

С точки зрения информативности эксперимента интервал варьирования надо брать достаточно большим, так как в противном случае нарушаются предпосылки регрессионного анализа.

Таблица 4.13

Последовательность проведения опытов

$y_{u1}$	$y_{u2}$	$y_{u3}$
10	9	73
25	33	76
52	1	35
34	67	48
80	17	39
29	27	49
37	54	20
5	64	47
42	24	40
63	61	4
2	16	65
8	26	53
19	50	3
23	60	15
6	7	38
31	13	11
43	62	59
12	79	70
66	68	77
14	57	36
69	81	45
18	30	74
32	56	21
55	75	28
46	78	44
22	58	51
71	41	72

Чрезмерному увеличению интервала варьирования препятствует требование адекватности модели. При больших интервалах варьирова-

ния реальная поверхность отклика может сильно отличаться от экспериментально полученной аппроксимирующей поверхности.

Таблица 4.14

Условные обозначения факторов, уровни и интервал их варьирования

Обозначение фактора	Интервал варьирования	Уровни варьирования		
		нижний «-»	основной «0»	верхний «+»
X1	2	50	60	70
X2	1	4	5	6
X3	10	50	60	70
X4	2	5	7	9
X5	150	200	350	500

#### 4.5.2. Обработка данных многофакторного эксперимента

Результаты многофакторного эксперимента представляют собой выборку из некоторой генеральной совокупности, т.е. являются случайными величинами. Основная задача обработки – оценка параметров принятой полиномиальной модели и проверка гипотезы об адекватности модели. Если принятая модель не отвечает требованию адекватности, следует использовать другую и провести иной эксперимент, изменяя интервалы варьирования факторов.

Процесс обработки данных многофакторного эксперимента включает в себя несколько этапов:

1. Оценка математических ожиданий и дисперсий отклика в отдельных точках факторного пространства или строках плана  $u = 1, 2, \dots, N$  ( $N = 27$ ).

Эти оценки называют построчными средними  $\bar{y}_u$  и построчными дисперсиями  $S_u^2$ .

2. Проверка однородности статистического материала в целях исключения грубых промахов.

3. Проверка воспроизводимости эксперимента.

4. Определение  $b$  – коэффициентов модели.

5. Проверка значимости  $b$ - коэффициентов.

6. Проверка адекватности модели.

Результаты дробного факторного эксперимента, реализованного по плану Хартли, представлены в табл. 4.15 и 4.16.

Таблица 4.15

Результаты статистической обработки данных для параметра  $E$

№	$Y_{cp}$	$S_u^2$	$S_u$	$r_{max}$	$r_{min}$
1	15,7	0,97	0,99	0,47	0,28
2	2,1	0,02	0,15	0,36	0,45
3	3,2	1,97	1,40	0,39	0,43
4	20,9	0,36	0,60	0,43	0,38
5	3,7	0,04	0,20	0,41	0,41
6	18,1	0,10	0,32	0,30	0,47
7	12,4	7,09	2,66	0,35	0,45
8	3,7	0,67	0,82	0,35	0,45
9	0,8	0,42	0,65	0,42	0,40
10	10,3	1,12	1,06	0,31	0,46
11	23,2	4,49	2,12	0,44	0,37
12	5,6	0,04	0,20	0,41	0,41
13	8,4	0,09	0,31	0,36	0,45
14	18,6	0,89	0,95	0,46	0,32
15	5,4	0,30	0,55	0,27	0,47
16	2,9	0,76	0,87	0,47	0,28
17	10,1	4,41	2,10	0,41	0,41
18	17,9	7,75	2,78	0,44	0,37
19	5,2	1,20	1,10	0,46	0,32
20	11,5	4,22	2,06	0,39	0,42
21	13,7	6,07	2,46	0,43	0,38
22	10,4	8,16	2,86	0,39	0,42
23	12,7	0,42	0,65	0,42	0,40
24	10,8	0,42	0,65	0,42	0,40
25	8,9	0,20	0,45	0,39	0,42
26	8,7	0,42	0,65	0,40	0,42
27	6,7	0,37	0,61	0,36	0,45

Таблица 4.16

Результаты статистической обработки данных для параметра П

№	$Y_{cp}$	$S_u^2$	$S_u$	$r_{max}$	$r_{min}$
1	2,61	0,419	0,65	0,46	0,32
2	0,14	0,0002	0,02	0,36	0,45
3	0,29	0,012	0,11	0,44	0,37
4	4,27	0,430	0,66	0,46	0,30
5	0,44	0,002	0,05	0,32	0,46
6	2,83	0,203	0,45	0,42	0,39
7	0,41	0,005	0,07	0,41	0,41
8	0,13	0,002	0,05	0,39	0,42
9	0,05	0,001	0,04	0,39	0,43
10	1,13	0,341	0,58	0,45	0,35
11	3,95	0,495	0,70	0,45	0,35
12	0,91	0,053	0,23	0,41	0,40
13	0,40	0,017	0,13	0,41	0,41
14	3,68	0,865	0,93	0,40	0,41
15	0,45	0,034	0,18	0,47	0,29
16	0,18	0,015	0,12	0,47	0,24
17	1,48	0,144	0,38	0,41	0,41
18	1,90	0,010	0,10	0,41	0,41
19	0,63	0,002	0,05	0,41	0,41
20	0,79	0,010	0,10	0,28	0,47
21	0,48	0,004	0,07	0,40	0,42
22	1,34	0,109	0,33	0,36	0,44
23	1,20	0,011	0,11	0,40	0,41
24	1,32	0,074	0,27	0,46	0,30
25	1,18	0,014	0,12	0,39	0,42
26	1,32	0,001	0,03	0,38	0,43
27	1,44	0,117	0,34	0,43	0,38

*Исключение грубых ошибок.* Грубые ошибки или промахи искажают результаты эксперимента и поэтому должны быть исключены. Имеется несколько способов обнаружения «промахов», но чаще применяется  $r$  – критерий. Для проверки основной статистической гипотезы  $H_0$  определяется расчетное значение  $r$  – критерия:

$$r_{min} = \frac{\bar{y}_u - y_{u1}}{S_u \sqrt{m(m-1)}},$$

если сомнительным считается *наименьшее* в строке значение  $y_{u1}$ ;

$$r_{max} = \frac{\bar{y}_{un} - y_u}{S_u \sqrt{m(m-1)}}.$$

Значение  $r$  – критерия берется из статистических таблиц с учетом степени свободы  $f_u = m - 1$  и уровня значимости  $\alpha$ . Если расчетное значение  $r_{min}$  и  $r_{max}$  окажется меньше табличного, гипотеза  $H_0$  принимается с надежностью вывода  $P = 1 - \alpha$ .

Проверка нулевой гипотезы при  $\alpha = 0,05$  и  $f_u = 2$  показала, что расчетные значения  $r$  – критерия меньше табличных:  $r_{табл} = 1,41$ . Можно утверждать, что экспериментальные данные не содержат сомнительных значений (грубых промахов).

*Проверка воспроизводимости эксперимента.* Цель проверки – определить, являются ли измерения отклика во всех точках (строках) плана равноточными. В данном случае проверка воспроизводимости эксперимента есть проверка однородности выборочных дисперсий  $S_u^2$  (см. табл. 4.12).

Для проверки однородности оценок  $S_u^2$  используем критерий Кохрена, так как число повторных опытов во всех строках постоянно ( $m = 3 = const$ ).

Результаты расчетов для параметра Е показали, что

$$G_p = 0,15 < G_{кр} = 0,29 ;$$

для параметра П -  $G_p = 0,26 < G_{кр} = 0,29$ .

Таким образом, гипотеза об однородности дисперсий принимается.

Выполнение требования однородности построчных дисперсий  $S_u^2$  позволяет определить дисперсию воспроизводимости как среднеарифметическое построчных дисперсий. В этом случае оценка дисперсии воспроизводимости имеет вид

$$S_{вос}^2\{y\} = \frac{1}{N} \sum_{u=1}^N S_u^2\{y\}$$

с числом степеней свободы  $\nu = N(m - 1)$ .

В результате расчетов

для параметра Е –  $S_{\text{вос}}^2\{y\} = 1,98$ ;

для параметра П –  $S_{\text{вос}}^2\{y\} = 0,13$ .

*Оценка коэффициентов уравнения модели.* Коэффициенты уравнения регрессии рассчитываем по формулам:

$$b_0 = a_1 \sum_{u=1}^N (x_{0u} y_u) - a_2 \sum_{i=1}^m \sum_{u=1}^N (x_{iu} x_{iu} y_u) , \quad b_i = a_3 \sum_{u=1}^N (x_{iu} y_u) ,$$

$$b_0 = a_1 \sum_{u=1}^N (x_{iu} x_{iu} y_u) + a_6 \sum_{i=1}^m \sum_{u=1}^N (x_{iu} x_{iu} y_u) - a_2 \sum_{u=1}^N (x_{0u} y_u) ,$$

$$b_0 = a_4 \sum_{u=1}^N (x_{iu} x_{iu} y_u) .$$

Значения коэффициентов см. в табл. 4.17.

Таблица 4.17

Значения коэффициентов  $a_1 - a_6$   
для расчета коэффициентов уравнения

Число опытов	$a_1$	$a_2$	$a_3$	$a_4$	$a_5$	$a_6$
27	0,13805	0,0303	0,05556	0,0625	0,5	-0,09091

Дисперсии коэффициентов уравнения регрессии вычисляются по формулам:

$$S^2(b_0) = a_1 S^2(\bar{y}), \quad S^2(b_i) = a_2 S^2(\bar{y}) ;$$

$$S^2(b_{ij}) = a_3 S^2(\bar{y}), \quad S^2(b_{ii}) = a_4 S^2(\bar{y}) ,$$

где  $a_1 = 0,11111$ ;  $a_2 = 0,02309$ ;  $a_1 = 0,03127$ ;  $a_1 = 0,00315$ .

В результате расчетов уравнение модели для параметра Е приняло вид:

$$\begin{aligned} y_E = & 10,785 + 5,940x_1 - 0,645x_2 - 0,739x_3 - 1,950x_4 + \\ & + 2,100x_5 - 0,813x_1x_2 - 0,638x_1x_3 - 0,888x_1x_4 + \\ & + 1,088x_1x_5 + 0,913x_2x_3 - 0,388x_2x_4 + 1,412x_2x_5 + \\ & + 0,538x_3x_4 - 0,038x_3x_5 + 0,288x_4x_5 + 0,683x_1^2 + b \\ & + 1,733x_2^2 + 0,683x_3^2 - 1,017x_4^2 - 3,167x_5^2 . \end{aligned} \quad (4.3)$$

В результате расчетов уравнение модели для параметра П приняло вид:

$$y_{\Pi} = 1,16 + 1,003x_1 - 0,242x_2 - 0,257x_3 - 0,337x_4 +$$

$$\begin{aligned}
& +0,441x_5 - 0,27x_1x_2 - 0,161x_1x_3 - 0,273x_1x_4 + \\
& +0,348x_1x_5 + 0,091x_2x_3 + 0,201x_2x_4 + 0,272x_2x_5 + \\
& +0,302x_3x_4 + 0,136x_3x_5 + 0,023x_4x_5 + 0,145x_1^2 - b \\
& -0,485x_2^2 + 0,151x_3^2 + 0,131x_4^2 + 0,261x_5^2 .
\end{aligned} \tag{4.4}$$

После определения оценок  $b$  коэффициентов уравнения модели необходимо проверить гипотезу об их значимости, другими словами, проверить, значимо ли полученная оценка отличается от нуля.

Вычисленные значения  $t$  – критерия (табл.4.18 и табл. 4.19) сравнивали с табличным  $t_{\text{табл}} = 2,0$  при заданном уровне значимости  $\alpha = 0,05$  и числа степеней свободы  $\nu = N(m - 1) = 27 \cdot (3 - 1) = 54$ . Результаты расчетов показали, что не все коэффициенты являются статистически значимыми.

Таблица 4.18

Коэффициенты модели и расчетные значения  $t$ - критерия  
для параметра Е

Коэффициент	$b_0$	$b_1$	$b_2$	$b_3$	$b_4$	$b_5$	$b_{12}$
t- критерий	20,2	27,8	3,0	3,5	9,1	9,8	3,3
Коэффициент	$b_{13}$	$b_{14}$	$b_{15}$	$b_{23}$	$b_{24}$	$b_{25}$	$b_{34}$
t- критерий	2,6	3,6	4,4	3,7	1,6	5,7	2,2
Коэффициент	$b_{35}$	$b_{45}$	$b_{11}$	$b_{22}$	$b_{33}$	$b_{44}$	$b_{55}$
t- критерий	0,2	1,2	8,6	21,9	8,6	12,9	40,1

Таблица 4.19

Коэффициенты модели и расчетные значения  $t$ - критерия  
для параметра П

Коэффициент	$b_0$	$b_1$	$b_2$	$b_3$	$b_4$	$b_5$	$b_{12}$
t- критерий	9,8	18,6	4,5	4,8	6,3	8,2	4,3
Коэффициент	$b_{13}$	$b_{14}$	$b_{15}$	$b_{23}$	$b_{24}$	$b_{25}$	$b_{34}$
t- критерий	2,6	4,4	5,6	1,4	3,2	4,3	4,8
Коэффициент	$b_{35}$	$b_{45}$	$b_{11}$	$b_{22}$	$b_{33}$	$b_{44}$	$b_{55}$
t- критерий	2,2	0,4	7,3	24,3	7,5	6,5	13,1

Математическая модель процесса пневмосепарации может быть представлена как уравнение регрессии с факторами и их взаимодействиями, имеющими только значимые коэффициенты уравнения:



$$y_E = 10,785 + 5,940x_1 - 0,645x_2 - 0,739x_3 - 1,950x_4 + \\ + 2,100x_5 - 0,813x_1x_2 - 0,638x_1x_3 - 0,888x_1x_4 + \\ + 1,088x_1x_5 + 0,913x_2x_3 + 1,412x_2x_5 + 0,538x_3x_4 + \\ + 0,683x_1^2 + 1,733x_2^2 + 0,683x_3^2 - 1,017x_4^2 - 3,167x_5^2. \quad (4.5)$$

$$y_{II} = 1,16 + 1,003x_1 - 0,242x_2 - 0,257x_3 - 0,337x_4 + \\ + 0,441x_5 - 0,27x_1x_2 - 0,161x_1x_3 - 0,273x_1x_4 + \\ + 0,348x_1x_5 + 0,201x_2x_3 + 0,272x_2x_5 + 0,302x_3x_4 + \\ + 0,136x_3x_5 + 0,145x_1^2 - 0,485x_2^2 + 0,151x_3^2 - \\ - 0,131x_4^2 - 0,261x_5^2. \quad (4.6)$$

Статистическая незначимость некоторых коэффициентов уравнения может быть вызвана различными и неравнозначными причинами: малым шагом варьирования, статистически незначимым влиянием на эффективность пневмосепарации каких-либо факторов или их взаимодействий, неправильным выбором уровней факторов в области факторного пространства, значительными ошибками эксперимента.

Следующим этапом статистического анализа является проверка гипотезы адекватности модели. Существует несколько методов проверки адекватности модели: использование  $F$  – критерия, критерия серий, критерия инверсий и др.

Проверка нулевой гипотезы об адекватности модели с помощью  $F$  – критерия показала, что для параметра  $E$  расчетное значение  $F$  – критерия равно  $F_p = 1,029$ . Так как  $F_p < F_{\text{табл}} = 2,1$ , то гипотезу об адекватности модели принимаем.

На рис. 4.2 представлены экспериментальные данные и расчетные значения параметра  $E$  по уравнению (4.5), а на рис. 4.3 – отклонения  $(\hat{y}_u - \bar{y}_u)$ .

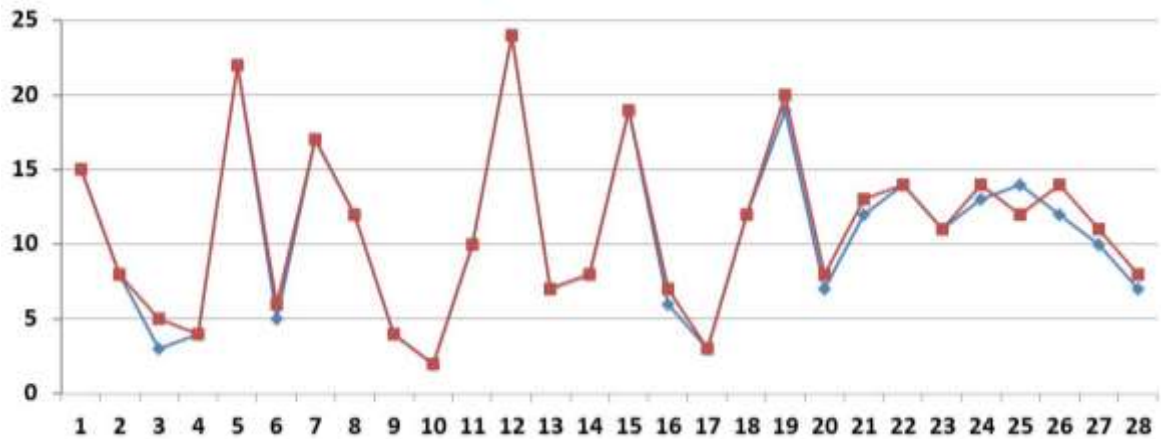


Рис 4.2 Экспериментальные (♦) и расчетные (■) данные параметра E

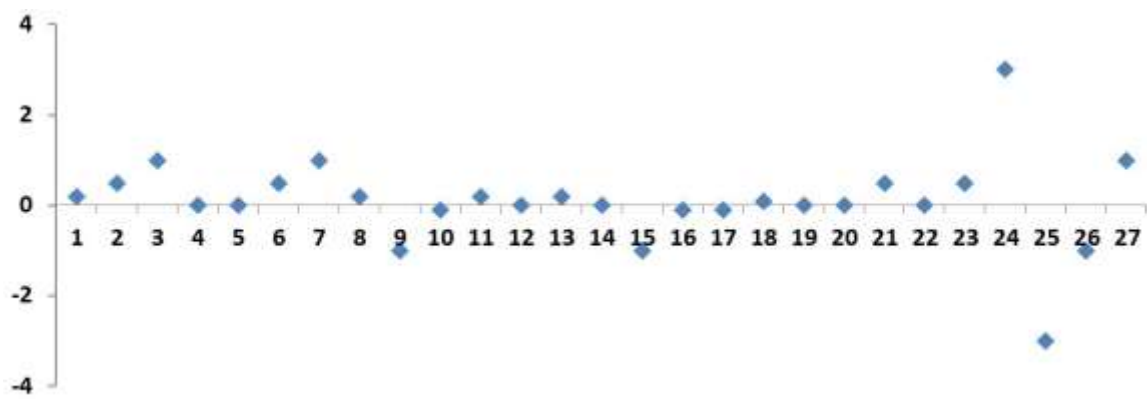


Рис. 4.3. График отклонений  $\hat{y}_u - \bar{y}_u$  для параметра E

Для параметра  $\Pi$  расчетное значение  $F$  – критерия равно  $F_p = 1,79$ . Так как  $F_p < F_{\text{табл}}$ , то гипотезу об адекватности модели принимаем.

На рис. 4.4 представлены экспериментальные данные и расчетные значения параметра  $\Pi$  по уравнению (4.6), а на рис. 4.5 – отклонения ( $\hat{y}_u - \bar{y}_u$ ).

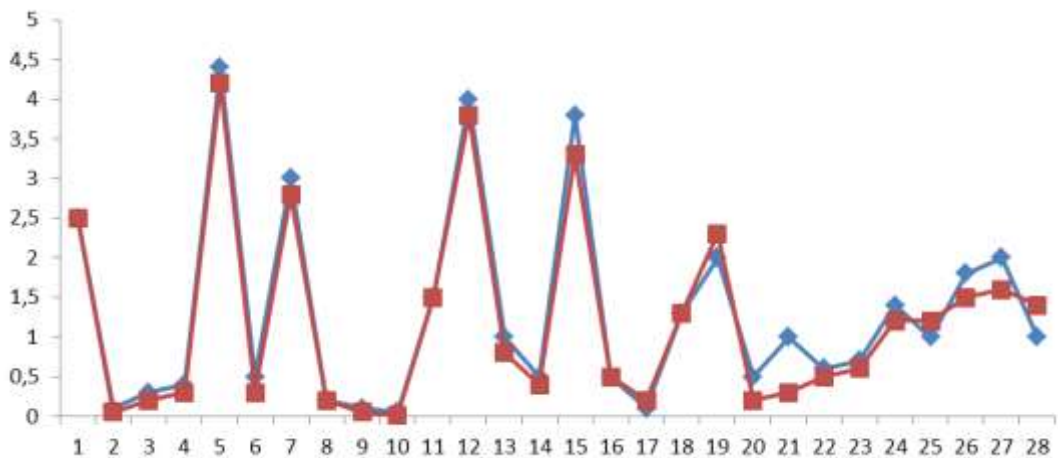


Рис. 4.4 Экспериментальные (♦) и расчетные (■) данные параметра  $\Pi$

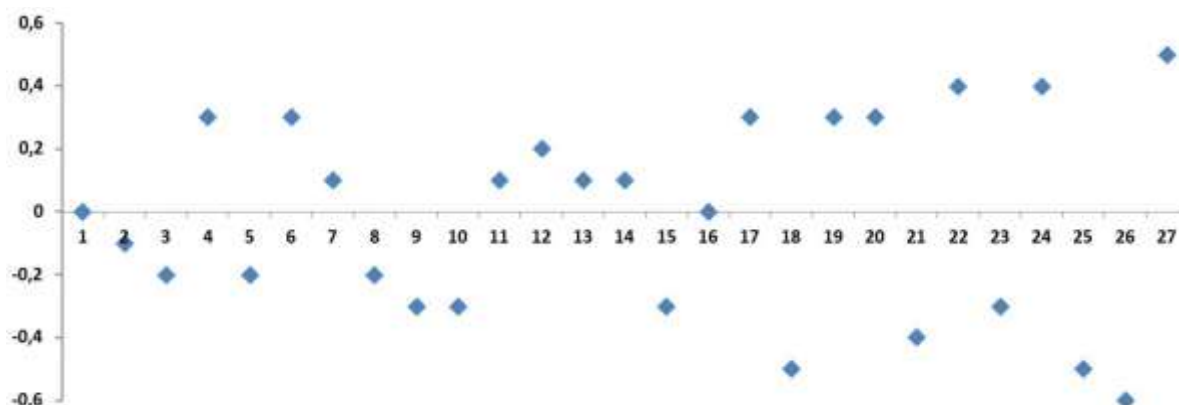


Рис. 4.5. График отклонений  $\hat{y}_i - \bar{y}_i$  для параметра  $\Pi$

Анализ остатков позволяет проверить нулевую гипотезу о статистической независимости остатков. Для проверки данной гипотезы используем критерий серий. Серией называется последовательность однотипных наблюдений, перед и после которой следуют наблюдения противоположного типа или же нет вообще никаких наблюдений. Будем сравнивать остатки с медианой, которая в данном случае равна  $M_{\text{ост}} = -0,209$  и  $M_{\text{пост}} = 0,041$ . Остатки, превышающие медиану, классифицируем как (+), а остатки меньше медианы, классифицируем как (-).

Для параметра E:

- + + + - + + + - - + - + - - - - + - - + - + + - + +  
 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16

последовательность из 27 остатков имеет  $r_E = 16$  серий;

для параметра  $\Pi$ :

+ - - + - + + - - - - + + + - - + - + + - + - + - - +  
 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17

последовательность из 27 остатков имеет  $r_{\Pi} = 17$  серий. Нулевая гипотеза принимается, если расчетное значение критерия серий находится в пределах:

$$r_{N/2; 1-\alpha/2} < r < r_{N/2; \alpha/2},$$

где  $N = 27$  – число наблюдений;  $\alpha = 0,05$  – уровень значимости.

Табулированные значения критерия серий:

$$r_{13; 0,975} = 8,$$

$$r_{13; 0,025} = 19,$$

$$r_{14; 0,975} = 9 ,$$

$$r_{14; 0,025} = 20 .$$

Так как  $r_E = 16$  и  $r_{\Pi} = 17$  попадают в интервал принятия гипотезы, то нет оснований сомневаться в независимости наблюдений. Следовательно, отклонения значений, рассчитанных по модели, от экспериментальных носят случайный характер.

Для построения математической модели процесса пневмосепарации в натуральных переменных необходимо, используя формулы кодирования, сделать соответствующие преобразования. После раскодирования уравнение модели пневмосепарации (для параметра  $E$ ) имеет вид:

$$\begin{aligned} E = & 51,88 + 0,62V - 21,87d - 1,16a + 3,63q + \\ & + 0,02h - 0,08Vd - 0,006Va - 4Vq + \\ & + 0,001Vh - 0,09da + 0,009dh + 0,03aq + \\ & + 0,007V^2 + 1,73d^2 + 0,007a^2 - 0,25q^2 - 0,00014h^2 . \end{aligned}$$

После раскодирования уравнение модели для параметра  $\Pi$  имеет вид:  $\Pi = -0,2125 + 0,17225V - 4,88995d - 0,24785a - 1,2165q -$

$$\begin{aligned} & - 5h - 0,027Vd - 0,00161Va - 0,01365Vq + \\ & + 0,000232Vh + 0,1005dq + 0,001813dh + 0,0151aq + \\ & + 0,000091ah + 0,00145V^2 - 0,485d^2 + 0,00151a^2 + \\ & + 0,03275q^2 - 0,0000116h^2 . \end{aligned}$$

При нахождении оптимальных значений использовалось стандартное программное обеспечение пакета Excel. В результате оптимизации получены оптимальные значения исследуемых параметров (табл. 4.20), при которых функция  $E$  принимает максимальное значение  $E = 26,2$ .

Таблица 4.20

Оптимальные значения параметров

| Параметр                                     | Значение |
|--|----------|
| $V$ – скорость выхода струи воздуха, м/с     | 70       |
| $d$ – диаметр отверстия, мм                  | 4        |
| $a$ – расстояние между отверстиями, мм       | 50       |
| $q$ – подача хлебной массы в молотилку, кг/с | 5        |
| $h$ – толщина вороха, мм                     | 392      |

## Список использованных источников

1. Мельников С.В. Механизация и автоматизация животноводческих ферм / С.В. Мельников. – М.: Колос, 1978. – 280 с.
2. Адлер Ю.П. Планирование эксперимента при поиске оптимальных условий / Ю.П. Адлер, Е.В. Маркова, Ю.В. Грановский. – М.: Наука, 1976. – 280 с.
3. Кетков Ю.П. Программирование на Бейсике / Ю.П. Кетков. – М.: Статистика, 1970. – 162 с.
4. Львовский Е.Н. Статистические методы построения эмпирических формул / Е.Н. Львовский. – М.: Высш. шк., 1982. – 240 с.
5. Вычислительная техника в инженерных и экономических расчетах: учеб. для вузов / Под ред. А.В. Петрова. – 2-е изд., перераб. и доп. – М.: Высш. шк., 1984. – 180 с.
6. Налимов В.В. Применение математической статистики при анализе вещества / В.В. Налимов. – М.: Физматгиз, 1960. – 430 с.
7. Федоров В.В. Теория оптимального эксперимента / В.В. Федоров. – М.: Наука, 1971. – 312 с.
8. Математическая статистика: учеб. / В.М. Иванова, В.Н. Калинина, Л.А. Нешумова и др. – М.: Высш. шк., 1981. – 371 с.
9. Мухачёв В.А. Планирование и обработка результатов эксперимента: учеб. пособие / В.А. Мухачёв / Томский государственный университет систем управления и радиоэлектроники. – Томск, 2007. – 118 с.
10. Реброва И.А. Планирование эксперимента: учеб. пособие / И.А. Реброва. – Омск: СиБАДИ, 2010. – 105 с.
11. Макаричев Ю.А. Методы планирования эксперимента и обработки данных: учеб. пособие / Ю.А. Макаричев, Ю.Н. Иванников / Самар. гос. техн. ун-т. – Самара, 2016. – 131 с.

## ПРИЛОЖЕНИЕ 1

### Равномерно распределенные случайные числа в интервале от 0 до 99

|    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| 10 | 09 | 73 | 25 | 33 | 76 | 52 | 01 | 35 | 86 | 34 | 67 | 35 | 48 | 76 | 80 | 95 |
| 90 | 91 | 17 | 39 | 29 | 27 | 49 | 49 | 37 | 54 | 20 | 48 | 05 | 64 | 89 | 47 | 42 |
| 96 | 24 | 80 | 52 | 40 | 37 | 20 | 63 | 61 | 04 | 02 | 00 | 82 | 29 | 16 | 65 | 08 |
| 42 | 26 | 89 | 53 | 19 | 64 | 50 | 93 | 03 | 23 | 20 | 90 | 25 | 60 | 15 | 95 | 33 |
| 47 | 64 | 35 | 08 | 03 | 36 | 06 | 99 | 01 | 90 | 25 | 29 | 09 | 37 | 67 | 07 | 15 |
| 38 | 31 | 13 | 11 | 65 | 88 | 67 | 67 | 43 | 97 | 04 | 43 | 62 | 76 | 59 | 12 | 80 |
| 79 | 99 | 70 | 80 | 15 | 73 | 61 | 47 | 64 | 03 | 23 | 66 | 53 | 98 | 95 | 11 | 68 |
| 77 | 12 | 17 | 14 | 68 | 33 | 66 | 06 | 57 | 47 | 17 | 34 | 07 | 27 | 68 | 50 | 36 |
| 69 | 73 | 61 | 70 | 65 | 81 | 33 | 98 | 85 | 11 | 19 | 92 | 91 | 70 | 31 | 06 | 01 |
| 08 | 05 | 45 | 57 | 18 | 24 | 06 | 35 | 30 | 34 | 26 | 14 | 86 | 79 | 90 | 74 | 39 |
| 23 | 40 | 30 | 97 | 32 | 85 | 26 | 97 | 76 | 02 | 02 | 05 | 16 | 56 | 92 | 68 | 66 |
| 57 | 48 | 18 | 73 | 05 | 38 | 52 | 47 | 18 | 62 | 38 | 85 | 79 | 63 | 57 | 33 | 21 |
| 35 | 05 | 32 | 54 | 70 | 48 | 90 | 55 | 35 | 75 | 48 | 28 | 46 | 82 | 87 | 09 | 83 |
| 49 | 12 | 56 | 24 | 73 | 79 | 64 | 57 | 53 | 03 | 52 | 96 | 47 | 78 | 35 | 80 | 83 |
| 42 | 82 | 60 | 93 | 52 | 03 | 44 | 35 | 27 | 38 | 84 | 35 | 98 | 52 | 01 | 77 | 67 |
| 14 | 90 | 56 | 86 | 07 | 22 | 10 | 94 | 05 | 58 | 60 | 97 | 09 | 34 | 33 | 50 | 50 |
| 07 | 37 | 98 | 11 | 80 | 50 | 54 | 31 | 39 | 80 | 82 | 77 | 32 | 50 | 72 | 56 | 82 |
| 48 | 29 | 40 | 52 | 42 | 01 | 52 | 77 | 56 | 78 | 51 | 83 | 45 | 20 | 96 | 34 | 06 |
| 28 | 89 | 80 | 83 | 18 | 74 | 67 | 00 | 78 | 18 | 47 | 54 | 06 | 10 | 68 | 71 | 17 |
| 78 | 17 | 88 | 68 | 54 | 02 | 00 | 86 | 50 | 75 | 84 | 01 | 36 | 76 | 66 | 79 | 51 |
| 90 | 36 | 47 | 64 | 93 | 29 | 60 | 91 | 10 | 62 | 99 | 59 | 46 | 73 | 48 | 37 | 51 |
| 79 | 49 | 69 | 91 | 82 | 60 | 89 | 28 | 93 | 78 | 56 | 13 | 68 | 23 | 47 | 83 | 41 |
| 13 | 65 | 48 | 11 | 76 | 74 | 17 | 46 | 85 | 09 | 50 | 58 | 04 | 77 | 69 | 74 | 73 |
| 03 | 95 | 71 | 86 | 40 | 21 | 81 | 65 | 44 | 80 | 12 | 43 | 56 | 35 | 17 | 72 | 70 |
| 70 | 15 | 45 | 31 | 82 | 23 | 74 | 21 | 11 | 57 | 82 | 53 | 14 | 38 | 55 | 37 | 63 |
| 74 | 35 | 09 | 98 | 17 | 77 | 45 | 27 | 72 | 14 | 43 | 23 | 60 | 02 | 10 | 45 | 52 |
| 16 | 42 | 37 | 96 | 28 | 60 | 26 | 55 | 69 | 91 | 62 | 68 | 03 | 66 | 25 | 22 | 91 |
| 48 | 36 | 93 | 68 | 72 | 03 | 76 | 62 | 11 | 39 | 90 | 94 | 40 | 05 | 64 | 18 | 09 |
| 89 | 32 | 05 | 05 | 14 | 22 | 56 | 85 | 14 | 46 | 42 | 75 | 67 | 88 | 96 | 29 | 77 |
| 88 | 22 | 54 | 38 | 21 | 45 | 98 | 91 | 49 | 91 | 45 | 23 | 68 | 47 | 92 | 76 | 86 |
| 46 | 16 | 28 | 35 | 54 | 94 | 75 | 08 | 99 | 23 | 37 | 08 | 92 | 00 | 48 | 80 | 33 |
| 69 | 45 | 98 | 26 | 94 | 03 | 68 | 58 | 70 | 29 | 73 | 41 | 35 | 53 | 14 | 03 | 33 |

## ПРИЛОЖЕНИЕ 2

Зависимость  $t_{\text{табл}}$  от числа  $n$  при  $\alpha = 0,05$

| $n$               | 4    | 8    | 12   | 16   | 24   | 32   | 48   | 64  |
|-------------------|------|------|------|------|------|------|------|-----|
| $t_{\text{табл}}$ | 2,78 | 2,31 | 2,18 | 2,12 | 2,06 | 2,04 | 2,01 | 2,0 |

Зависимость  $F_{\text{табл}}$  от числа степеней свободы числителя ( $\nu_1$ )  
и знаменателя ( $\nu_2$ ) при  $\alpha = 0,05$

| $\nu_2$ | $\nu_1$ |       |       |       |       |       |       |       |
|---------|---------|-------|-------|-------|-------|-------|-------|-------|
|         | 4       | 8     | 12    | 16    | 24    | 30    | 48    | 60    |
| 1       | 225     | 239   | 244   | 246   | 249   | 250,2 | 252   | 253   |
| 2       | 19,12   | 19,37 | 19,41 | 19,43 | 19,45 | 19,46 | 19,47 | 19,48 |
| 3       | 9,12    | 8,84  | 8,74  | 8,69  | 8,64  | 8,62  | 8,59  | 8,57  |
| 4       | 6,39    | 6,04  | 5,91  | 5,84  | 5,77  | 5,74  | 5,70  | 5,68  |
| 5       | 5,19    | 4,82  | 4,68  | 4,60  | 4,53  | 4,50  | 4,45  | 4,42  |
| 6       | 4,53    | 4,15  | 4,00  | 3,92  | 3,84  | 3,81  | 3,76  | 3,72  |
| 7       | 4,12    | 3,73  | 3,57  | 3,49  | 3,41  | 3,38  | 3,32  | 3,28  |
| 8       | 3,84    | 3,44  | 3,28  | 3,20  | 3,12  | 3,08  | 3,02  | 3,00  |
| 9       | 3,63    | 3,23  | 3,07  | 2,98  | 2,90  | 2,86  | 2,80  | 2,77  |
| 10      | 3,48    | 3,07  | 2,91  | 2,82  | 2,74  | 2,70  | 2,66  | 2,61  |

Учебное издание

**Борисова** Людмила Викторовна

**Димитров** Валерий Петрович

**Зубрилина** Елена Михайловна

**ОСНОВЫ ТЕОРИИ ЭКСПЕРИМЕНТА.  
ПОСТРОЕНИЕ МАТЕМАТИЧЕСКИХ МОДЕЛЕЙ**

---

Подписано в печать \_\_\_\_\_ 2021.

Формат 60×84/16. Объем 6,6 усл. п.л.

Тираж 100 экз. Заказ №145

---

ООО «ДГТУ-ПРИНТ»

Адрес полиграфического предприятия:

344000, Ростов-на-Дону, пл. Гагарина, 1