МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное бюджетное образовательное
учреждение высшего образования

«Уральский государственный лесотехнический университет»
(УГЛТУ)

МАТЕМАТИЧЕСКОЕ ПЛАНИРОВАНИЕ ЭКСПЕРИМЕНТОВ И АНАЛИЗ ИХ РЕЗУЛЬТАТОВ С ПРИМЕНЕНИЕМ КОМПЬЮТЕРНЫХ ПРОГРАММ

Учебное пособие

Екатеринбург
2023

УДК 001.891.5:519.242(075.8)
ББК 35.10в643я73
М34

Рецензенты:

кафедра химии и процессов горения УрИ ГПС МЧС России, канд. хим. наук, доцент *А. В. Кокшаров*;

Н. П. Кулик, старший научный сотрудник Института высокотемпературной электрохимии Уральского отделения РАН, канд. хим. наук

Авторы: В. В. Глухих, А. Е. Шкуро, А. В. Артемов,
О. Ф. Шишлов, П. С. Кривоногов

М34 **Математическое планирование экспериментов и анализ их результатов с применением компьютерных программ** : учебное пособие / В. В. Глухих, А. Е. Шкуро, А. В. Артемов [и др.] ; Министерство науки и высшего образования Российской Федерации, Уральский государственный лесотехнический университет. – Екатеринбург : УГЛТУ, 2023. – 9 Мб. – 1 электрон. опт. диск (CD-ROM). – Текст : электронный.

ISBN 978-5-94984-864-7

Пособие предназначено для обучающихся по направлениям подготовки 18.03.01 и 18.04.01 «Химическая технология», магистров и аспирантов.

В пособии рассмотрены вопросы математического планирования экспериментов и анализа их результатов с применением Microsoft Excel, Statistica и других компьютерных программ. Приведены примеры составления планов эксперимента и решения оптимизационных задач.

Пособие предлагается использовать на практических занятиях по дисциплинам «Прикладные научные исследования» и «Решение оптимизационных задач», а также при подготовке выпускных квалификационных работ.

Издается по решению редакционно-издательского совета Уральского государственного лесотехнического университета.

УДК 001.891.5:519.242(075.8)
ББК 35.10в643я73

Мин. системные требования : IBM Intel Celeron 1,3 ГГц ; Microsoft Windows XP SP3 ; Видеосистема Intel HD Graphics ; дисковод, мышь.

ISBN 978-5-94984-864-7

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ

Учебное пособие является дополнением к учебнику В. В. Глухих «Прикладные научные исследования» [1] и предназначено для формирования умений обучающихся (бакалавров, магистров, аспирантов) применять на практике свои теоретические знания математического планирования экспериментов и владения современными компьютерными программами для математической обработки и статистического анализа результатов экспериментальных исследований.

Экспериментальные исследования (эксперимент) в различных отраслях науки состоят из следующих обязательных элементов, выполняемых в такой последовательности:

1) формулировка целей эксперимента;
2) планирование эксперимента;
3) реализация эксперимента;
4) анализ результатов эксперимента.

Планирование эксперимента подразделяют на две части: методическую и организационную.

В методической части составляют и анализируют план и методику проведения эксперимента, выбирают средства измерения, экспериментальные образцы, материалы, установки, исследователей.

В организационной части решают вопросы материально-технического обеспечения эксперимента (подготовку к работе средств измерения, установок, исследователей, компьютерные программы для обработки и анализа данных и др.).

Выбор плана эксперимента зависит:

– от целей и задач научно-исследовательской работы (НИР) и эксперимента;

– методов анализа результатов эксперимента, которые планируется применить;

– объема имеющихся материальных ресурсов (финансовых, человеческих и др.);

– временных ресурсов (ограничений продолжительности научных исследований) и др.

Основной целью применения математических методов планирования эксперимента является получение математического описания поведения объекта (математических моделей объекта) при изменении параметров эксперимента (входных факторов) с предполагаемым их влиянием на свойства объекта даже при отсутствии сведений о фундаментальных закономерностях этого поведения.

Математические модели объектов, получаемые с помощью методов математического планирования эксперимента, называют экспериментально-статистическими. Эти модели применимы на практике с принятой при их построении вероятностью только в той области изменения входных факторов, которая использовалась в эксперименте для построения экспериментально-статистических моделей.

Математическое планирование эксперимента и получение экспериментально-статистических моделей объектов позволяет решать следующие задачи:

– дает информацию о влиянии изменяемых в эксперименте входных факторов на свойства (отклики) объекта;

– позволяет количественно определять и прогнозировать значения откликов объекта при конкретных значениях входных факторов;

– может помочь при решении оптимизационных задач для откликов объекта.

К достоинствам математических планов эксперимента относят возможности сокращения при реализации исследовательских работ времени и материальных средств на их выполнение.

Недостатком математических планов эксперимента является высокая вероятность применимости на практике экспериментально-статистических моделей объектов только в той области изменений входных факторов, которая была исследована в этом эксперименте.

1. МАТЕМАТИЧЕСКОЕ ПЛАНИРОВАНИЕ ЭКСПЕРИМЕНТОВ ДЛЯ ПОЛУЧЕНИЯ ЭКСПЕРИМЕНТАЛЬНО-СТАТИСТИЧЕСКИХ МОДЕЛЕЙ ОБЪЕКТА

Прежде чем приступить к математическому планированию эксперимента, необходимо убедиться в том, что опыты воспроизводимы, т. е. рассчитать и проанализировать ошибки и погрешности измерений. Для этой цели проводят несколько серий параллельных опытов в рассматриваемой области изменения входных факторов с целью определения воспроизводимости опытов в эксперименте.

По мнению С. А. Семенова [2], если никакими способами невозможно достигнуть воспроизводимости опытов в эксперименте, то стандартные математические методы планирования к такому эксперименту применять нельзя. В его учебном пособии приведены конкретные примеры оценки воспроизводимости опытов в химическом эксперименте и практические рекомендации по улучшению воспроизводимости опытов.

При проведении эксперимента по математическому плану для объектов с неизвестным или малоизученным алгоритмом их поведения используется принцип «черного ящика». Суть этого принципа заключается в том, что исследователь, не зная об истинных закономерностях поведения объекта, описывает его с помощью статистических математических моделей с заданной доверительной вероятностью, выбранной исследователем.

Образно говоря, «ударяя» по исследуемому объекту изменением входных факторов Z_i в ходе эксперимента и измеряя его отклики (свойства) на эти «удары» Y_j на фоне действия случайных факторов, получают после его завершения экспериментально-статистическую математическую модель (зависимость U), пригодную для описания и прогнозирования поведения объекта с заданной вероятностью P:

$$Y_j = U_j(Z_i).$$

Выбор математического плана эксперимента зависит от того, какой вид зависимости φ желают получить: качественный или количественный.

Зависимость U является качественной, если она выражается словами, например: «Z_i влияет на Y_j», «увеличение Z_i уменьшает значение Y_j» и др.

Зависимость U является количественной, если она представляет собой математическое выражение.

Так как результаты измерений значений Z_i и Y_j являются случайными величинами, то для установления зависимости U необходимо использовать соответствующие методы математической статистики.

Для получения экспериментально-статистических моделей объекта составляют математические планы полных факторных экспериментов (ПФЭ) и дробных факторных экспериментов (ДФЭ).

Планы дробных факторных экспериментов по сравнению с планами полных факторных экспериментов являются менее затратными по времени и материальным ресурсам при реализации исследовательских работ из-за меньшего числа опытов в эксперименте. Но при их применении может быть меньшая вероятность подтверждения на практике полученных закономерностей откликов объекта по сравнению с таковой закономерностей экспериментально-статистических моделей объекта, полученных по результатам ПФЭ.

Математические планы экспериментов делятся на различные виды:

– планы для дисперсионного анализа результатов экспериментов (для получения качественных зависимостей);

– планы для регрессионного анализа результатов экспериментов (для получения количественных зависимостей);

– планы для решения оптимизационных задач;

– планы для изучения зависимости свойств объекта от соотношения его компонентов;

– робастные планы;

– и др.

При составлении математических планов экспериментов используют различные принципы математической комбинаторики для формализованных (кодированных) значений входных факторов x_i с использованием, например, следующих терминов и их обозначений:

- N – общее число опытов в эксперименте;
- k – число входных факторов;
- m – число уровней входных факторов;
- n – число повторений каждого опыта в эксперименте.

Соотношения кодированных значений входных факторов x_i с их натуральными значениями Z_i описываются следующими математическими формулами:

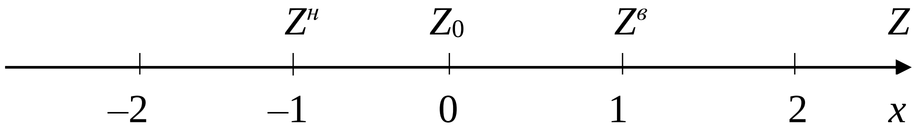
$$x_i = \frac{Z_i^0}{\Delta Z_i}; \quad Z_i^0 = \frac{Z_i^{\text{в}} + Z_i^{\text{н}}}{2}; \quad \Delta Z_j = \frac{Z_i^{\text{в}} - Z_i^{\text{н}}}{2},$$

где Z_i^0 – натуральное значение фактора в центре (середине) выбранной (заданной) области изменения (варьирования) фактора,

$Z_i^{\text{в}}$ и $Z_i^{\text{н}}$ – соответственно верхняя и нижняя границы факторной области, которой будут соответствовать кодированные значения факторов, равные соответственно +1 и –1;

ΔZ_j – шаг варьирования фактора.

В соответствии с этими формулами натуральному значению фактора Z соответствует $x = 0$ на числовой оси кодированных значений:



Переход от кодированных значений факторов к натуральным осуществляют по формуле

$$Z_j = Z_j^0 + x_j \Delta Z_j.$$

Математические планы экспериментов могут иметь названия:

- по фамилиям авторов планов (при этом один и тот же план может обозначаться фамилиями разных авторов);
- математическим формулам расчета общего числа опытов плана;
- именам вида или группы планов.

По количеству значений входных факторов (уровней факторов) математические планы экспериментов делятся на двухуровневые и многоуровневые (число уровней больше двух), а по доле комбинаций уровней входных факторов – планы полных факторных экспериментов (ПФЭ) и дробных факторных экспериментов (ДФЭ).

Планы полных многофакторных экспериментов (ПФЭ) предусматривают их проведение при всех возможных комбинациях уровней всех входных факторов. При многофакторном эксперименте одновременно изменяются три и более факторов.

При математическом планировании эксперимента выбирают такие принципы математической комбинаторики, чтобы исключить в плане дублирование опытов с одинаковым сочетанием значений входных факторов.

Общее число опытов (без их повторения) в эксперименте по плану ПФЭ рассчитывается по формулам

$N = m^k$;

$N = 2^k$ (для двухуровневого плана эксперимента);

$N = 3^k$ (для трехуровневого плана эксперимента).

Общее число опытов при их повторении в эксперименте по плану ПФЭ рассчитывается по формулам

$N = n2^k$ (для двухуровневого плана эксперимента);

$N = n3^k$ (для трехуровневого плана эксперимента).

В табл. 1 и 2 приведены примеры планов ПФЭ двухфакторного двухуровневого и двухфакторного трехуровневого экспериментов без повторения опытов по одному из принципов математической комбинаторики с кодированными значениями факторов. Именем этих планов ПФЭ является формула расчета их числа опытов.

При повторении опытов общее их число увеличивается и повышается вероятность подтверждения закономерностей экспериментально-статистических моделей объекта на практике.

Для статистического анализа результатов эксперимента в пакетах прикладных программ можно использовать вышеприведенные таблицы с натуральными значениями входных факторов и среднеарифметическими значениями откликов объекта.

Для анализа результатов ПФЭ можно применять дисперсионный и регрессионный анализ.

Таблица 1

План эксперимента 2^2

Номер опыта	Входные факторы				Отклики объекта	
	Кодированные значения		Натуральные значения		Y_1	Y_2
	x_1	x_2	Z_1	Z_2		
1	1	1				
2	1	–1				
3	–1	1				
4	–1	–1				

Таблица 2

План эксперимента 3^2

Номер опыта	Входные факторы				Отклики объекта	
	Кодированные значения		Натуральные значения		Y_1	Y_2
	x_1	x_2	Z_1	Z_2		
1	1	1				
2	1	0				
3	1	–1				
4	0	1				
5	0	0				
6	0	–1				
7	–1	1				
8	–1	0				
9	–1	–1				

Планы ДФЭ из-за значительно меньшего числа опытов по сравнению с планами ПФЭ широко применяются в промышленных экспериментах, а также при отсеивающих лабораторных экспериментах, когда необходимо изучить достаточно большое число факторов при небольшом числе опытов и определить те факторы, которые оказывают наиболее сильное влияние на свойство Y.

При необходимости (желании) уменьшить число опытов в эксперименте по сравнению с таковым плана ПФЭ математические планы ДФЭ составляют в зависимости от предпочитаемого метода статистического анализа его результатов.

Математическое планирование эксперимента – это постоянно совершенствующийся и обновляющийся раздел математики. Выдержки из истории и теории планирования эксперимента описаны в книгах [3, 4]. Важными показателями качества составленных математических планов эксперимента является их ортогональность и оптимальность [3–5].

Теоретические вопросы оценки оптимальности составленных планов ДФЭ, в том числе с помощью программы STATGRAPHICS, приведены в книгах [3, 4].

1.1. Планы дробных факторных экспериментов для дисперсионного анализа их результатов

Сокращение числа опытов в эксперименте, т. е. переход от ПФЭ к ДФЭ, всегда приводит к снижению точности дисперсионного анализа результатов эксперимента.

Существуют различные принципы составления и типы планов ДФЭ для дисперсионного анализа результатов эксперимента:

– планы, составленные по принципу латинских квадратов и кубов;

– планы Плакетта – Бермана;

– другие планы.

Одними из самых экономичных по числу опытов и эффективных для многофакторного дисперсионного анализа из известных планов ДФЭ являются планы Плакетта – Бермана, удовлетворяющие критериям ортогональности и D-оптимальности [3, 4].

Применяются различные способы математической комбинаторики при построении планов Плакетта – Бермана. Рассмотрим одни из них.

Пример 1. Необходимо составить план эксперимента с минимальным числом опытов без их повторений для изучения методом дисперсионного анализа качественной зависимости свойства объекта Y от изменений десяти факторов ($k = 10$), каждый из которых имеет два уровня ($m = 2$).

Выполним расчеты необходимого числа опытов для различных видов планов эксперимента.

Для проведения ПФЭ необходимо будет выполнить следующее число опытов $N_{ПФЭ}$:

$$N_{ПФЭ} = m^k = 2^{10} = 1024.$$

Из известных ортогональных и D-оптимальных планов ДФЭ рассмотрим возможное число опытов в двухуровневом эксперименте по планам, составленным по принципу дробных реплик $N_{ДР}$ и планам Плакетта – Бермана $N_{ПБ}$:

$$N_{ДР} = 2^{k-a} = 2^{10-a}; \; N_{ПБ} = 4b,$$

где $a = 1, 2, 3, ..., 10$ (соответственно $N_{ДР}$ равно 512, 256, 128, 64, 32, 16, 8, 4, 2), а b равно 1, 2, 3, 4 (соответственно $N_{ПБ}$ равно 4, 8, 12, 16).

Из таких двухуровневых планов можно выбирать только «насыщенные» планы, т. е. те, для которых выполняется соотношение

$$N \geq k+1 \geq 10+1 \geq 11.$$

Требованиям этого соотношения и минимального числа опытов лучше всех удовлетворяет план Плакетта – Бермана с $N_{ПБ} = 12$.

Построим такой план с кодированными значениями факторов, обозначая знаком «+1» максимальное значение фактора, а знаком «–1» его минимальное значение.

При построении данного плана (табл. 3) в соответствии с одним из вариантов способов математической комбинаторики в ячейки последнего опыта (№ 12) внесем кодированное значение «–1» для всех факторов [5].

Таблица 3

План Плакетта – Бермана с 12 опытами без их повторений

Номер опыта i	Кодированные значения факторов												Y
	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}*	x_{12}*	
1	+1	–1	+1	–1	–1	–1	+1	+1	+1	–1	+1	+1	
2	+1	+1	–1	+1	–1	–1	–1	+1	+1	+1	–1	+1	
3	–1	+1	+1	–1	+1	–1	–1	–1	+1	+1	+1	–1	
4	+1	–1	+1	+1	–1	+1	–1	–1	–1	+1	+1	+1	
5	+1	+1	–1	+1	+1	–1	+1	–1	–1	–1	+1	+1	
6	+1	+1	+1	–1	+1	+1	–1	+1	–1	–1	–1	+1	
7	–1	+1	+1	+1	–1	+1	+1	–1	+1	–1	–1	–1	
8	–1	–1	+1	+1	+1	–1	+1	+1	–1	+1	–1	–1	
9	–1	–1	–1	+1	+1	+1	–1	+1	+1	–1	+1	–1	
10	+1	–1	–1	–1	+1	+1	+1	–1	+1	+1	–1	+1	
11	–1	+1	–1	–1	–1	+1	+1	+1	–1	+1	+1	–1	
12	–1	–1	–1	–1	–1	–1	–1	–1	–1	–1	–1	–1	

---

* Фиктивные факторы, которые служат для оценки случайных ошибок при анализе результатов данного эксперимента [5].

Затем в столбце плана для x_1 в одиннадцати незаполненных ячейках расположим шесть ($N/2$) знаков «+1» и пять ($N/2–1$) знаков «–1». При заполнении ячеек следующего столбца для x_2 воспользуемся комбинацией значений в левом соседнем столбце x_1: поместим в первую ячейку столбца x_2 значение x_1 в предпоследней (одиннадцатой) ячейке, а в оставшиеся пустые ячейки столбца x_2 – значения x_1 в десяти верхних ячейках. Аналогичным способом математической комбинаторики заполним столбцы для всех факторов.

Правильность построенного плана Плакетта – Бермана по использованному способу математической комбинаторики подтверждается двумя признаками [5]:

– диагональным расположением одинаковых знаков в ячейках плана;

– равенством количества знаков «+1» и знаков «–1» (по шесть) в каждом столбце плана.

План с натуральными значениями факторов получим из плана с кодированными значениями путем замены знаков «+1» и «–1» на соответствующие им максимальные и минимальные натуральные значения для каждого фактора.

1.2. Планы дробных факторных экспериментов для регрессионного анализа их результатов

Планы Плакетта – Бермана применяют также и для регрессионного анализа результатов эксперимента, но только для получения экспериментально-статистических моделей объекта в виде линейных уравнений регрессии, которые не содержат в своем составе эффекты взаимодействия входных факторов [5].

Планы дробных факторных экспериментов, позволяющие получать только линейные уравнения регрессии, называют планами первого порядка, а планы, позволяющие получать и линейные и нелинейные уравнения регрессии, – планами второго порядка.

Рассмотрим применение методов математического планирования эксперимента для проведения регрессионного анализа полученных экспериментальных данных.

Регрессионный анализ (РА) – метод математической статистики, который позволяет выявлять приближенные количественные зависимости свойств объекта Y_j от натуральных (зависимость U) или кодированных значений (зависимость φ) входных факторов, оказывающих влияние на эти свойства. Полученные при регрессионном анализе результатов эксперимента уравнения регрессии являются экспериментально-статистическими моделями исследованного объекта в форме приближенных зависимостей, выраженных в виде конкретных математических функций:

$$Y_j = U_j \ (Z_1, Z_2, \ \ldots, Z_i, \ldots, Z_k),$$

$$Y_j = j_j \ (x_1, x_2, \ \ldots, x_i, \ldots, x_k).$$

Проводить регрессионный анализ можно **только для количественных значений *Y* и *Z*** **Z_i**.

При регрессионном анализе решают две основные задачи.

1. Ищут с помощью метода приближения уравнение регрессии, наиболее точно описывающее по результатам измерения свойств объекта истинную зависимость этих свойств при различных натуральных значениях входных факторов:

$$Y = U(Z_1,\ Z_2,\ ...,\ Z_i,\ ...,\ Z_k) + \theta + \varepsilon,$$

где θ – ошибки, связанные с несовпадением уравнения регрессии с истиной зависимостью, ε – случайные ошибки эксперимента.

2. Оценивают суммарные ошибки ($\theta + \varepsilon$).

Порядок проведения регрессионного анализа (его тип) зависит от плана эксперимента. Различают классический регрессионный анализ (КРА) и регрессионный анализ при математическом планировании эксперимента (РАМПЭ).

Общим требованием к планированию любого эксперимента для проведения КРА и РАМПЭ является выполнение условия для числа уровней каждого входного фактора $m_l \geq 2$.

С помощью метода «черного ящика» получают экспериментально-статистические модели объектов в виде полиномов различной степени.

Известно, что любую функцию (в том числе φ и *U*) можно разложить в ряд Тейлора и представить в виде конкретного полинома определенной степени (конечного отрезка ряда Тейлора) вида, например:

$$Y = \varphi_2(x) = b_0 + \sum_{j=1}^{k} b_j x_j + \sum_{u,\,j=1(u \neq j)}^{k} b_{uj} x_u x_j + \sum_{j=1}^{k} b_{jj} x_j^2 + \sum_{g,u,\,j=1(g \neq u \neq j)}^{k} b_{guj} x_g x_u x_j + ...,$$

или полинома другой степени, где *b* – выборочные коэффициенты ряда Тейлора для кодированных значений входных факторов.

По результатам эксперимента можно определить вид полинома только с выборочными коэффициентами, которые характеризуют:

b_0 – величину *Y* при нулевом значении всех кодированных значений входных факторов (свободный член);

b_1, b_2, ..., b_i, ..., b_k –линейные эффекты влияния соответствующих кодированных значений входных факторов на величину *Y*;

b_{12}, b_{13}, ..., b_{1j}, ..., b_{1k}, b_{23}, b_{34}, ..., b_{2i}, ..., $b_{(k-1)i}$, ..., $b_{(k-1)k}$ – парные эффекты влияния соответствующих кодированных значений входных факторов на величину *Y* (эффекты «взаимодействия» двух соответствующих факторов);

b_{11}, b_{22}, ..., b_{ii}, ..., b_{kk} – квадратичные эффекты влияния соответствующих кодированных значений входных факторов на величину Y;

b_{123}, b_{124}, ..., b_{1uj}, ..., b_{234}, b_{235}, ..., b_{2ui}, ..., $b_{(k-2)(k-1)k}$ – тройные эффекты влияния соответствующих кодированных значений входных факторов на величину Y (эффекты «взаимодействия» трех соответствующих факторов) и т. д.

Наиболее удобно планировать эксперимент математическими методами для кодированных значений факторов x_i, получаемых из натуральных значений (Z_i) по следующим формулам:

$$x_i = \frac{Z_i - Z_i^0}{\Delta Z_i}; \quad Z_i^0 = \frac{Z_i^{\max} + Z_i^{\min}}{2}; \quad \Delta Z_i = \frac{Z_i^{\max} + Z_i^{\min}}{2},$$

где Z_i^0 – натуральное значение фактора в центре (середине) выбранной (заданной) факторной области.

Выбор плана эксперимента для применения РАМПЭ в отличие от планирования экспериментов для проведения КРА определяется видом выбранного семейства функций (видом полинома).

После завершения эксперимента для проведения РАМПЭ выполняют следующие действия:

- выбирают вид полинома (отрезок ряда Тейлора) для поиска уравнения регрессии;
- для выбранного полинома с помощью метода наименьших квадратов (МНК) рассчитывают параметры функции (выборочные коэффициенты уравнения регрессии);
- проверяют рассчитанные выборочные коэффициенты уравнения регрессии на значимость (равенство нулю);
- корректируют вид исходной функции, исключая из нее члены с незначимыми коэффициентами;
- оценивают ошибки, допускаемые при описании истинной зависимости с помощью найденного уравнения регрессии: проверяют адекватность уравнения регрессии с помощью распределения Фишера (функция f) или рассчитывают вероятность описания зависимости (U или φ) функцией f;
- если точность найденного уравнения регрессии не удовлетворяет, то выбирают, планируют и реализуют другой план эксперимента для поиска уравнения регрессии в другом семействе полиномов (например, полиномов более высокого порядка).

РАМПЭ в отличие от КРА имеет следующие особенности:

- выбирается только один класс функций – полиномы;
- используется только один метод приближения – МНК;
- выполняется меньшее количество этапов РА.

Обычно поиск уравнения регрессии начинают в семействе самых простых полиномов первого и второго порядка. По названиям степеней полиномов называют и планы эксперимента для применения РАМПЭ (планы первого и второго порядка).

1.2.1. Математические планы эксперимента первого порядка

Планы первого порядка позволяют находить линейные уравнения регрессии и нелинейные уравнения с членами, учитывающими эффекты «взаимодействия» (влияния одновременного изменения) кодированных значений входных факторов:

$$Y = b_0 + \sum\nolimits_{i=1}^{k} b_i x_i = b_0 + b_1 x_1 + b_2 x_2 + \ldots + b_i x_i + \ldots + b_k x_k;$$

$$Y = b_0 + \sum\nolimits_{i=1}^{k} b_i x_i + \sum\nolimits_{u,i=1(u \neq i)}^{k} b_{ui} x_u x_i + \sum\nolimits_{g,u,i=1(g \neq u \neq i)}^{k} b_{gui} x_g x_u x_i + \ldots =$$

$$= b_0 + b_1 x_1 + \ldots b_k x_k + b_{12} x_1 x_2 + b_{13} x_1 x_3 + \ldots b_{(k-1)k} x_{k-1} x_k b_{123} x_1 x_2 x_3 +$$

$$+ b_{124} x_1 x_2 x_4 + \ \ldots + b_{(k-2)(k-1)k} x_{k-2} x_{k-1} x_k .$$

Для удобства программирования расчетов в состав уравнения регрессии можно вводить фиктивную переменную $x_0 = +1$ во всех опытах эксперимента, тогда

$$Y = \sum\nolimits_{i=0}^{k} b_i x_i ;$$

$$Y = \sum\nolimits_{j=0}^{k} b_i x_i + \sum\nolimits_{u,i=1(u \neq i)}^{k} b_{ui} x_u x_i + \sum\nolimits_{g,u,i=1(g \neq u \neq i)}^{k} b_{gui} x_g x_u x_i + \ldots$$

Для РАМПЭ наибольшее распространение получили двухуровневые ($m_l = m = 2$) ортогональные и D-оптимальные планы первого порядка $2^{(k\text{-}a)}$. При таких планах все факторы в кодированном виде могут иметь только два значения ($x_i = +1$ и $x_i = -1$). Тип плана обозначает формулу для расчета числа его опытов без их повторения N: $N = m^{(k-a)} = 2^{(k-a)}$, где $k > a$ при a, равном 0, 1, 2, 3, ..., k–2.

При $a = 0$ план является планом полного факторного эксперимента 2^k, а при $a > 0$ – планом дробного факторного эксперимента $2^{(k-a)}$. При ПФЭ все выборочные коэффициенты уравнения регрессии являются достаточно точными, «несмешанными» оценками соответствующих генеральных коэффициентов $b_d \approx \beta_d$.

Условием ортогональности плана эксперимента является выполнение условия

$$\sum_{c=1}^{N} x_{cu} x_{cj} = 0 \text{ при } u \neq j \text{ и } u, j, \text{ равных } 0, 1, 2, ..., k.$$

Планы, отвечающие условиям ортогональности, позволяют любой коэффициент уравнения регрессии рассчитывать по одной формуле:

$$b_d = \frac{\sum_{c=1}^{N} x_{c,d} y_c}{N},$$

где c – номер опыта в плане эксперимента;

b_d – коэффициент, учитывающий эффект факторов, значения которых приведены в столбце x_d плана эксперимента;

y_c – свойства объекта, измеренные при проведении соответствующего опыта;

N – число опытов в эксперименте.

Ортогональные планы экспериментов обладают свойством D-оптимальности. ***D*-оптимальные планы** обеспечивают минимальную и одинаковую ошибку в оценке всех коэффициентов уравнения регрессии (S_b^2), определяемую по формуле

$$S_{b_d}^2 = S_b^2 = \frac{S_{воспр}^2}{N},$$

где $S_{воспр}^2$ – дисперсия воспроизводимости, характеризующая случайные ошибки всех опытов эксперимента.

Для D-оптимальных планов должны выполняться следующие условия:

$$\sum_{c=1}^{N} x_{c,j} = 0 \text{ при } j, \text{ равном } 1, 2, ..., k;$$

$$\sum_{c=1}^{N} x_{c,j}^2 = N \text{ при } j, \text{ равном } 0, 1, 2, ..., k.$$

Выбор плана эксперимента начинается с расчета необходимого числа опытов $N_{необх}$ или его задания $N_{зад}$. При этом при расчете выбранного числа опытов в эксперименте N **для двухуровневых планов первого порядка** должны выполняться следующие соотношения:

$$N \geq N_{необх};$$
$$N_{необх} \geq k+1;$$
$$N_{необх} \geq L+1;$$
$$N_{зад} \geq N,$$

где L – общее число коэффициентов в выбранном семействе полиномов (число отрезков ряда Тейлора).

При расчете $N_{необх}$ задаются видом полинома (типом и числом коэффициентов уравнения регрессии L). При задании числа опытов рассчитывают количество входных факторов, которые можно одновременно изменять в эксперименте, и число коэффициентов полинома, которые можно определить по результатам эксперимента с заданным числом опытов:

$$k_{\max} = N_{зад} - 1;$$
$$L_{\max} = N_{зад} - 1.$$

Пример 2. Необходимо исследовать влияние на свойство объекта ***Y*** четырех кодированных значений факторов (k = 4) и описать их зависимость уравнением регрессии в виде следующего нелинейного полинома (L = 11):

$$Y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4 + b_{13}x_1x_3 + b_{14}x_1x_4 + b_{23}x_2x_3 + b_{24}x_2x_4 + b_{34}x_3x_4.$$

Тогда совместное выполнение соотношений даст:

$$N_{необх} \geq k+1 \geq 4+1 \geq 5;$$
$$N_{необх} \geq L+1 \geq 11+1 \geq 12;$$
$$N \geq N_{необх} \geq 12.$$

Следовательно, для оценки влияния четырех факторов минимальное число опытов в эксперименте должно быть равно пяти, а для получения уравнения регрессии с одиннадцатью членами число опытов должно быть не менее двенадцати.

Первоначально выберем план из группы планов первого порядка $2^{(k–a)}$. Для четырех факторов (k = 4) возможна реализация плана полного факторного эксперимента типа 2^k ($N = 2^4 = 16$) и двух планов дробного факторного эксперимента типа 2^{k-a}: полуреплика ПФЭ типа 2^{k-1} ($N = 2^{4-1} = 8$) и четверть реплики ПФЭ типа 2^{k-2} ($N = 2^{4-2} = 4$).

Очевидно, что нашему условию ($N \geq 12$) отвечает только план первого порядка типа 2^k с $N = 16$. План ПФЭ типа 2^4 позволяет **получить наиболее точные оценки** коэффициентов уравнения регрессии.

Для построения ортогонального и D-оптимального плана ПФЭ типа 2^4 воспользуемся одним из распространенных приемов математической комбинаторики, заключающемся в следующем:

- делается заготовка плана в виде таблицы (плана-матрицы эксперимента), в которой предусматривается не менее N строк и $L_{\max}$ столбца для x_d;
- в первый столбец таблицы заносят номера строк, соответствующие номерам опытов;
- во второй столбец (этот столбец не является обязательным при построении матрицы плана эксперимента) – кодированные значения фиктивного фактора x_0 (во всех строках плана $x_0 = +1$);
- в третий столбец – кодированные значения первого фактора x_1 в виде последовательного чередования друг за другом значений (+1) и (–1);
- в следующем, четвертом столбце для x_2, выбранная комбинация чередований в предыдущем столбце знаков (+1) и (–1) **удваивается**, например: после двух знаков (+1) следуют два знака (–1);
- по аналогичному принципу удвоения комбинации чередования знаков предыдущего столбца заполняются и последующие столбцы для всех оставшихся факторов x_i, столбцы для оценки эффектов «взаимодействия» факторов (x_{12}, x_{13}, x_{23} и др.) заполняются путем перемножения знаков для соответствующих факторов в соответствующих строках таблицы ($x_{12} = x_1x_2$, $x_{13} = x_1x_3$ и т. д.);
- правильность составления плана проверяется по выполнению условия его D-оптимальности $\sum_{c=1}^{N} x_{c,j} = 0$ (для всех столбцов, кроме столбца для x_0).

Построенный по этому приему план приведен в табл. 4. Данный план является ортогональным и D-оптимальным.

План с натуральными значениями факторов Z_i строится исходя из плана с кодированными значениями путем замены знаков (+1) и (–1) на соответствующие им максимальные и минимальные натуральные значения для данных факторов в их факторной области.

Довольно часто на практике приходится задаваться не видом полинома, а числом опытов из-за дефицита ресурсов для проведения эксперимента (времени, средств и др.). В этом случае выбор плана эксперимента начинают с расчета параметров полинома, которые возможно определить при заданном числе опытов $N_{зад}$.

Таблица 4

План 2^4

Номер опыта	Кодированные значения факторов x_d											Свойство объекта Y
	x_0^*	x_1	x_2	x_3	x_4	x_{12}	x_{13}	x_{14}	x_{23}	x_{24}	x_{34}	
1	1	1	1	1	1	1	1	1	1	1	1	
2	1	–1	1	1	1	–1	–1	–1	1	1	–1	
3	1	1	–1	1	1	–1	1	1	–1	–1	1	
4	1	–1	–1	1	1	1	–1	–1	–1	–1	–1	
5	1	1	1	–1	1	1	–1	1	–1	1	1	
6	1	–1	1	–1	1	–1	1	–1	–1	1	–1	
7	1	1	–1	–1	1	–1	–1	1	1	–1	1	
8	1	–1	–1	–1	1	1	1	–1	1	–1	–1	
9	1	1	1	1	–1	1	1	–1	1	–1	1	
10	1	–1	1	1	–1	–1	–1	1	1	–1	–1	
11	1	1	–1	1	–1	–1	1	–1	–1	1	1	
12	1	–1	–1	1	–1	1	–1	1	–1	1	–1	
13	1	1	1	–1	–1	1	–1	–1	–1	–1	1	
14	1	–1	1	–1	–1	–1	1	1	–1	–1	–1	
15	1	1	–1	–1	–1	–1	–1	–1	1	1	1	
16	1	–1	–1	–1	–1	1	1	1	1	1	–1	

---

* Необязательный столбец.

Допустим, что $N_{зад} = 10$, $k = 4$ и уравнение регрессии необходимо получить в следующем виде ($L = 5$):

$$y = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + b_4 x_4.$$

Оценим возможности двухуровнего плана первого порядка с $N = 10$:

$$k_{\max} = N_{зад} - 1 = 10 - 1 = 9;$$

$$L_{\max} = N_{зад} - 1 = 10 - 1 = 9.$$

Из данных равенств следует, что план с 10 опытами позволяет решить нашу задачу ($k = 4$ и $L = 5$), так как при десяти опытах можно оценить в эксперименте одновременное влияние до девяти факторов ($k \leq k_{\max}$) по уравнению регрессии в виде полинома с девятью коэффициентами ($L \leq L_{\max}$).

При $N_{зад}$ = 10 из планов первого порядка возможна реализация только плана ДФЭ, так как план ПФЭ для четырех факторов насчитывает 16 опытов.

Воспользуемся методом дробных реплик для построения планов ДФЭ типа $2^{(k–a)}$. Наиболее близким по числу опытов к $N_{зад}$ = 10 является полуреплика (1/2 часть) плана ПФЭ, т. е. план ДФЭ $2^{(4–1)}$ с числом опытов N = 8. Проверка показывает, что план $2^{(4–1)}$ пригоден для решения поставленной задачи, так как выполняются следующие соотношения:

$$N \geq k + 1 \geq 4 + 1 \geq 5 \ (8 > 5);$$
$$N \geq L + 1 \geq 5 + 1 \geq 6 \ (8 > 6).$$

Линейный план дробного факторного эксперимента типа $2^{(4–2)}$ с числом опытов N = 4 не пригоден для решения поставленных задач, так как:

$$N \geq k + 1 \geq 4 + 1 \geq 5 \ (4 < 5!);$$
$$N \geq L + 1 \geq 5 + 1 \geq 6 \ (4 < 6!).$$

Поскольку план ДФЭ представляет собой часть опытов плана ПФЭ, то необходимо решить, какой именно набор опытов из плана ПФЭ использовать в плане ДФЭ $2^{(4-1)}$. От этого набора будет зависеть точность определения эффектов влияния факторов на свойство Y (так называемая «смешиваемость» коэффициентов). Если взять для плана ДФЭ первую (опыты 1–8) или вторую половину (опыты 9–16) плана ПФЭ (см. табл. 1), то полученный план ДФЭ будет не четырехфакторный, а трехфакторный (фактор x_4 постоянен) и не будет обладать свойствами ортогональности и D-оптимальности для столбца x_4 ($\sum_{c=1}^{N} x_{cu} x_{c4} = 0$ и $\sum_{c=1}^{N} x_{c4} = 0$), т. е. лишен преимуществ метода математического планирования эксперимента по минимальной точности получаемых оценок.

Построение планов ДФЭ начинают по тому же приему математической комбинаторики, что и при построении планов ПФЭ для числа факторов, равных разности (k–a). В нашем случае k–a = 4–1 = 3. Поэтому построим первоначально заготовку плана ДФЭ типа $2^{(4–1)}$ в виде плана ПФЭ типа 2^3, предусмотрев в нем не менее 5 колонок для x_d (табл. 5).

При заполнении столбца для фактора x_4 **принцип удвоения** чередований уровней, применяемый для построения планов ПФЭ, не подходит, так как его использование в данном столбце даст только знаки (+1) и такой план не будет являться ортогональным и *D*-оптимальным. Если же для заполнения столбца x_4 воспользоваться произведением двух и более других факторов в одной строке плана (так называемым **генерирующим соотношением** [5]), то тогда план будет и ортогональным и *D*-оптимальным.

Таблица 5

Заготовка плана ДФЭ типа $2^{(4-1)}$

Номер опыта	Кодированные значения факторов x_d						Y
	x_1	x_2	x_3	x_4	x_{12}	…	
1	+1	+1	+1				
2	–1	+1	+1				
3	+1	–1	+1				
4	–1	–1	+1				
5	+1	+1	–1				
6	–1	+1	–1				
7	+1	–1	–1				
8	–1	–1	–1				

К выбору генерирующего соотношения нужно подходить осознанно, так как оно определяет «смешанность» (точность) коэффициентов уравнения регрессии, полученных по составленному плану.

Составим и проанализируем все возможные варианты генерирующего соотношения для x_4:

$$x_4 = x_1x_2(\text{I});\; x_4 = x_1x_3(\text{II});\; x_4 = x_2x_3(\text{III});\; x_4 = x_1x_2x_3(\text{IV}).$$

Для данных генерирующих соотношений рассчитаем **определяющие контрасты** [5] путем умножения левой и правой частей соответствующего генерирующего соотношения на x_4:

$$x_4^2 = x_1x_2x_4;\; x_4^2 = x_1x_3x_4;\; x_4^2 = x_2x_3x_4;\; x_4^2 = x_1x_2x_3x_4.$$

Так как $x_4 = \pm 1$, то $x_4^2 = 1$ и определяющие контрасты можно выразить следующими равенствами:

$$1 = x_1x_2x_4;\; 1 = x_1x_3x_4;\; 1 = x_2x_3x_4;\; 1 = x_1x_2x_3x_4.$$

Перемножив левые и правые части определяющих контрастов на каждый фактор, можно определить в плане столбцы с одинаковым порядком чередования знаков (+1) и (–1), например для фактора x_1:

$$x_1 = x_1^2 x_2 x_4 = x_2 x_4; x_1 = x_1^2 x_3 x_4 = x_3 x_4;$$

$$x_1 = x_1 x_2 x_3 x_4; x_1 = x_1^2 x_2 x_3 x_4 = x_2 x_3 x_4.$$

Эти равенства показывают, что при генерирующих соотношениях I–IV выборочный коэффициент b_1 будет служить оценкой влияния на y не только фактора x_1, но и других:

$$b_1 \approx b_1 + b_{24}; b_1 \approx b_1 + b_{34}; b_1 \approx b_1 + b_{1234}; b_1 \approx b_1 + b_{234}.$$

Эффекты «взаимодействия» трех и более факторов обычно близки к нулю и ими можно пренебрегать [2]:

$$b_1 \approx b_1 + b_{24}; b_1 \approx b_1 + b_{34}; b_1 \approx b_1; b_1 \approx b_1.$$

Поэтому можно считать, что «несмешанные» оценки эффекта влияния фактора x_1 на свойство Y могут быть получены при реализации плана ДФЭ с генерирующими соотношениями III и IV для фактора x_4. Результаты проверки на «смешиваемость» остальных эффектов приведены в табл. 6.

Таблица 6

Параметры проверки разрешающей силы
дробной реплики плана $2^{(4–1)}$

Параметр	Выражение для определения параметра			
Генерирующее соотношение	$x_4 = x_1x_2$	$x_4 = x_1x_3$	$x_4 = x_2x_3$	$x_4 = x_1x_2x_3$
Определяющий контраст	$1 = x_1x_2x_4$	$1 = x_1x_3x_4$	$1 = x_2x_3x_4$	$1 = x_1x_2x_3x_4$
Оценки точности («смешанности») коэффициентов уравнения регрессии	$b_0 \approx \beta_0+\beta_{124}$ $b_1 \approx \beta_1 + \beta_{24}$ $b_2 \approx \beta_2 + \beta_{14}$ $b_3 \approx \beta_3 + \beta_{1234}$ $b_4 \approx \beta_4 + \beta_{12}$ $b_{12} \approx \beta_{12}+\beta_4$ $b_{13} \approx \beta_{13}+\beta_{234}$ $b_{14} \approx \beta_{14}+\beta_2$ $b_{23} \approx \beta_{23}+\beta_{134}$ $b_{24} \approx \beta_{24}+\beta_1$ $b_{34} \approx \beta_{34}+\beta_{123}$	$b_0 \approx \beta_0+\beta_{134}$ $b_1 \approx \beta_1 + \beta_{34}$ $b_2 \approx \beta_2 + \beta_{1234}$ $b_3 \approx \beta_3 + \beta_{14}$ $b_4 \approx \beta_4 + \beta_{13}$ $b_{12} \approx \beta_{12}+\beta_{234}$ $b_{13} \approx \beta_{13}+\beta_4$ $b_{14} \approx \beta_{14}+\beta_3$ $b_{23} \approx \beta_{23}+\beta_{124}$ $b_{24} \approx \beta_{24}+\beta_{123}$ $b_{34} \approx \beta_{34}+\beta_1$	$b_0 \approx \beta_0+\beta_{234}$ $b_1 \approx \beta_1 + \beta_{1234}$ $b_2 \approx \beta_2 + \beta_{34}$ $b_3 \approx \beta_3 + \beta_{24}$ $b_4 \approx \beta_4 + \beta_{23}$ $b_{12} \approx \beta_{12}+\beta_{134}$ $b_{13} \approx \beta_{13}+\beta_{124}$ $b_{14} \approx \beta_{14}+\beta_{123}$ $b_{23} \approx \beta_{23}+\beta_4$ $b_{24} \approx \beta_{24}+\beta_3$ $b_{34} \approx \beta_{34}+\beta_2$	$b_0 \approx \beta_0+\beta_{1234}$ $b_1 \approx \beta_1 + \beta_{234}$ $b_2 \approx \beta_2 + \beta_{134}$ $b_3 \approx \beta_3 + \beta_{124}$ $b_4 \approx \beta_4 + \beta_{123}$ $b_{12} \approx \beta_{12}+\beta_{34}$ $b_{13} \approx \beta_{13}+\beta_{24}$ $b_{14} \approx \beta_{14}+\beta_{23}$ $b_{23} \approx \beta_{23}+\beta_{14}$ $b_{24} \approx \beta_{24}+\beta_{13}$ $b_{34} \approx \beta_{34}+\beta_{12}$

Данные табл. 6 показывают, что при любом генерирующем соотношении точными («несмешанными») будут пять коэффициентов. Для генерирующего соотношения I точными будут все коэффициенты, оценивающие эффект фактора x_3, при II – эффект фактора x_2, при III – эффект фактора x_1, а при IV – линейные эффекты всех факторов.

Так как по заданию нам необходимо получить уравнение регрессии с линейными эффектами всех факторов и эти эффекты должны быть наиболее точными, то выбираем генерирующее соотношение IV и в соответствии с ним заполняем колонку плана для x_4 (см. табл. 5).

Этот план, с $N = 8$, является ортогональным и D-оптимальным. По данному плану есть возможность оценить еще 6 эффектов парного влияния факторов, однако, как показывают данные табл. 6, расчеты приведут к получению «смешанных» коэффициентов уравнения регрессии (т. е. неточно отражающих парное влияние соответствующих факторов), так как комбинации знаков (изменения значений факторов в опытах) совпадают для x_{14} и x_{23}, x_{13} и x_{24} и др.

Для получения заданного уравнения регрессии для точной оценки всех линейных эффектов влияния входных факторов подходит только IV генерирующее соотношение, в соответствии с которым и заполним столбец x_4 в плане эксперимента для 8 опытов (табл. 7).

Таблица 7

План ДФЭ типа $2^{(4-1)}$ с генерирующим соотношением
$x_4 = x_1x_2x_3$

Номер опыта	Кодированные значения факторов							Y
	x_1	x_2	x_3	x_4	x_{12} и x_{34}	x_{13} и x_{24}	x_{14} и x_{23}	
1	+1	+1	+1	+1	+1	+1	+1	
2	–1	+1	+1	–1	–1	–1	+1	
3	+1	–1	+1	–1	–1	+1	–1	
4	–1	–1	+1	+1	+1	–1	–1	
5	+1	+1	–1	–1	+1	–1	–1	
6	–1	+1	–1	+1	–1	+1	–1	
7	+1	–1	–1	+1	–1	–1	+1	
8	–1	–1	–1	–1	+1	+1	+1	
9	0	0	0	0	0	0	0	
10	0	0	0	0	0	0	0	

Два дополнительных опыта в эксперименте (опыты № 9 и № 10) можно использовать как повторные для оценки дисперсии воспроизводимости эксперимента, если сделать допущение, что и другие опыты плана имеют такие же случайные ошибки. Дисперсия воспроизводимости $S^2_{воспр}$ по результатам этих двух опытов может быть использована для оценки ошибки в определении коэффициентов уравнения регрессии и их значимости, а также проверки адекватности найденного уравнения регрессии.

В соответствии с общепринятыми рекомендациями запланируем опыты для определения $S^2_{воспр}$ при нулевых кодированных значениях всех исследуемых факторов, т. е. в центре области изменения факторов (см. табл. 7).

1.2.2. Математические планы эксперимента второго порядка (композиционные планы эксперимента)

Если после реализации планов эксперимента первого порядка по результатам регрессионного анализа не были получены адекватные линейные экспериментально-статистические модели объекта, то планируют и реализуют дополнительные опыты по **композиционному плану эксперимента** второго порядка. Ядром композиционных математических планов экспериментов являются планы полных или дробных факторных экспериментов первого порядка. По сравнению с планами первого порядка композиционные планы экспериментов второго порядка содержат дополнительные опыты в «звездных» точках [5].

В отличие от планов первого порядка (линейных планов), планы второго порядка не удовлетворяют одновременно нескольким критериям оптимальности, что вынуждает выбирать план с характеристиками, которые наиболее точно отвечают задачам исследования [4].

Существуют различные виды композиционных планов второго порядка: трехуровневые планы Бокса (3^k), планы Бокса – Уилсона, Бокса – Хантера, Коно и др. Эти планы позволяют найти уравнение регрессии в следующем семействе полиномов второго порядка, например для кодированных значений входных факторов:

$$Y = b_0 + \sum\nolimits_{i=1}^{k} b_i x_i + \sum\nolimits_{u,i=1(u \neq i)}^{k} b x_u x_i + \sum\nolimits_{g,u,i=1(g \neq u \neq i)}^{k} b x_g x_u x_i + \sum\nolimits_{i=1}^{k} b_{ii} x_i^2 + \ldots$$

Для получения квадратичного уравнения регрессии необходимо, чтобы число уровней каждого фактора было не менее трех ($m_l \geq 3$).

Трехуровневые планы полного факторного эксперимента носят имя планы Бокса. Достоинством планов Бокса является высокая точность в определении эффектов влияния факторов. Планы Бокса не являются ортогональными и *D*-оптимальными. К недостаткам планов Бокса относится также большое число опытов, намного превышающее максимально возможное число коэффициентов в квадратичном полиноме *L* (табл. 8).

Таблица 8

Параметры планов Бокса

Параметр	Значение параметра при числе входных факторов *k*				
	2	3	4	5	6
Число опытов $N = 3^k$	9	27	81	243	729
L	6	10	15	21	28

В планах Бокса все факторы имеют три кодированные значения: 1; 0 и (–)1. Область факторного пространства для плана Бокса 3^2 приведена на рис. 1.

На рис. 1 видно, что опыты в планах Бокса находятся в областях всех четырех граней куба и в центре между его гранями. Планы Бокса называют планами на кубе.

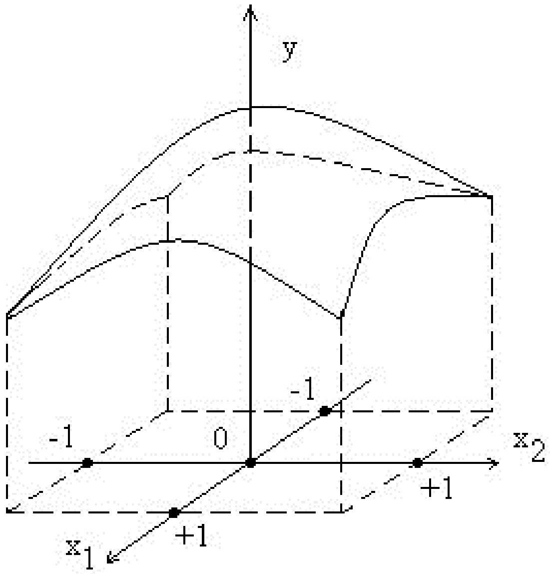


Рис. 1. Факторное пространство входных факторов в плане двухфакторного трехуровневого эксперимента 3^2 с кодированными значениями факторов

Матрица плана Бокса для трехуровневого двухфакторного эксперимента без повторения опытов с кодированными значениями факторов приведена в табл. 9.

Таблица 9

Матрица плана эксперимента 3^2

Номер опыта	Кодированные значения факторов x_d					Свойство объекта Y
	x_1	x_2	x_1x_2	x_1^2	x_2^2	
1	–1	–1	1	1	1	
2	0	–1	0	0	1	
3	1	–1	–1	1	1	
4	–1	0	0	0	0	
5	0	0	0	1	0	
6	1	0	0	0	0	
7	–1	1	–1	1	1	
8	0	1	0	0	1	
9	1	1	1	1	1	

Пример 3. Для получения уравнения регрессии, описывающего влияние на предел текучести (Y, кгс/см^2) пресс-композиции температуры (Z_1, °С) и времени (Z_2, мин) прессования, был составлен по выбранному приему математической комбинаторики и реализован план Бокса трехуровневого двухфакторного эксперимента без повторений опытов. Результаты эксперимента представлены в табл. 10.

Таблица 10

Результаты эксперимента по плану Бокса типа 3^2 с кодированными значениями факторов

Номер опыта	Кодированные значения факторов, x_d					Y, кгс/см^2
	x_1	x_2	x_1x_2	x_1^2	x_2^2	
1	–1	–1	1	1	1	20,8
2	0	–1	0	0	1	11,6
3	1	–1	–1	1	1	9,4
4	–1	0	0	0	0	19,4
5	0	0	0	1	0	12,3
6	1	0	0	0	0	9,0
7	–1	1	–1	1	1	4,1
8	0	1	0	0	1	9,6
9	1	1	1	1	1	7,4

Как наиболее экономные по числу опытов популярностью пользуются ортогональные композиционные планы второго порядка Бокса – Уилсона, включающие в себя составной частью планы первого порядка: типа 2^k (при $k < 5$) и $2^{(k–1)}$ (при $k \geq 5$). **Планы Бокса – Уилсона не являются *D*-оптимальными**.

Общее число опытов плана Бокса – Уилсона $N_{БУ}$ без повторения опытов рассчитывается по следующим формулам:

$$N_{БУ} = N_{ПФЭ} + N^* + n_0 = \ 2^k + \ 2k + n_0 (k < 5);$$
$$N_{БУ} = N_{ДФЭ} + N^* + n_0 = \ 2^{k-1} + \ 2k + n_0 (k \geq 5),$$

где $N_{ПФЭ}$ и $N_{ДФЭ}$ – число опытов плана первого порядка;

N^* – число опытов в «звездных» точках;

n_0 – число опытов при нулевых кодированных значениях всех исследуемых факторов **задается исследователем (любое целое число не менее 2)**.

Построение плана Бокса – Уилсона начинается с построения входящего в его состав плана первого порядка.

После заполнения всех строк плана первого порядка ($N_{ПФЭ}$ или $N_{ДФЭ}$) заполняют $2k$ строк для «звездных точек» плана (N^*) и строки с нулевыми кодированными значениями факторов (n_0).

Звездные точки располагаются на координатных осях соответствующих кодированных значений факторов на расстоянии $\lambda \geq 1$ от начала координат. В этом случае факторное пространство для кодированных значений двух факторов выглядит следующим образом (рис. 2).

Композиционные планы второго порядка (со «звездными точками») предусматривают проведение эксперимента в большей области факторного пространства по сравнению с планами Бокса (см. рис. 2) и их называют планами на поверхности.

При составлении планов Бокса – Уилсона для получения ортогонального плана величина λ (величина «звездного плеча») рассчитывается по специальным формулам:

$$l^4 + 2^k l^2 - 2^{(k-1)} \left(k + \ 0,5n_0\right) = 0 \ (k < 5);$$
$$l^4 + 2^{(k-1)} l^2 - 2^{(k-2)} \left(k + \ 0,5n_0\right) = 0 \ (k \geq 5).$$

На основании этих формул составлены таблицы для λ^2 при различных величинах k и n_0. Так, например, при $k = 2$ величина $\lambda^2 = 1$ (при $n_0 = 1$), $\lambda^2 = 1,160$ (при $n_0 = 2$) и $\lambda^2 = 1,606$ (при $n_0 = 5$).

После заполнения строк для «звездных точек» плана Бокса – Уилсона заполняют строки с нулевыми кодированными значениями всех исследуемых факторов (число строк равно n_0).

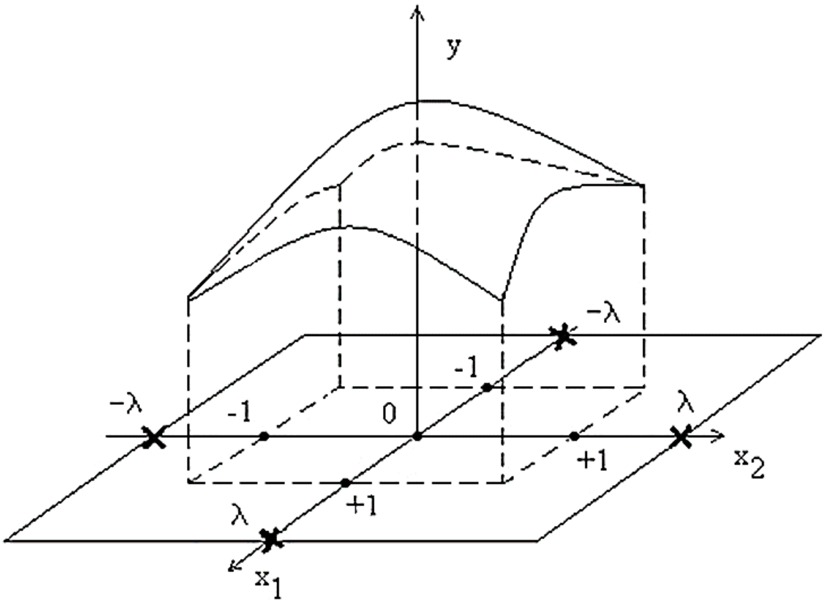


Рис. 2. Факторное пространство входных факторов в плане двухфакторного четырехуровневого эксперимента Бокса – Уилсона с кодированными значениями факторов

Алгоритмы проведения РАМПЭ по планам Бокса – Уилсона изложены в книге [5].

Начинать составление планов Бокса – Уилсона для натуральных значений факторов необходимо с анализа физического и химического смысла факторного пространства. Так, например, если при планировании вы задаете область изменения времени реакции от 1 до 11 мин, то крайним натуральным значениям необходимо присвоить кодированные значения соответственно ($-\lambda$) и ($+\lambda$), а затем, исходя из этого, рассчитать $Z_{\min}$ и $Z_{\max}$ для соответствующих кодированных значений $x = -1$ и $x = 1$. Если же принять, что $x = -1$ соответствует $Z_{\min} = 1$ мин, то даже для $x = \lambda = -1$ значение $Z_{(-\lambda)} = \lambda \Delta Z + Z^0 =$ $= -1 \cdot 10 + 6 = -4$ мин, что не имеет физического смысла.

Пример плана Бокса – Уилсона для $k = 2$ и $n_0 = 2$ приведен в табл. 11.

Ротатабельные планы Бокса – Хантера позволяют получать более точное описание поверхности отклика объекта по сравнению с ортогональными планами Бокса – Уилсона. При составлении ротатабельных планов используются такие же принципы математической комбинаторики, как и при составлении планов Бокса – Уилсона, только величина «звездного плеча» λ рассчитывается по другим формулам и число опытов в центре плана n_0 не задается исследователем, а рассчитывается исходя из числа входных факторов. Алгоритмы проведения РАМПЭ по планам Бокса – Хантера изложены в книге [3].

Таблица 11

План Бокса – Уилсона для k = 2 и n_0 = 2

Номер опыта	Значения факторов								Y
	кодированные						натуральные		
	x_0	x_1	x_2	x_{12}	x_1'	x_2'	Z_1	Z_2	
1	+1	+1	+1	+1	0,368	0,368			
2	+1	–1	+1	–1	0,368	0,368			
3	+1	+1	–1	–1	0,368	0,368			
4	+1	–1	–1	+1	0,368	0,368			
5	+1	1,162	0	0	0,528	–0,632			
6	+1	–1,162	0	0	0,528	–0,632			
7	+1	0	1,162	0	–0,632	0,528			
8	+1	0	–1,162	0	–0,632	0,528			
9	+1	0	0	0	–0,632	–0,632			
10	+1	0	0	0	–0,632	–0,632			

Для объектов исследования, имеющих большие значения дисперсии воспроизводимости результатов экспериментов (например, в производственных условиях), используются робастные планы экспериментов (планы Тагучи).

2. МАТЕМАТИЧЕСКОЕ ПЛАНИРОВАНИЕ ЭКСПЕРИМЕНТА ДЛЯ РЕШЕНИЯ ОПТИМИЗАЦИОННЫХ ЗАДАЧ

Оптимизационные задачи в большинстве случаев формулируются как задачи поиска экстремальных значений функции отклика объекта (максимумов или минимумов) или их заданных значений. Например, часто приходится решать задачи поиска оптимальных условий производства продукции с **максимальной** производительностью или **минимальной** себестоимостью.

Поведение объекта во многих случаях приходится описывать несколькими функциями отклика. Очень редко удается найти такое сочетание значений всех влияющих факторов, при которых достигаются все желаемые экстремумы функций отклика объекта. Большинство влияющих факторов можно изменять только в реальных пределах: концентрации реагентов не могут быть отрицательными, температуры процессов не могут превышать безопасные значения и т. д. Исследователю нужны и реальные значения функции отклика объекта: неотрицательные значения выхода продукции, степени очистки, себестоимости продукции и др. Поэтому в большинстве случаев оптимизационные задачи решают при условии различных ограничений на величину влияющих факторов и значений функций откликов объектов, т. е. проводят **поиск рациональных значений** влияющих факторов.

Если свойство (отклик) объекта описывается полиномом второго порядка, то для решения оптимизационных задач полезным является исследовать соответствующую поверхность отклика. Для этого переходят от полинома второго порядка к стандартному каноническому уравнению.

Первый этап канонического преобразования – перенос начала координат в особую точку поверхности отклика – центр поверхности. Координаты центра поверхности *S* определяются решением следующей системы уравнений:

$$\left.\begin{array}{l} dY / dx_1 = 0, \\ dY / dx_2 = 0, \\ \dots\dots\dots\dots \\ dY / dx_k = 0. \end{array}\right\}$$

Если поверхность отклика имеет центр, то в него переносят начало координат. При этом в уравнении поверхности исчезают члены, содержащие линейные эффекты факторов, и изменяется величина свободного члена.

Второй этап канонического преобразования – это поворот координатных осей в новом центре таким образом, что в уравнении поверхности исчезают члены с эффектами взаимодействия факторов. После завершения канонических преобразований получают стандартное каноническое уравнение поверхности следующего вида:

$$Y - Y_s = \omega_{11} G_1^2 + \omega_{22} G_2^2 + \ldots + \omega_{kk}^2 G_k^2,$$

где Y_s – значение отклика объекта в центре поверхности;

ω_{11}, ω_{22}, ..., ω_{kk} – коэффициенты канонической формы уравнения регрессии;

G_1, G_2, ..., G_k – канонические переменные, являющиеся функциями факторов x_k.

Поверхности второго порядка классифицируют по их каноническим формам следующим образом.

1. **Эллиптический параболоид** (рис. 3 и 4) – если все коэффициенты ω_{jj} имеют одинаковые знаки: все знаки положительные – поверхность имеет минимум; все знаки отрицательные – поверхность имеет максимум.

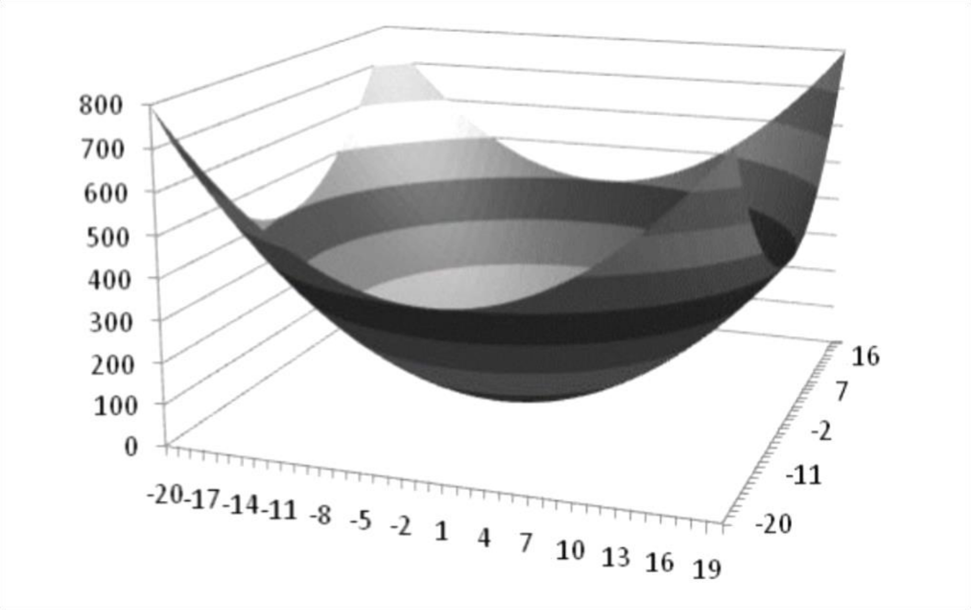


Рис. 3. Эллиптический параболоид
при всех положительных ω_{jj}

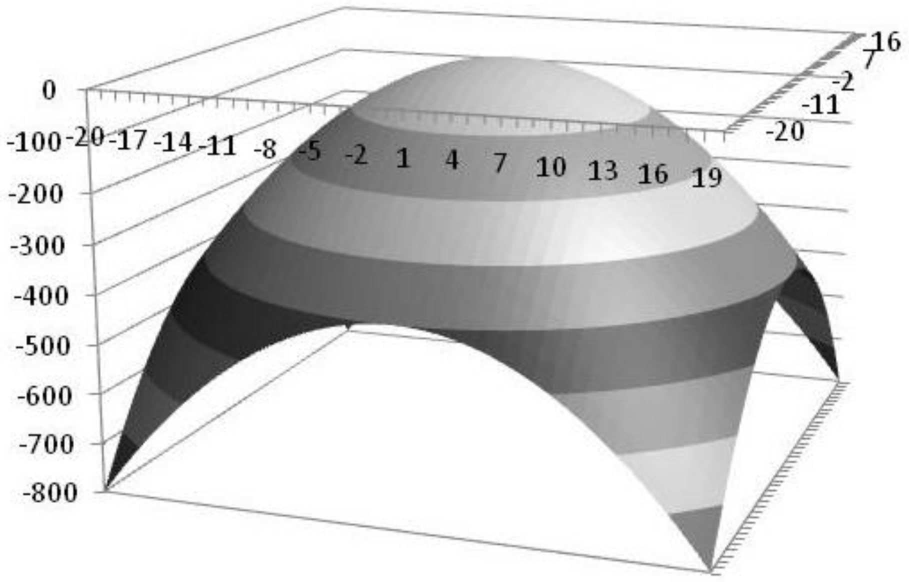


Рис. 4. Эллиптический параболоид
при всех отрицательных ω_{jj}

2. **Гиперболический параболоид**, или «**седло»**, с «мини-максом» в центре (рис. 5) – если коэффициенты ω_{jj} имеют разные знаки (положительные и отрицательные).

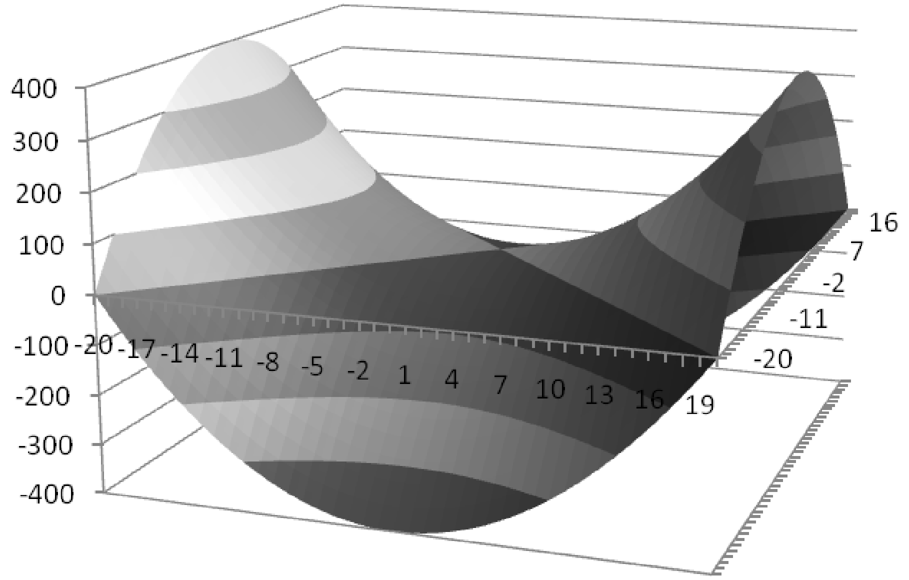


Рис. 5. Гиперболический параболоид, «седло»,
при положительном знаке ω_{11} и отрицательном знаке ω_{22}

3. Поверхность «**возрастающее возвышение**», или «**гребень»**, с центром, находящимся далеко за областью экспериментирования (рис. 6), – если один или несколько коэффициентов ω_{jj} (***но не все!***) близки к нулю.

С алгоритмами проведения канонических преобразований уравнений регрессии второго порядка и примерами анализа поверхности отклика можно познакомиться в книге [5].

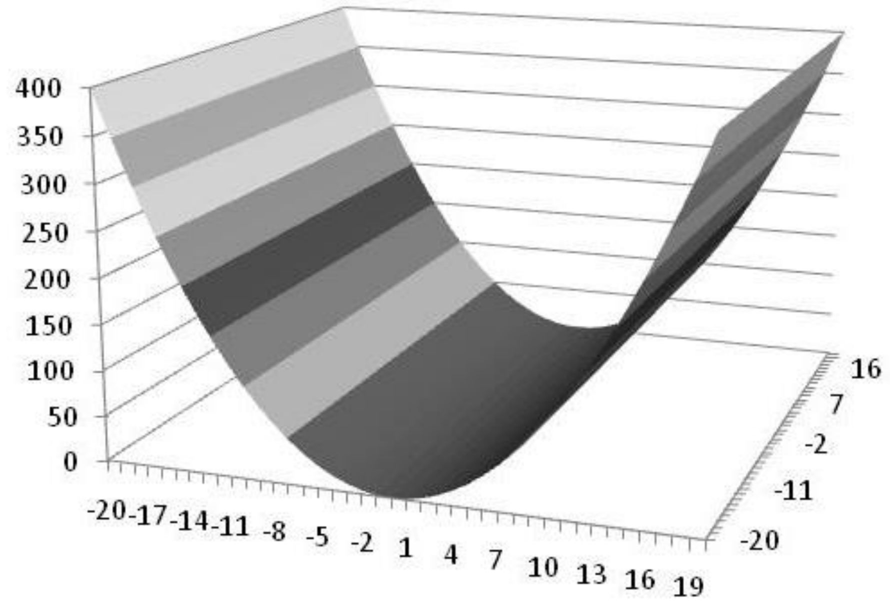


Рис. 6. Поверхность «гребень» при нулевом значении ω_{22}

Довольно часто уравнение регрессии невозможно преобразовать в каноническую форму и провести анализ поверхности отклика объекта. Для поиска рациональных значений в исследованной области факторного пространства по найденным уравнениям регрессии можно применить математические методы, которые реализованы в различных программных продуктах. В программе MS Excel для поиска экстремальных значений линейных уравнений регрессии применяется симплексный метод, а для нелинейных – другие методы.

Алгоритмы и примеры решения оптимизационных задач детально освещены в книгах [3–6].

Довольно часто в исследованной части факторного пространства математические методы не находят экстремума уравнения регрессии (нет однозначного решения) или найденное рациональное значение свойства объекта не устраивает исследователя (слишком низкое или высокое). Поэтому возникает задача проведения поиска рациональных значений входных факторов в неисследованной области факторного пространства при минимальном количестве дополнительных опытов.

При традиционном поиске рациональных условий все факторы в эксперименте, кроме одного, поддерживают на постоянном уровне. При этом зачастую обнаруживается некоторое экстремальное значение функции отклика объекта для данного постоянного значения фактора (частный экстремум), которое может не совпадать с максимальным значением этой функции (рис. 7).

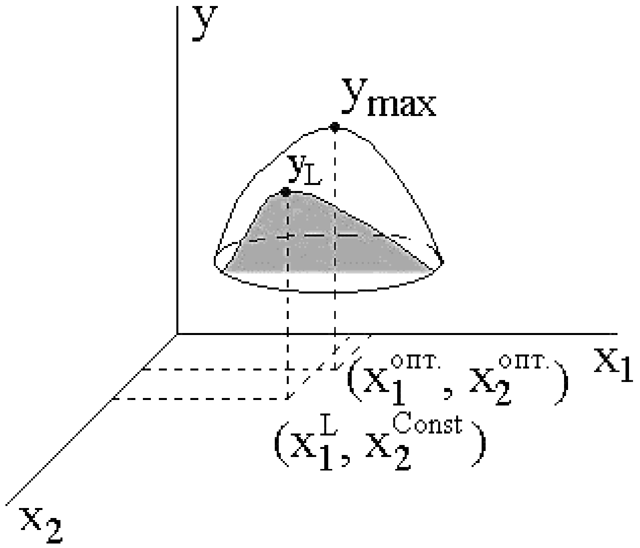


Рис. 7. Сечение поверхности отклика объекта y при постоянном значении фактора x_2

Для получения более полной информации о нахождении экстремальных значений свойств объекта рекомендуется проводить поиск рациональных условий при одновременном изменении нескольких факторов, используя специальные методы математического планирования эксперимента.

Все методы математического планирования эксперимента для решения оптимизационных задач делят на две группы:

- методы, требующие знаний уравнения регрессии функции отклика объекта;
- методы, не требующие таких знаний.

К группе методов, требующих знаний уравнения регрессии функции отклика объекта, относятся:

– метод крутого восхождения или наискорейшего спуска по поверхности функции отклика объекта;

– метод обобщенной функции желательности;

– и др.

К методу, не требующему знания уравнения регрессии, относится метод симплекс-планирования эксперимента.

При экспериментальном поиске экстремума свойств объекта следует всегда предполагать, что найденный экстремум будет локальным, а не глобальным. Глобальным экстремумом называют такой, для которого характерно самое максимальное (или минимальное) значение свойства объекта (рис. 8).

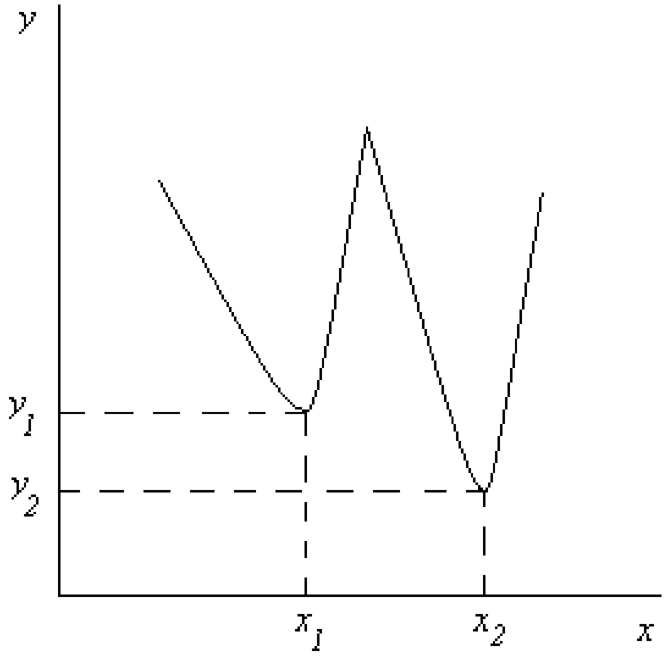


Рис. 8. Значения входных факторов x_1 и x_2 для локального (Y_1) и глобального (Y_2) экстремумов минимального значения функции отклика объекта

Для поиска глобального экстремума (который может и не существовать!) требуется выполнение дополнительных опытов в другой области факторного пространства.

2.1. Метод крутого восхождения или наискорейшего спуска по поверхности функции отклика объекта

Применение данного метода требует знания функции отклика объекта в виде исходного уравнения регрессии. Для этого выбирают входные факторы и области их изменения (Z_i^0, Z_i), планируют и реализуют план эксперимента первого или второго порядка.

После нахождения исходного уравнения регрессии определяют шаг и направление поиска экстремума функции отклика объекта. Для этого один из факторов, оказывающий наибольшее влияние на Y, принимают за базовый и для него выбирают шаг движения, который должен быть не больше шага варьирования этого фактора при получении исходного уравнения регрессии ΔZ_i, например, если за базовый фактор взять Z_1, то для базового шага движения ΔZ_i^* должно выполняться соотношение

$$\Delta Z_1^* \le \Delta Z_1.$$

Шаги движения остальных факторов рассчитывают следующим образом:

$$\delta = \frac{\Delta Z_1^*}{b_1 \Delta Z_1}; \Delta Z_1^* = \delta b_i Z_i,$$

где b_i – линейные эффекты факторов (с учетом их знаков!) в исходном уравнении регрессии.

Движение к экстремуму функции отклика объекта начинают из центра плана эксперимента, использованного при получении исходного уравнения регрессии ($Z_i = Z_i^0 \Delta Z_j^{0*}$, $x_i = 0$).

При поиске ***максимума*** функции отклика объекта (**метод крутого восхождения**) для определения условий проведения последующего опыта к координатам предыдущего опыта в факторном пространстве **прибавляют** шаги движения по каждому фактору. При поиске ***минимума*** (**метод наискорейшего спуска**) из координат предыдущего опыта **вычитают** шаги движения.

Движение к экстремуму прекращают в следующих случаях.

1. Значения функции отклика объекта или хотя бы одного фактора вышли за пределы допустимых значений.
2. Найден экстремум (возможно, локальный) функции отклика объекта.

В первом случае оптимизацию заканчивают. Во втором случае проводят дополнительные эксперименты по получению нового уравнения регрессии в области обнаруженного экстремума (рис. 9) и поиску нового экстремума на основе нового уравнения регрессии.

К недостаткам метода крутого восхождения или наискорейшего спуска следует отнести значительное число экспериментов при его реализации.

С конкретными примерами применения метода крутого восхождения или наискорейшего спуска можно познакомиться в литературе, например в книгах [3, 4].

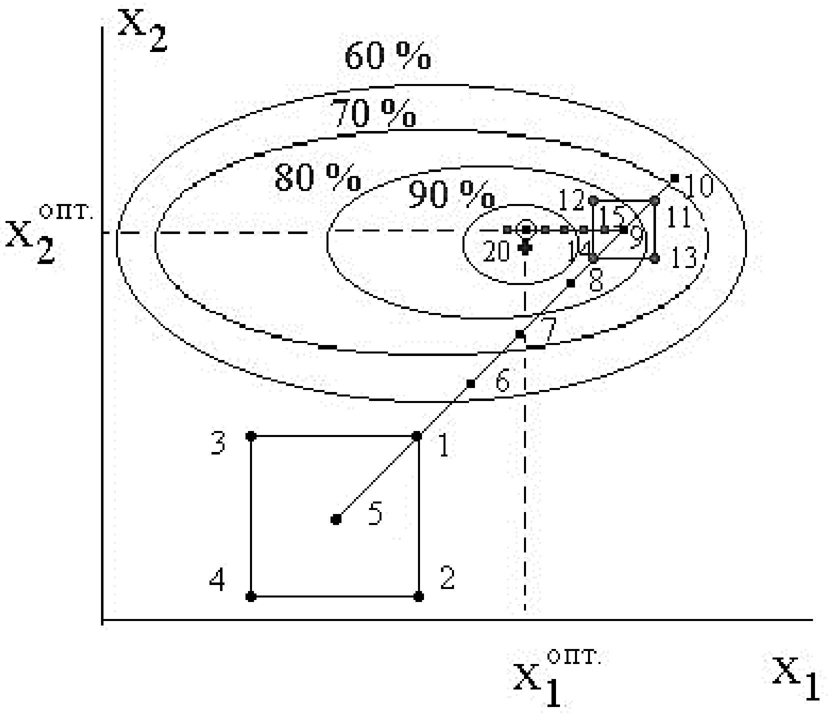


Рис. 9. Крутое восхождение по поверхности функции отклика объекта

Название метода произошло от названия геометрической фигуры «регулярный симплекс», т. е. правильный выпуклый многогранник. Если исследуется влияние на свойство объекта k факторов, то факторное пространство можно задать в виде регулярного симплекса с $(k + 1)$ вершиной. Так, например, для двух факторов $(k = 2)$ факторное пространство можно задать в виде правильного треугольника (правильный многогранник с 3 вершинами), а для трех факторов $(k = 3)$ – в виде тетраэдра (правильный многогранник с 4 вершинами).

2.2. Метод симплекс-планирования

Разработаны простые в построении матрицы планов эксперимента с использованием для исследований факторного пространства регулярных симплексов. Центр регулярных симплексов совпадает с центром факторного пространства (координаты центра начального плана эксперимента задаются в кодированных значениях как $x_i = 0$). Одна из вершин исходного симплекса лежит на одной из координатных осей этого пространства. Для такого случая построение плана эксперимента начинается с составления матрицы E с кодированными значениями факторов (табл. 12).

Таблица 12

Исходный симплекс-план эксперимента
с кодированными значениями факторов

Номер опыта *n*	Кодированные значения факторов								*Y*
	x_1	x_2	x_3	…	x_i	…	x_{k-1}	x_k	
1	e_1	e_2	e_3	…	e_i	…	e_{k-1}	e_k	
2	$-1e_1$	e_2	e_3	…	e_i	…	e_{k-1}	e_k	
3	0	$-2e_2$	e_3	…	e_i	…	e_{k-1}	e_k	
4	0	0	$-3e_3$	…	e_i	…	e_{k-1}	e_k	
…	0	0	0	…	e_i	…	e_{k-1}	e_k	
$n+1$	0	0	0	0	$-ne_i$	…	e_{k-1}	e_k	
…	0	0	0	0	0	…	e_{k-1}	e_k	
k	0	0	0	0	0	0	$-(k-1)e_{k-1}$	e_k	
$k+1$	0	0	0	0	0	0	0	$-ke_k$	

При длине стороны симплекса, равной 1, конкретные цифровые значения x_i рассчитываются по формуле

$$x_i = e_i = \sqrt{\frac{1}{2i\left(i+1\right)}}.$$

Высота такого симплекса h_k, равная расстоянию от вершины до противоположной грани, равна:

$$h_k = \frac{k+1}{\sqrt{2k(k+1)}}.$$

Число опытов в исходном симплекс-плане N всегда на единицу больше числа исследуемых факторов k ($N = k + 1$).

После завершения эксперимента по исходному симплекс-плану оценивают полученные значения отклика объекта и определяют номер опыта с наихудшим для исследователя значением Y (минимальным или максимальным в зависимости от цели оптимизации). По координатам наихудшего опыта h рассчитывают координаты нового дополнительного ($k + 2$) опыта:

$$x_i^{(k+2)} = 2x_i^{(c)} - x_i^{(h)};\ x_i^{(c)} = \frac{\sum_{n=1}^{k+1} x_j^{(i)}}{k} \text{ при } n \neq h.$$

После реализации нового опыта ($k + 2$) вновь анализируют значения Y и, если значение Y^{k+2} лучше значения Y^h, снова определяют номер наихудшего опыта во всей совокупности проведенных опытов (исключая из рассмотрения опыт h) и рассчитывают координаты следующего дополнительного ($k + 3$) опыта. Если новый опыт не приводит к получению лучшего значения Y (симплекс «зацикливается»), то эксперименты заканчивают и за рациональное значение факторов принимают координаты опыта, в котором получено наилучшее значение Y.

Визуально процедура поиска экстремума функции отклика объекта симплекс-методом представлена на рис. 10.

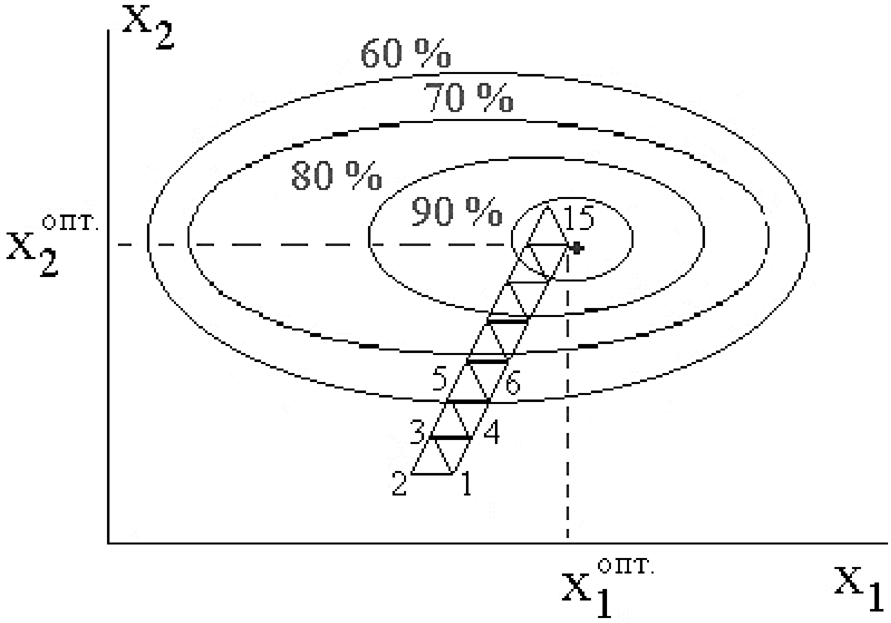


Рис. 10. Поиск рациональных условий симплекс-методом

Достоинства метода симплекс-планирования

1. Экономное число опытов из-за необязательности их повторения.

2. Возможность увеличения числа исследуемых факторов на любом этапе исследований без значительного увеличения числа опытов.

3. Возможность при оптимизации одного свойства объекта учитывать изменения других свойств.

4. Возможность получения уравнения регрессии по симплекс-планам при $k = 4a - 1$, где a равно 1, 2, 3, …, n.

Недостатки метода симплекс-планирования:

- позволяет найти только один экстремум функции отклика объекта, а для поиска других экстремумов необходимо повторять реализацию исходного симплекс-плана в другой области факторного пространства;
- эффективность поиска экстремумов функции отклика объекта зависит от величины выбранного интервала варьирования факторов.

2.3. Особенности планирования эксперимента в производственных условиях

При переходе от лабораторных условий к промышленным результаты исследований довольно часто плохо воспроизводятся в основном из-за резко увеличивающегося числа случайных факторов. Поэтому приходится продолжать эксперименты и при промышленном производстве продукции.

С целью уменьшения доли выпуска бракованной продукции при промышленных экспериментах применяют небольшие интервалы варьирования факторов по сравнению с лабораторными исследованиями. Планы промышленных экспериментов составляют таким образом, чтобы путем осторожного изменения факторов получить информацию об изменениях отклика объекта и при этом выпустить качественную продукцию.

Такой подход к планированию промышленных экспериментов был предложен Боксом и получил название «метод эволюционного планирования эксперимента», или адаптационной оптимизации [7].

Рассмотренные методы математического планирования эксперимента и обработки экспериментальных данных не являются обязательными для научных исследований в химии и химической технологии. К выбору методов планирования эксперимента и обработке его результатов необходимо подходить осознанно исходя из поставленных целей и задач. В каждом конкретном исследовании могут применяться специальные методы. Так, например, метод симплекс-планирования широко применяется не только для решения оптимизационных задач, но и при изучении зависимостей свойств смесей химических веществ от их состава, в том числе при составлении диаграмм «Состав – свойство» [4].

Наука «Математика» продолжает искать и разрабатывать новые методы планирования и анализа результатов эксперимента. Поэтому нужно не останавливаться в изучении и освоении методов математического планирования эксперимента и обработки экспериментальных данных.

Для математического планирования эксперимента используется большое число разнообразных и достаточно непростых приемов математической комбинаторики. При письменном составлении планов эксперимента с использованием приемов математической комбинаторики могут допускаться ошибки. Также ошибки могут быть допущены экспериментатором при анализе результатов эксперимента различными методами математической статистики. Вероятность совершения математических ошибок при математическом планировании эксперимента значительно меньше при использовании для этих целей авторитетных и популярных компьютерных программ, проверенных на практике математиками и экспериментаторами.

3. ПРИМЕНЕНИЕ КОМПЬЮТЕРНЫХ ПРОГРАММ ДЛЯ СОСТАВЛЕНИЯ МАТЕМАТИЧЕСКИХ ПЛАНОВ ЭКСПЕРИМЕНТА И АНАЛИЗА ИХ РЕЗУЛЬТАТОВ

3.1. Применение компьютерных программ для составления математических планов эксперимента

В настоящее время составление математических планов эксперимента достаточно полно представлено в пакетах компьютерных программ Statistica и Statgraphics Centurion различных версий.

Рассмотрим работу пакета программ Statistica (версия 10) для компьютерного составления математических планов эксперимента для достижения различных целей после их реализации. Компьютерное составление математических планов эксперимента в пакете компьютерных программ Statgraphics Centurion описано в книгах [3, 4].

В данном пособии рассматривается работа с пакетом компьютерных программ Statistica версии 10.0 на русском языке (далее – Statistica).

В Statistica предусмотрена возможность настройки множества характеристик и интерфейса программы в соответствии с предпочтениями пользователя. Например, пользователь может установить запись результатов своих действий в Statistica в виде файлов рабочих книг, отчетов различного формата в меню «Сервис» в разделе «Параметры» [8].

3.1.1. Применение пакета программ Statistica для составления математических планов дробного факторного эксперимента Плакетта – Бермана

Запускаем на компьютере в пакете компьютерных программ Statistica версии 10.0 командный файл STATISTICA 10 RU. На экране появляется стартовая панель (рис. 11), и на ней открываем меню «Анализ» (рис. 12).

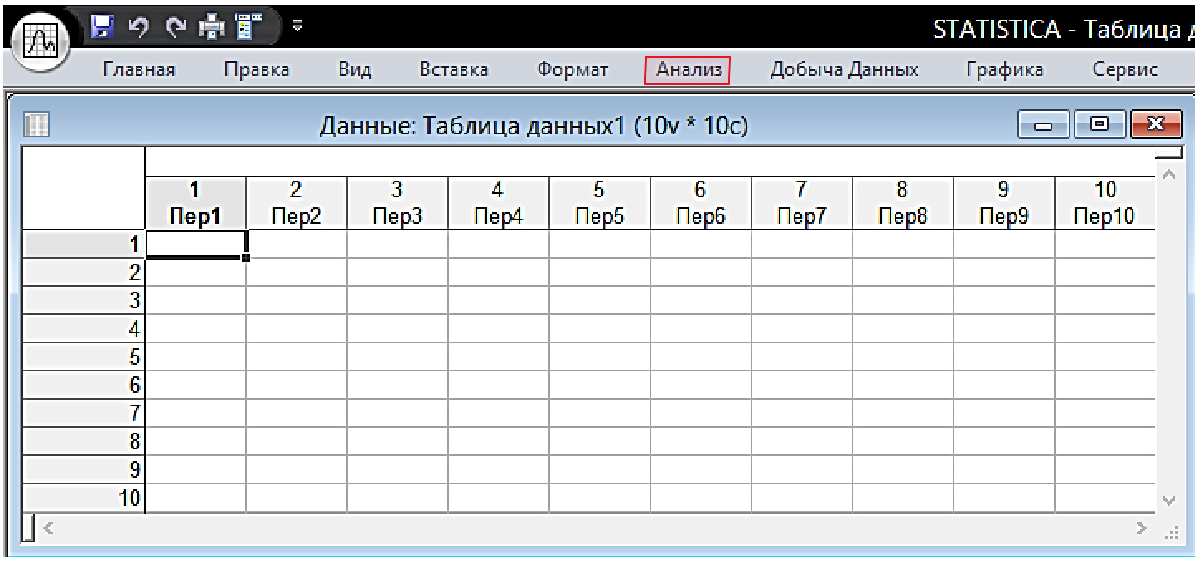


Рис. 11. Экран компьютера со стартовой панелью при запуске Statistica

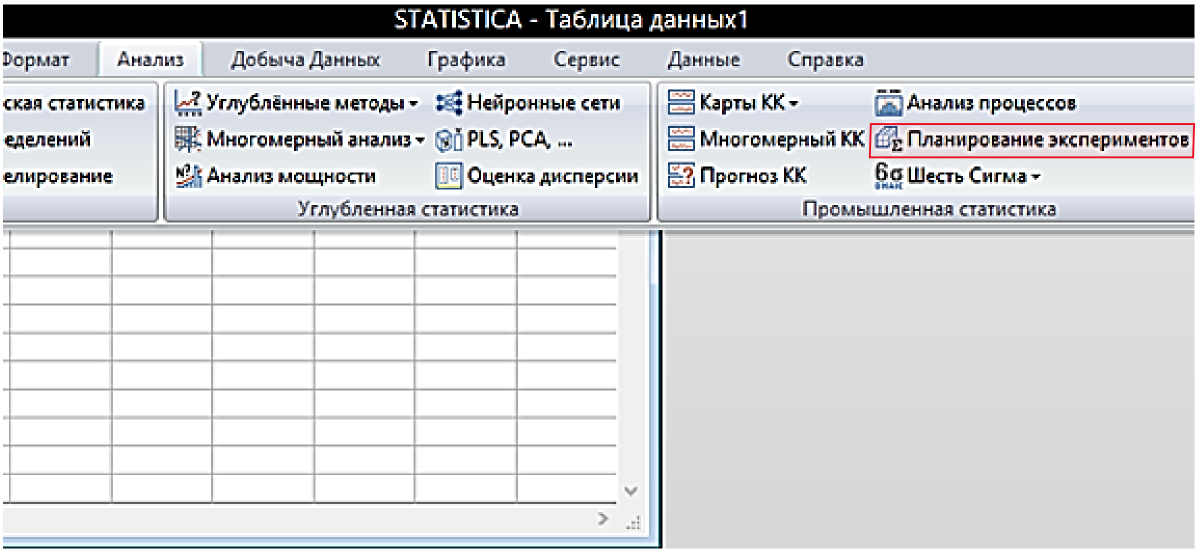
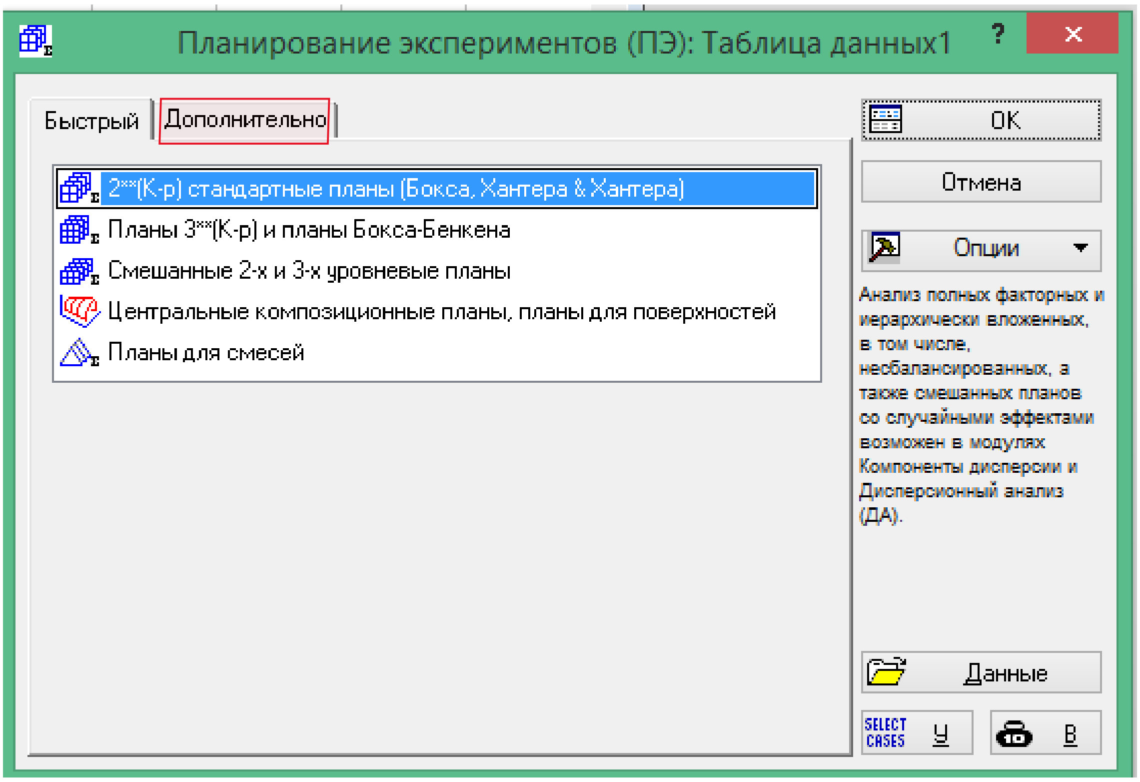


Рис. 12. Экран компьютера после открытия меню «Анализ»

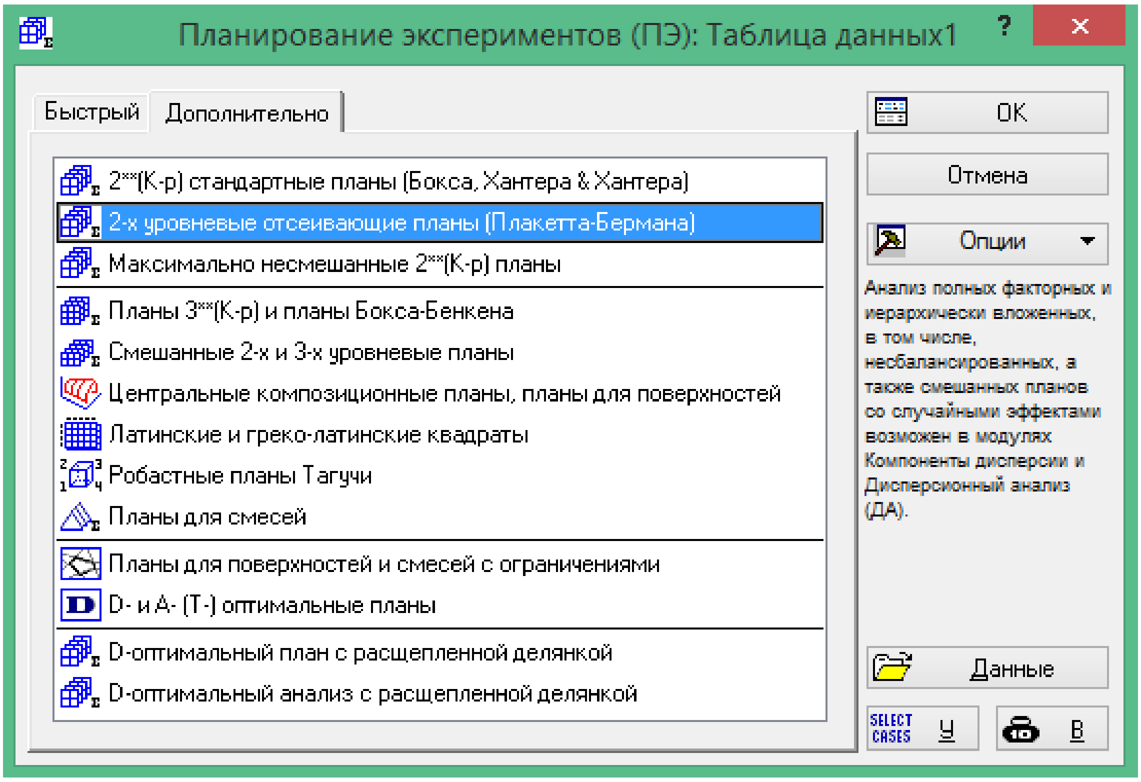
В появившемся окне «Планирование эксперимента» (рис. 13) нажимаем кнопку «Дополнительно».

В новом окне (рис. 14) выбрав двухуровневые отсеивающие планы (Плакетта – Бермана) и нажав кнопку «ОК», переходим в раздел «Построение плана» (рис. 15).



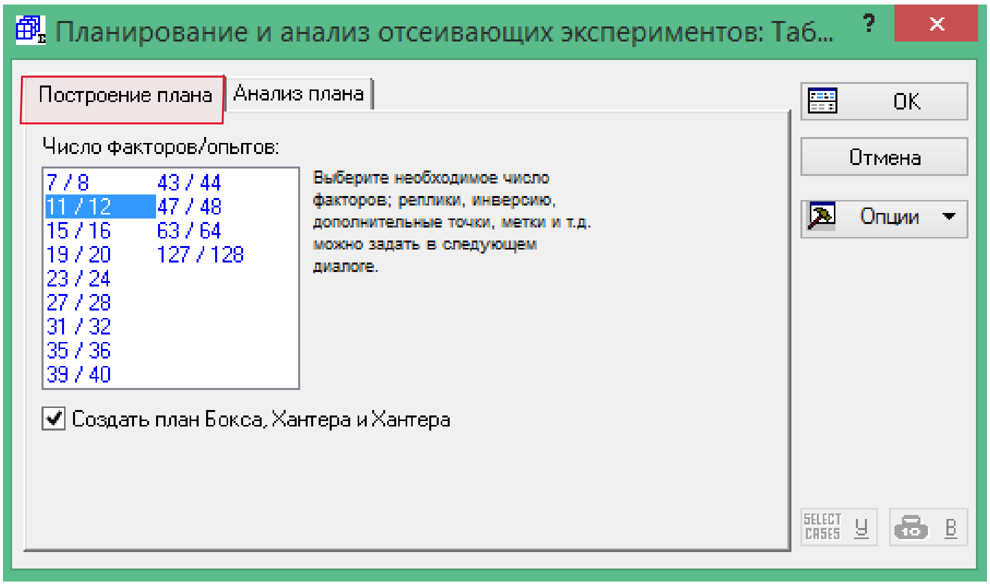
Анализ полных факторных и иерархически вложенных, в том числе, несбалансированных, а также смешанных планов со случайными эффектами возможен в модулях Компоненты дисперсии и Дисперсионный анализ (ДА).

Рис. 13. Экран компьютера после открытия меню «Планирование экспериментов»



Анализ полных факторных и иерархически вложенных, в том числе, несбалансированных, а также смешанных планов со случайными эффектами возможен в модулях Компоненты дисперсии и Дисперсионный анализ (ДА).

Рис. 14. Экран компьютера после открытия в меню «Планирование экспериментов» раздела «Дополнительно»



Планирование и анализ отсеивающих экспериментов: Таб...

Число факторов/опытов:

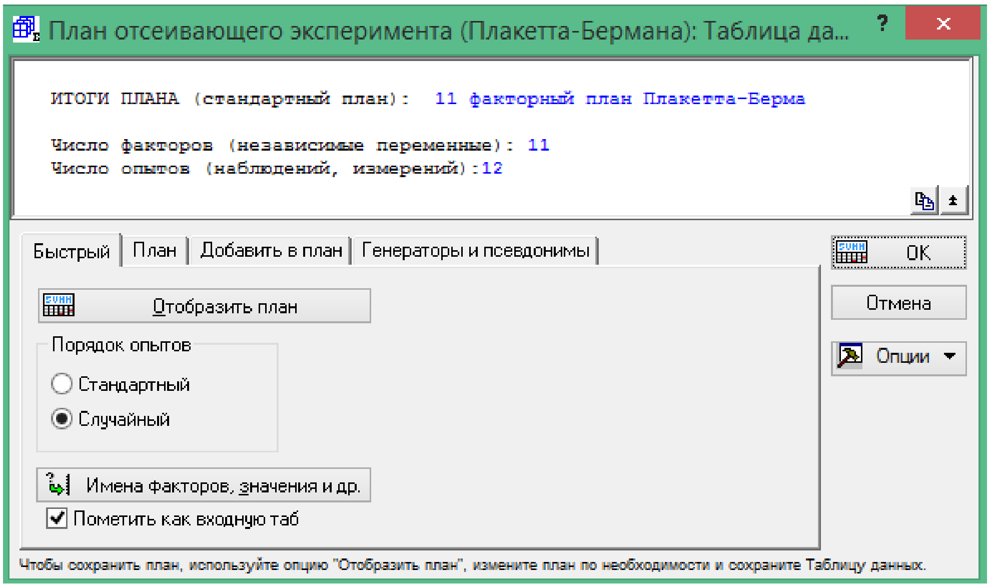
7 / 8	43 / 44
11 / 12	47 / 48
15 / 16	63 / 64
19 / 20	127 / 128
23 / 24	
27 / 28	
31 / 32	
35 / 36	
39 / 40	

Выберите необходимое число факторов; реплики, инверсию, дополнительные точки, метки и т.д. можно задать в следующем диалоге.

☑ Создать план Бокса, Хантера и Хантера

Рис. 15. Выбор числа опытов в плане Плакетта – Бермана

Выбираем, например, план для 11 входных факторов (12 опытов) и, нажав кнопку «ОК», переходим в раздел отображения плана Плакетта – Бермана (рис. 16), где определяем порядок последовательности проведения опытов в эксперименте (стандартный или случайный).



План отсеивающего эксперимента (Плакетта-Бермана): Таблица да...

```
ИТОГИ ПЛАНА (стандартный план):  11 факторный план Плакетта-Берма

Число факторов (независимые переменные): 11
Число опытов (наблюдений, измерений):12
```

Порядок опытов
○ Стандартный
● Случайный

☑ Пометить как входную таб

Чтобы сохранить план, используйте опцию "Отобразить план", измените план по необходимости и сохраните Таблицу данных.

Рис. 16. Выбор порядка опытов в плане Плакетта – Бермана

По умолчанию предлагается выбранный план эксперимента сформировать в форме вводной таблицы для проведения статистического анализа результатов эксперимента в Statistica. По умолчанию имена факторов в Statistica обозначаются прописными английскими буквами A, B, C, D и т. д., а кодированные значения этих факторов цифрами –1 и 1. Нажав кнопку «Имена факторов, значения и др.» (см. рис. 16), можно изменить их имена и значения, например, на необходимые имена и натуральные значения входных факторов.

Нажав кнопку «ОК» (см. рис. 16), получаем составленный план эксперимента Плакетта – Бермана для 12 опытов с 11 входными факторами с их кодированными значениями (рис. 17), который в Statistica называется «Таблица данных». Когда закроем файл таблицы данных «Рабочая книга…» (см. рис. 17), появится вопрос от программы «Сохранить изменения в Рабочая книга…?» Ответив на этот вопрос утвердительно, можно сохранить файл с составленным планом эксперимента.

Структура таблицы данных Statistica совпадает со структурой таблиц MS Excel, что упрощает перемещение содержания этих таблиц из одной программы в другую. Таблицы данных Statistica с планами Плакетта – Бермана могут быть использованы в ней для проведения многофакторного дисперсионного и регрессионного анализов результатов экспериментов.

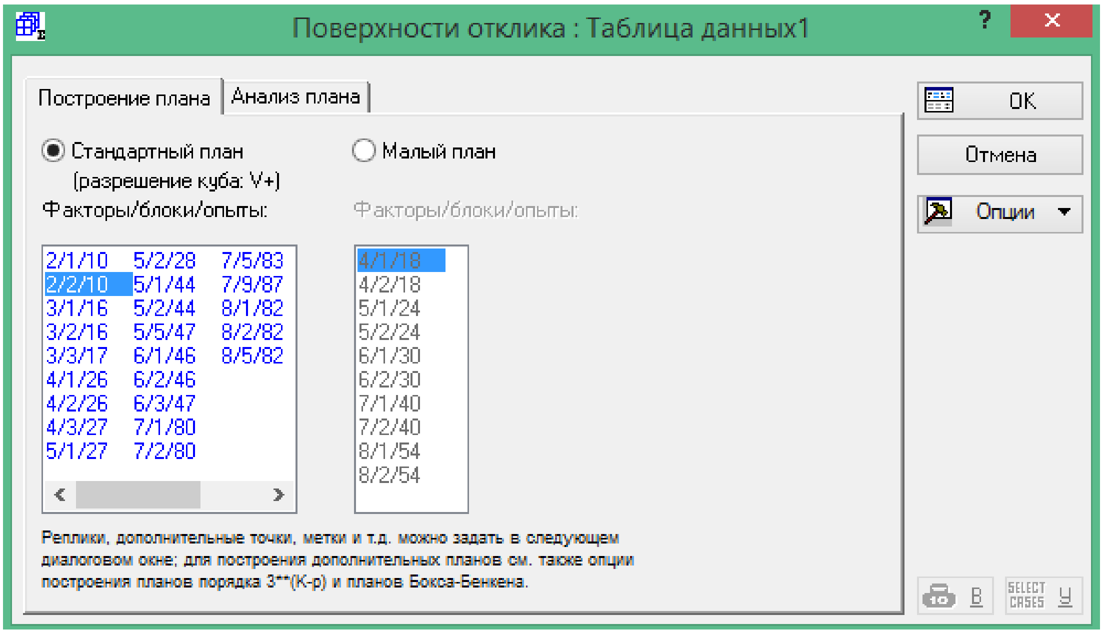
Планир. план	План: 11 факторный отсеив. план (Таблица данных1)					
	F	G	H	I	J	K
6	1,00000	-1,00000	1,00000	-1,00000	-1,00000	-1,00000
2	-1,00000	-1,00000	1,00000	1,00000	1,00000	-1,00000
9	1,00000	-1,00000	1,00000	1,00000	-1,00000	1,00000
11	1,00000	1,00000	1,00000	-1,00000	1,00000	1,00000
8	-1,00000	1,00000	1,00000	-1,00000	1,00000	-1,00000
1	-1,00000	1,00000	1,00000	1,00000	-1,00000	1,00000
4	1,00000	-1,00000	-1,00000	-1,00000	1,00000	1,00000
12	-1,00000	-1,00000	-1,00000	-1,00000	-1,00000	-1,00000
3	-1,00000	-1,00000	-1,00000	1,00000	1,00000	1,00000
10	1,00000	1,00000	-1,00000	1,00000	1,00000	-1,00000
5	-1,00000	1,00000	-1,00000	-1,00000	-1,00000	1,00000
7	1,00000	1,00000	-1,00000	1,00000	-1,00000	-1,00000

Рис. 17. План Плакетта – Бермана для двенадцати опытов
(11 входных факторов с их кодированными значениями) со случайным
порядком последовательности проведения опытов в эксперименте

Полученный в Statistica план Плакетта – Бермана для 11 входных факторов с их кодированными значениями со случайным порядком последовательности проведения опытов в эксперименте аналогичен плану, представленному в табл. 3 со стандартным порядком последовательности проведения опытов в эксперименте.

3.1.2. Применение пакета программ Statistica для составления математических планов эксперимента второго порядка

Выполняем в Statistica действия, описанные выше и представленные на рис. 12–14. На экране компьютера (см. рис. 14) математические планы второго порядка полного и дробного факторного эксперимента находятся в разделе «Центральные композиционные планы, планы для поверхностей». Для построения рототабельного плана Бокса – Хантера полного факторного эксперимента с двумя опытами в центре плана на экране монитора (см. рис. 14) в разделе «Центральные композиционные планы, планы для поверхностей» выбираем «Центральные композиционные планы, планы для поверхностей» и нажимаем кнопку «ОК». В появившемся окне (рис. 18) выбираем стандартный план 2/2/10 (что означает план для двух входных факторов с двукратным повторением двух опытов и десятью опытами) и нажимаем кнопку «ОК».



● Стандартный план (разрешение куба: V+)

Факторы/блоки/опыты:

2/1/10
2/2/10
3/1/16
3/2/16
3/3/17
4/1/26
4/2/26
4/3/27
5/1/27
5/2/28
5/1/44
5/2/44
5/5/47
6/1/46
6/2/46
6/3/47
7/1/80
7/2/80
7/5/83
7/9/87
8/1/82
8/2/82
8/5/82

○ Малый план

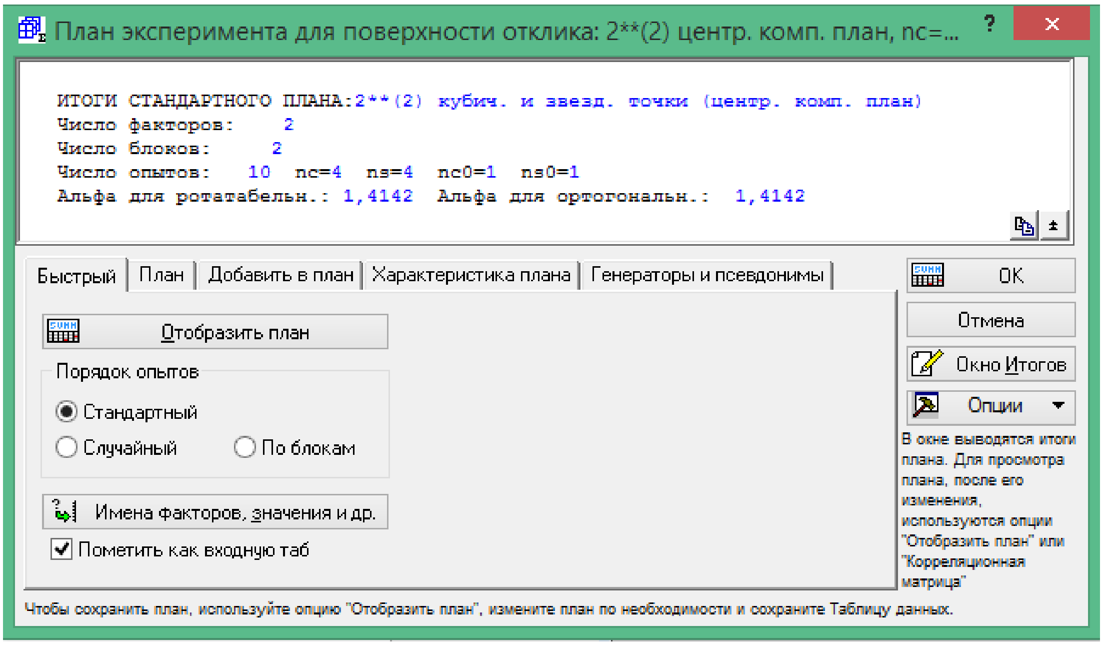
Факторы/блоки/опыты:

4/1/18
4/2/18
5/1/24
5/2/24
6/1/30
6/2/30
7/1/40
7/2/40
8/1/54
8/2/54

Реплики, дополнительные точки, метки и т.д. можно задать в следующем диалоговом окне; для построения дополнительных планов см. также опции построения планов порядка 3**(K-p) и планов Бокса-Бенкена.

Рис. 18. Выбор плана Бокса – Хантера полного факторного эксперимента для двух входных факторов с двукратным повторением двух опытов и десятью опытами

В появившемся окне (рис. 19) выбираем план со стандартным порядком опытов и нажимаем кнопку «ОК». Составленный в Statistica рототабельный план Бокса – Хантера полного факторного эксперимента для двух входных факторов с двумя опытами в центре плана представлен на рис. 20.



```
ИТОГИ СТАНДАРТНОГО ПЛАНА:2**(2) кубич. и звезд. точки (центр. комп. план)
Число факторов:     2
Число блоков:     2
Число опытов:   10  nc=4  ns=4  nc0=1  ns0=1
Альфа для ротатабельн.: 1,4142  Альфа для ортогональн.:  1,4142
```

Порядок опытов

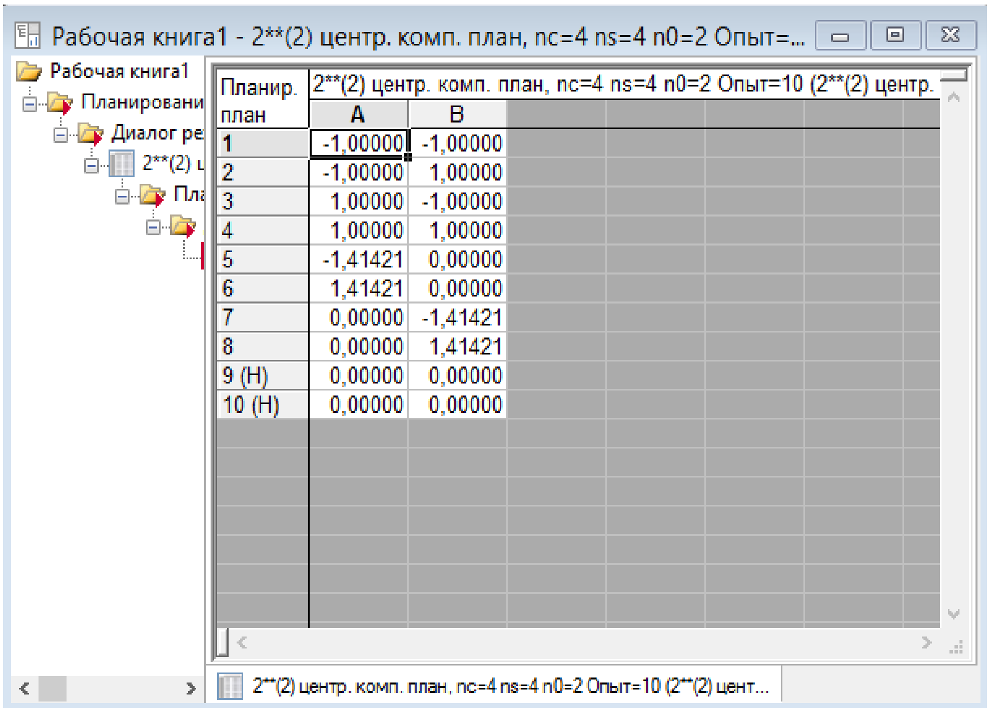
● Стандартный
○ Случайный ○ По блокам

☑ Пометить как входную таб

В окне выводятся итоги плана. Для просмотра плана, после его изменения, используются опции "Отобразить план" или "Корреляционная матрица"

Чтобы сохранить план, используйте опцию "Отобразить план", измените план по необходимости и сохраните Таблицу данных.

Рис. 19. Выбор стандартного плана Бокса – Хантера



2**(2) центр. комп. план, nc=4 ns=4 n0=2 Опыт=10 (2**(2) центр.

Планир. план	A	B
1	-1,00000	-1,00000
2	-1,00000	1,00000
3	1,00000	-1,00000
4	1,00000	1,00000
5	-1,41421	0,00000
6	1,41421	0,00000
7	0,00000	-1,41421
8	0,00000	1,41421
9 (H)	0,00000	0,00000
10 (H)	0,00000	0,00000

Рис. 20. Рототабельный план Бокса – Хантера
полного факторного эксперимента для двух входных факторов
с двукратным повторением двух опытов

Полученный план Бокса – Хантера отличается от плана Бокса – Уилсона для такого же числа входных факторов и числа опытов в центре факторной области (см. табл. 11) только большей величиной «звездного плеча», т. е. значением «звездных точек» (1,41421).

4. ПРИМЕНЕНИЕ КОМПЬЮТЕРНЫХ ПРОГРАММ ДЛЯ СТАТИСТИЧЕСКОГО АНАЛИЗА РЕЗУЛЬТАТОВ ЭКСПЕРИМЕНТА

Для статистического анализа данных (в том числе и результатов эксперимента) разработано и используется большое количество компьютерных программ. Рассмотрим работу программы Microsoft Excel пакета Microsoft Office 365 (версия 16) и пакета программ Statistica (версия 10) для расчета параметров статистического анализа результатов экспериментов.

4.1. Расчет в программах Microsoft Excel и Statistica параметров для статистического анализа случайных ошибок измерений результатов эксперимента

Математические расчеты в широко применяемых компьютерных программах являются безошибочными и автоматизированными. Прежде чем приступать к статистическому анализу количественных результатов экспериментов, полученных с помощью измерений, необходимо по требованиям математической статистики рассчитать и оценить случайные ошибки и погрешности измерений.

Расчеты параметров и статистический анализ случайных ошибок измерений экспериментальных данных выполняют следующим образом [1]:

- проверяют наличие грубых ошибок прямых многократных измерений результата эксперимента Y при числе повторных измерений не менее 3 по критерию Q для числа повторных измерений n и заданной доверительной вероятности P;
- исключают из полученных экспериментальных данных Y_n результаты, являющиеся грубыми ошибками прямых измерений для заданной вероятности P;
- для результатов неоднократных прямых измерений, не являющихся грубыми ошибками, рассчитывают их среднее арифметическое значение $\overline{Y}$, выборочную дисперсию единичных значений S_Y^2, выборочное абсолютное стандартное отклонение единичных значений S_Y и другие статистические параметры.

Исходя из величин случайных и систематических ошибок измерений, обусловленных точностью использованных средств измерений (класса точности приборов или минимальных значений шкал измерений этих приборов), рассчитывают погрешности измерений.

В MS Excel есть функции (математические, статистические), автоматизирующие расчеты некоторых параметров для статистического анализа случайных ошибок измерений экспериментальных данных. Например, такими функциями для автоматизации расчетов являются:

- СРЗНАЧ (расчет среднего арифметического значения ряда чисел);
- ДИСП.В (расчет выборочной дисперсии единичных значений);
- СТЬЮДЕНТ.ОБР.2Х (расчет квантиля распределения Стьюдента для 1–*P* и *n*–1).

Рассмотрим пример расчетов в MS Excel параметров для статистического анализа случайных ошибок измерений экспериментальных данных.

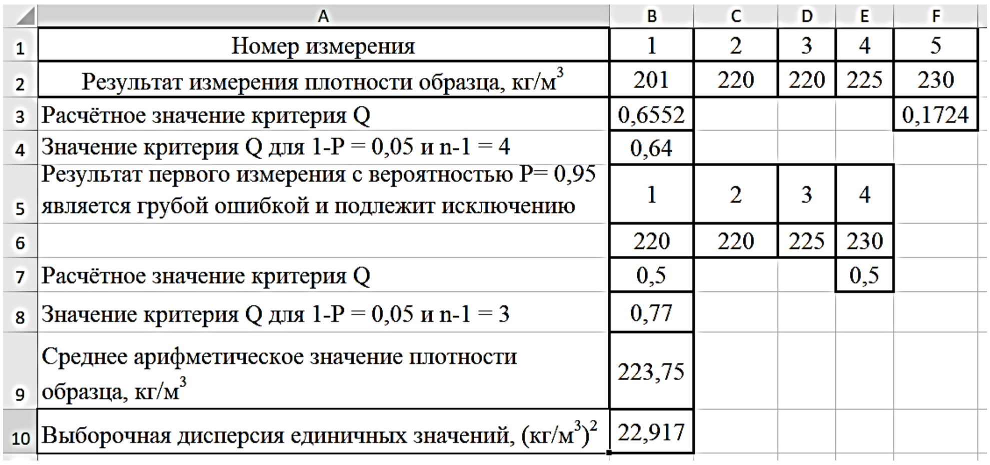
Пример 4. В эксперименте полученный полимерный композит был распилен на пять образцов и для каждого из этих образцов было проведено прямое измерение их плотности ядерным плотномером. На листе MS Excel в строку 2 (см. рис. 19) вводим в ячейки результаты измерений плотности образцов композита (Y_1 – Y_5) в виде возрастающего математического ряда значений.

На грубые ошибки проверяем оба крайних значения плотности композита (ячейки B2–F2). В строку 3 столбца А записываем название рассчитываемого параметра Q_p для первого и пятого результатов измерений (Q_1 и Q_5). Для доказательства с вероятностью P = 0,95 признания этих единичных результатах измерения площади образца композита (1 и 5) грубыми случайными ошибками проведенных измерений в ячейки B2 и F2 вводим формулу расчета параметра Q_p с адресами ячеек, содержащих необходимые исходные данные (введенные в MS Excel формулы расчетов приведены на рис. 20):

$$Q_p = \left|Y_n - Y_{n-1}\right| \;/\; \left|Y_n - Y_1\right|.$$

Для вероятности P = 0,95 (1– P = 0,05) и n = 5 (числа повторных измерений плотности образца композита) значение критерия Q равно 0,64 [9]. Для измерения № 1 полученное расчетное значение параметра Q_p = 0,6552 (рис. 21).

Так как оно больше значения критерия Q = 0,64, то этот результат измерения с вероятностью 0,95 следует считать грубой ошибкой и исключить его из дальнейшего рассмотрения.



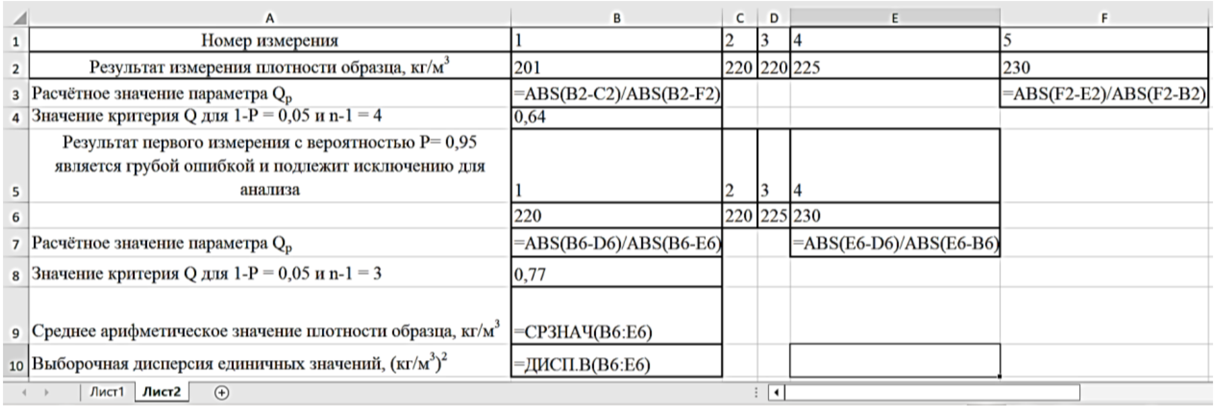
	A	B	C	D	E	F
1	Номер измерения	1	2	3	4	5
2	Результат измерения плотности образца, кг/м^3	201	220	220	225	230
3	Расчётное значение критерия Q	0,6552				0,1724
4	Значение критерия Q для 1-P = 0,05 и n-1 = 4	0,64				
5	Результат первого измерения с вероятностью P= 0,95 является грубой ошибкой и подлежит исключению	1	2	3	4	
6		220	220	225	230	
7	Расчётное значение критерия Q	0,5			0,5	
8	Значение критерия Q для 1-P = 0,05 и n-1 = 3	0,77				
9	Среднее арифметическое значение плотности образца, кг/м^3	223,75				
10	Выборочная дисперсия единичных значений, (кг/м^3)^2	22,917				

Рис. 21. Результаты расчетов в MS Excel параметров для обнаружения грубых ошибок измерений оценки их случайных ошибок

Оставшаяся совокупность четырех результатов измерений плотности композита с вероятностью 0,95 не содержит грубых ошибок, и для них были рассчитаны в MS Excel с помощью встроенных функций (рис. 22) в меню «Формулы» среднее арифметическое значение плотности композита и выборочная дисперсия единичных значений плотности $S_{\bar{Y}}^2$, необходимая для расчетов выборочного стандартного отклонения плотности $S_{\bar{Y}}$ и возможной предельной случайной ошибки ее измерения $\xi_{случ}, \bar{Y}, пр,$ по формуле

$$\xi_{случ}, \bar{Y}, пр = S_{\bar{Y}} \, | t_{P,n} |,$$

где $t_{P,n}$ – квантиль распределения Стьюдента, рассчитываемый в MS Excel по формуле статистической функции СТЬЮДЕНТ.ОБР.2Х $(1 – P; n – 1)$.



	A	B	C	D	E	F
1	Номер измерения	1	2	3	4	5
2	Результат измерения плотности образца, кг/м^3	201	220	220	225	230
3	Расчётное значение параметра Q_p	=ABS(B2-C2)/ABS(B2-F2)				=ABS(F2-E2)/ABS(F2-B2)
4	Значение критерия Q для 1-P = 0,05 и n-1 = 4	0,64				
5	Результат первого измерения с вероятностью P= 0,95 является грубой ошибкой и подлежит исключению для анализа	1	2	3	4	
6		220	220	225	230	
7	Расчётное значение параметра Q_p	=ABS(B6-D6)/ABS(B6-E6)			=ABS(E6-D6)/ABS(E6-B6)	
8	Значение критерия Q для 1-P = 0,05 и n-1 = 3	0,77				
9	Среднее арифметическое значение плотности образца, кг/м^3	=СРЗНАЧ(B6:E6)				
10	Выборочная дисперсия единичных значений, (кг/м^3)^2	=ДИСП.В(B6:E6)				

Рис. 22. Формулы расчетов в MS Excel параметров для обнаружения грубых ошибок измерений оценки их случайных ошибок в меню «Формулы» в разделе «Показать формулы»

Вышеназванные статистические функции программы MS Excel находятся и в описательной статистике системы Statistica.

4.2. Расчет в программах Microsoft Excel и Statistica параметров для корреляционного анализа результатов эксперимента

При корреляционном анализе результатов эксперимента учитывается, что между переменными (входными факторами и свойствами объектов *Y*) может проявляться функциональная связь как математическая функция зависимости свойства объекта от значений входного фактора, так и корреляционная связь, когда свойство объекта реагирует на величину входного фактора, изменяющуюся по своему закону (стохастическая связь). В качестве меры зависимости между переменными используют коэффициент парной линейной корреляции, изменяющийся в пределах от –1 до +1. Чем ближе значение коэффициента корреляции к 1, тем зависимость между переменными сильнее. Если коэффициент корреляции равен 0, то переменные не зависят друг от друга.

Для количественных переменных рассчитывают и анализируют коэффициент корреляции Пирсона *r*, а для качественных переменных – коэффициенты корреляции Спирмена, тау Кендала, коэффициент гамма. Коэффициент корреляции Спирмена рекомендуется использовать, если переменные количественные (закон распределения которых неизвестен или не является нормальным) и (или) качественные (порядковые). Коэффициент тау Кендела рекомендуется использовать, если хотя бы одна переменная качественная (порядковая). Коэффициент гамма рекомендуется использовать, если переменные содержат много повторяющихся значений [10].

Принято считать [10] линейную корреляцию:

- слабой, если $| r | \leq 0{,}25$;
- умеренной, если $0{,}25 \leq | r | \leq 0{,}75$;
- сильной, если $| r | \geq 0{,}75$.

В MS Excel для расчета коэффициента корреляции Пирсона для двух переменных используют либо функцию КОРРЕЛ в меню «Формулы», либо в меню «Данные» раздел «Анализ данных» подраздел «Корреляция».

В Statistica можно одновременно рассчитать коэффициенты корреляции Пирсона для нескольких переменных (получить корреляционную матрицу).

Пример 5. Рассмотрим пример применения Statistica для расчета параметров парного корреляционного анализа 4 пар переменных – влияния содержания в составе полимерного композита пластификатора дибутилфталата (ДБФ, мас. %) на его физико-механические свойства: плотность (Y_1, кг/м^3), твердость по Бринеллю (Y_2, МПа), число упругости (Y_3, %), пластичность (Y_4, %). Для этого вносим в таблицу данных Statistica результаты эксперимента (рис. 23).

	1 ДБФ	2 Y1	3 Y2	4 Y3	5 Y4	6 Пер6	7 Пер7	8 Пер8	9 Пер9	10 Пер10
1	0	1371	126.6	82.5	17.5					
2	6	1355	116.7	81	19					
3	6	1348	111	68.7	31.3					
4	9	1287	80.7	64.2	35.8					
5	13	1263	22.9	57.8	42.2					
6	16	1300	21	60.7	39.3					

Рис. 23. Таблица исходных данных

В меню «Анализ» открываем раздел «Основные статистики и таблицы» и после его открытия выделяем подраздел «Парные и частные корреляции» (рис. 24) и нажимаем кнопку «ОК».

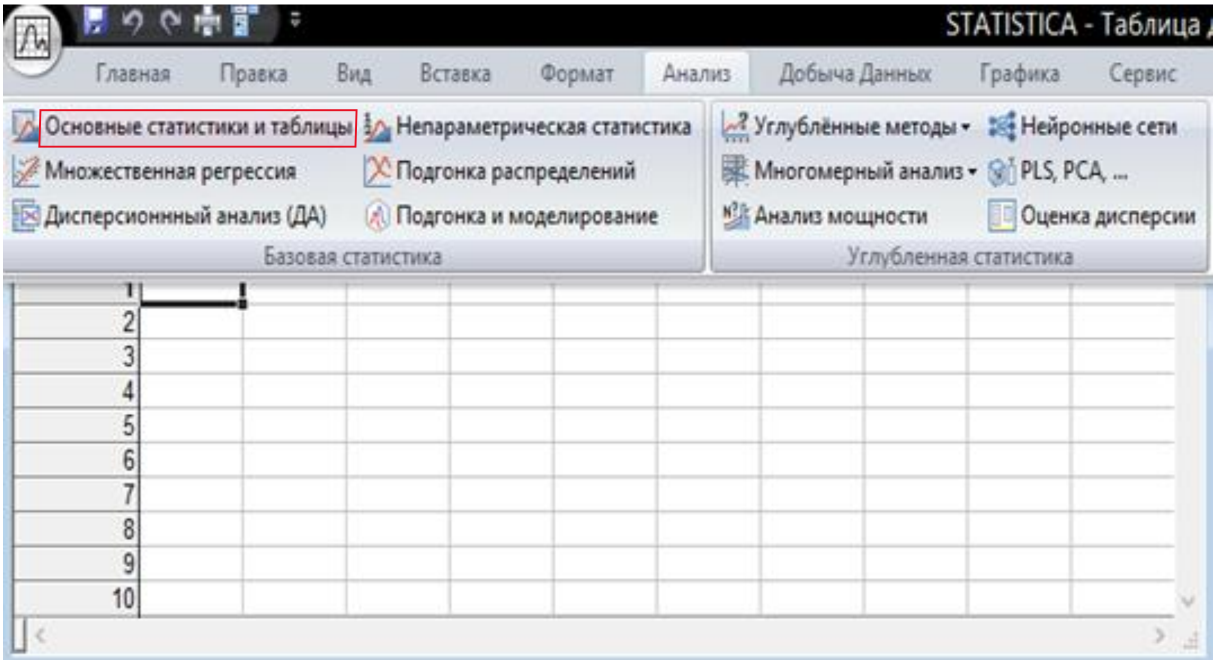


Рис. 24. Экран компьютера после открытия меню «Анализ»

В появившемся окне (рис. 25) выделяем подраздел «Парные и частные корреляции» и нажимаем кнопку «ОК».

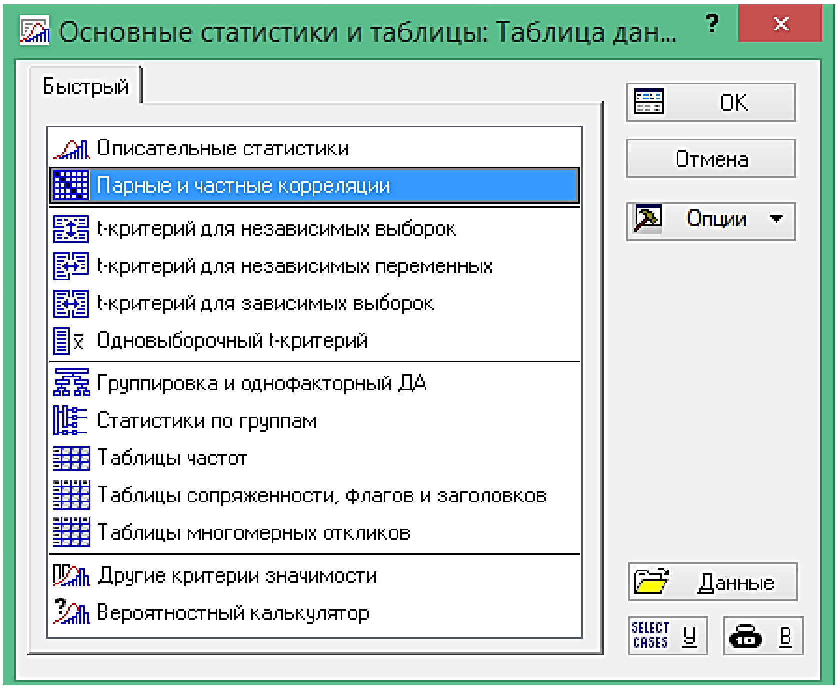
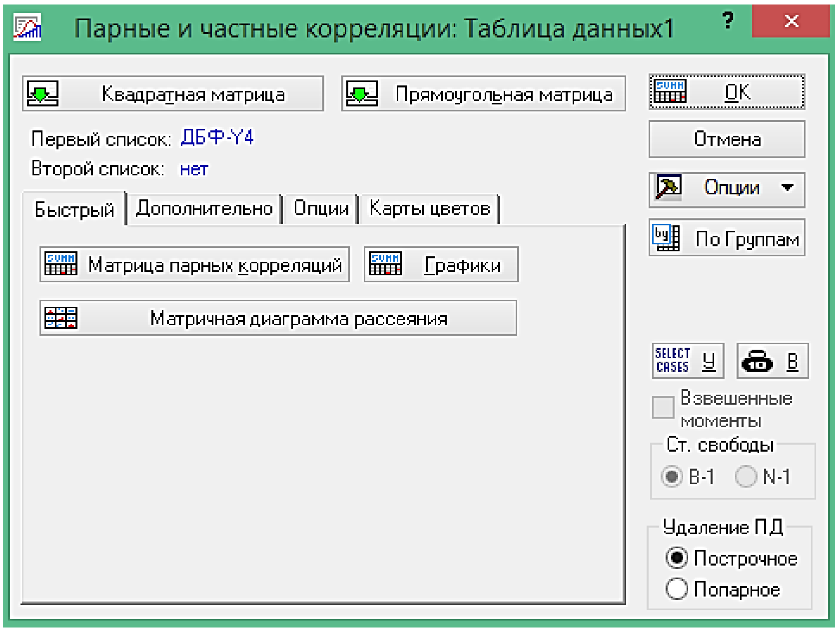


Рис. 25. Экран компьютера при открытии раздела «Основные статистики и таблицы»

В открывшемся окне (рис. 26) оставляем задаваемые параметры «Квадратная матрица» и «Матрица парных корреляций» и нажимаем кнопку «ОК».



Первый список: ДБФ-Y4
Второй список: нет

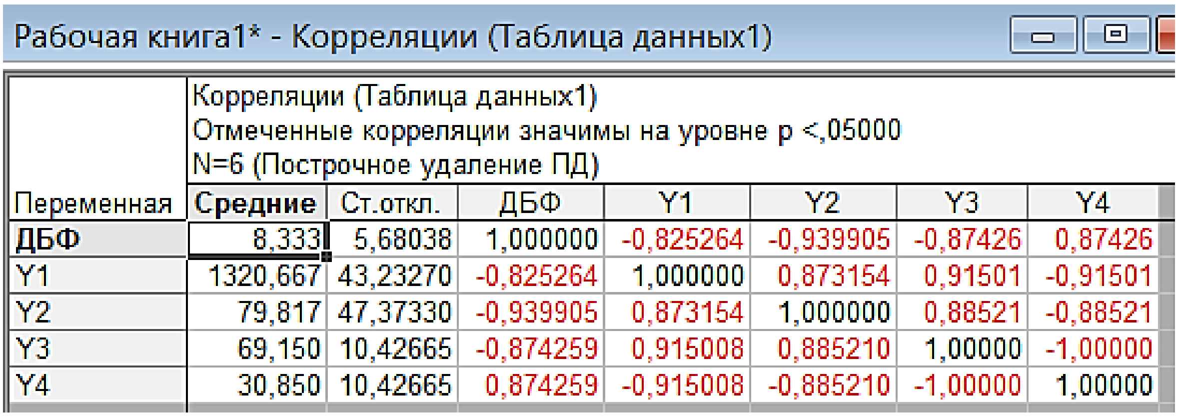
☐ Взвешенные моменты

Ст. свободы
● B-1 ○ N-1

Удаление ПД
● Построчное
○ Попарное

Рис. 26. Экран компьютера при открытии подраздела «Парные и частные корреляции»

Открывается таблица с набором рассчитанных коэффициентов парной линейной корреляции Пирсона (рис. 27).



Корреляции (Таблица данных1)
Отмеченные корреляции значимы на уровне p <,05000
N=6 (Построчное удаление ПД)

Переменная	Средние	Ст.откл.	ДБФ	Y1	Y2	Y3	Y4
ДБФ	8,333	5,68038	1,000000	-0,825264	-0,939905	-0,87426	0,87426
Y1	1320,667	43,23270	-0,825264	1,000000	0,873154	0,91501	-0,91501
Y2	79,817	47,37330	-0,939905	0,873154	1,000000	0,88521	-0,88521
Y3	69,150	10,42665	-0,874259	0,915008	0,885210	1,00000	-1,00000
Y4	30,850	10,42665	0,874259	-0,915008	-0,885210	-1,00000	1,00000

Рис. 27. Результаты расчета коэффициентов
парной линейной корреляции Пирсона

Из полученных результатов расчета коэффициентов парной линейной корреляции Пирсона (см. рис. 27) можно сделать вывод, что с вероятностью больше 0,95 (уровень значимости $p = 1 - P < 0{,}05000$) увеличение содержания пластификатора ДБФ от 0 до 16 мас. % в полимерном композите оказывает **сильное** влияние на все четыре показателя его исследованных физико-механических свойств ($r > 0{,}75$ выделяется в Statistica красным шрифтом). При этом с увеличением содержания пластификатора ДБФ у композита увеличивается пластичность, а остальные показатели свойств снижаются.

В MS Excel нет встроенных функций для расчета коэффициентов корреляции Спирмена, тау Кендала, коэффициента гамма. В Statistica такие возможности есть [10].

4.3. Расчет в программах Microsoft Excel и Statistica параметров для дисперсионного анализа результатов эксперимента

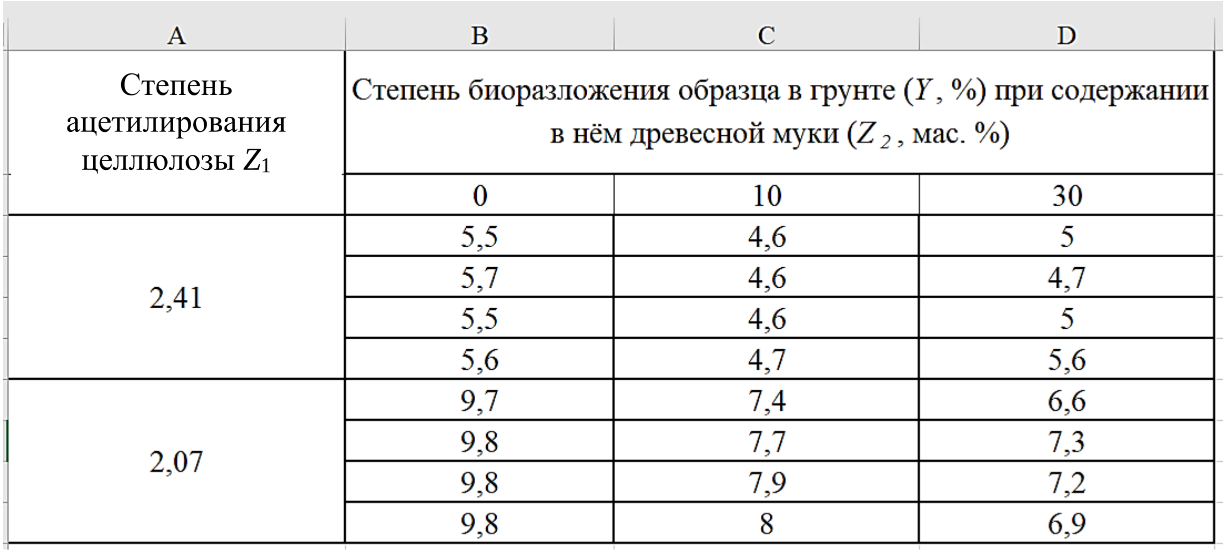
В MS Excel в меню «Данные» в разделе «Анализ данных» есть инструменты для дисперсионного анализа результатов эксперимента: «Однофакторный дисперсионный анализ», «Двухфакторный дисперсионный анализ без повторений» и «Двухфакторный дисперсионный анализ с повторениями». Формулы, алгоритмы расчетов в MS Excel детально описаны в книге Р. Вадзинского [11].

Рассмотрим пример расчетов в MS Excel параметров для дисперсионного анализа результатов эксперимента.

Пример 6. В эксперименте для оценки влияния на степень биоразложения в грунте материалов на основе пластифицированного ацетата целлюлозы степени ацетилирования целлюлозы Z_1 и содержания в них древесной муки Z_2 были получены из них методом горячего прессования лабораторные диски.

Каждый из полученных дисков был распилен на четыре образца и для каждого из этих образцов была определена степень их биоразложения по величине потери массы при выдержке в активированном грунте в течение 120 сут (Y, U).

Для расчетов в MS Excel параметров двухфакторного дисперсионного анализа с четырехкратным повторением результатов измерений степени биоразложения в грунте полученных образцов была сформирована следующая таблица исходных данных (рис. 28).



A	B	C	D
Степень ацетилирования целлюлозы Z_1	Степень биоразложения образца в грунте (Y, %) при содержании в нём древесной муки (Z_2, мас. %)		
	0	10	30
2,41	5,5	4,6	5
	5,7	4,6	4,7
	5,5	4,6	5
	5,6	4,7	5,6
2,07	9,7	7,4	6,6
	9,8	7,7	7,3
	9,8	7,9	7,2
	9,8	8	6,9

Рис. 28. Таблица результатов эксперимента для расчета в MS Excel параметров двухфакторного дисперсионного анализа

После выбора инструмента анализа данных «Двухфакторный дисперсионный анализ с повторениями» (рис. 29) и нажатия кнопки «ОК» появляется окно (рис. 30), в котором необходимо ввести входной интервал ячеек с исходными данными для дисперсионного анализа, включая пустую ячейку \$A\$2 (\$A\$2:\$D\$10), число строк для выборки (число повторений измерений $n = 4$), Альфа (уровень значимости $p = 1–P$) и адрес вывода результатов расчета параметров для дисперсионного анализа результатов эксперимента.

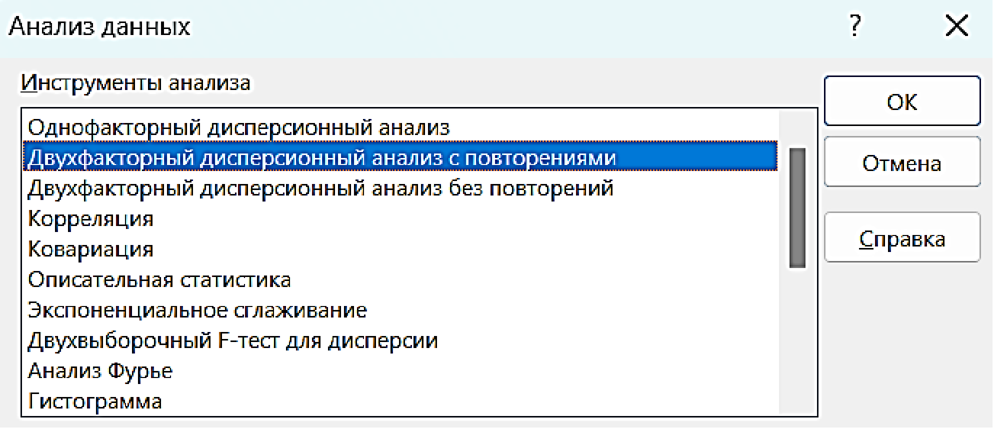
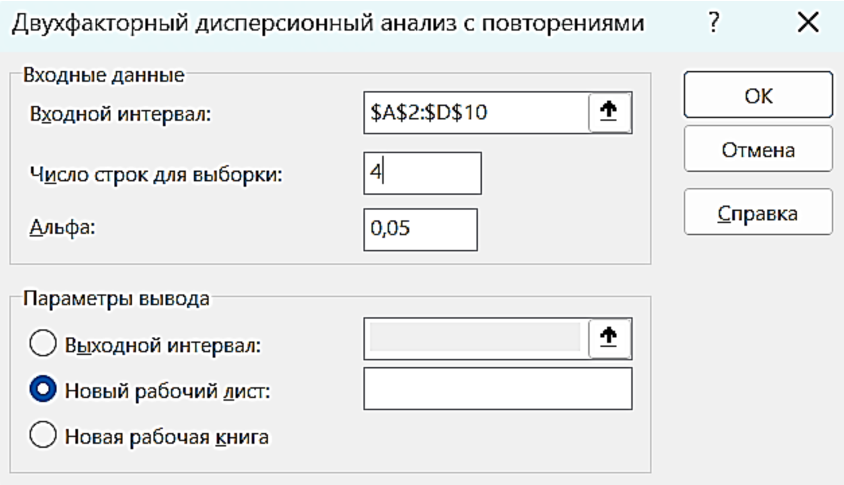


Рис. 29. Окно выбора инструмента анализа данных



Входные данные

Входной интервал: $A$2:$D$10

Число строк для выборки: 4

Альфа: 0,05

Параметры вывода

○ Выходной интервал:

● Новый рабочий лист:

○ Новая рабочая книга

Рис. 30. Окно на экране компьютера с введением данных, необходимых при расчете параметров для дисперсионного анализа результатов эксперимента

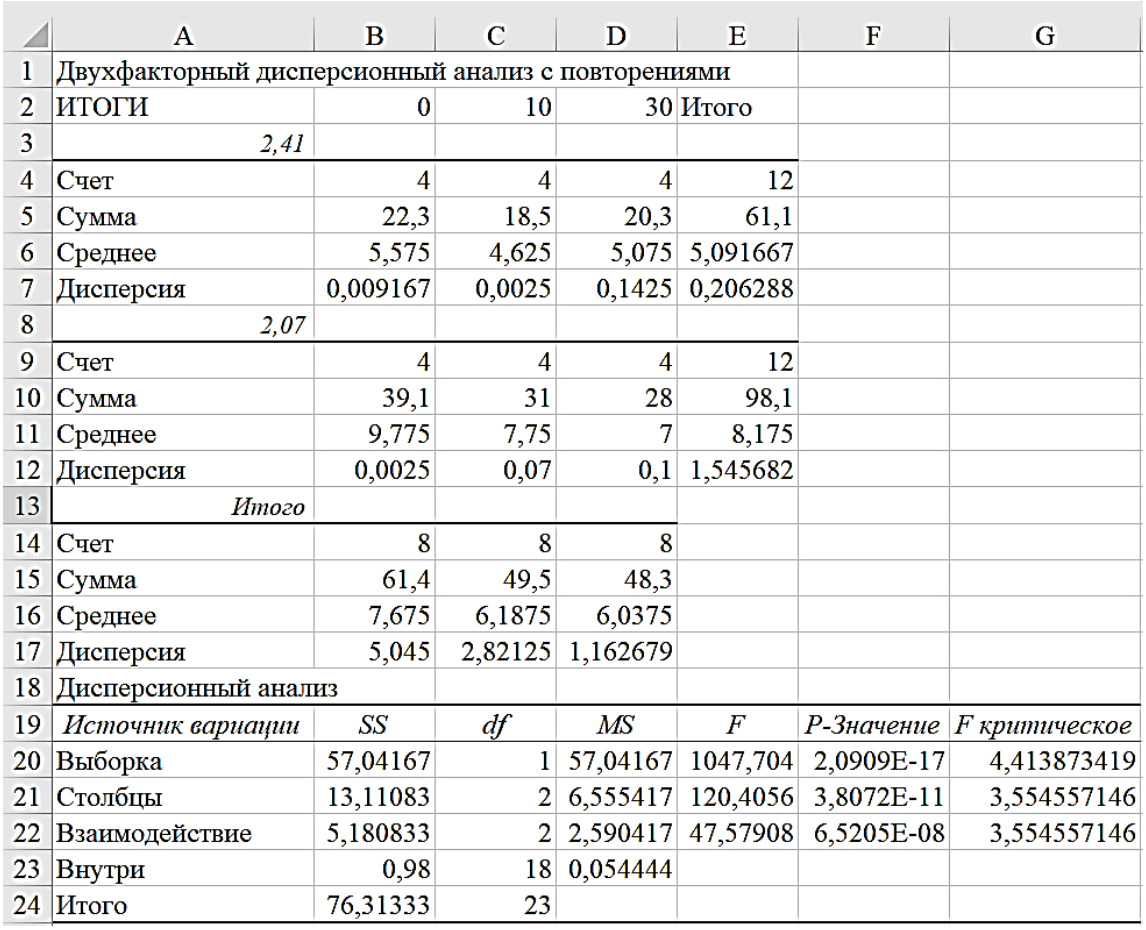
После нажатия кнопки «ОК» (рис. 31) появляется таблица с результатами расчета параметров двухфакторного дисперсионного анализа [11] с четырехкратным повторением результатов измерений степени биоразложения в грунте полученных образцов (см. рис. 31):

- число повторений измерений (Счет);
- сумма результатов измерений (Сумма);
- среднеарифметическое значение результатов измерений (Среднее);
- выборочная дисперсия единичных значений результатов измерений (Дисперсия);
- сумма квадратов отклонений (*SS*);
- число степеней свободы (*df*):
- выборочная внутригрупповая дисперсия (*MS*);
- значение критерия *F*;

- уровень значимости $p = 1–P$ (*P-Значение*);
- критическое значение распределения Фишера при уровне значимости p (*F критическое*).

На рис. 31 приведены полученные результаты расчетов статистических параметров для каждого уровня факторов Z_1 и Z_2, а также параметров для двухфакторного дисперсионного анализа. Параметры (строки 19–24) характеризуют:

- влияние фактора Z_1 на свойство образца Y (строка 20 – Выборка);
- влияние фактора Z_2 на свойство образца Y (строка 21 – Столбцы);
- влияние парного взаимодействия факторов (одновременного изменения) обоих факторов (Z_1Z_2) на свойство образца Y (строка 22 – Взаимодействие);
- выборочное стандартное отклонение результатов измерений для всех опытов эксперимента (Внутри).



	A	B	C	D	E	F	G
1	Двухфакторный дисперсионный анализ с повторениями						
2	ИТОГИ	0	10	30	Итого		
3	*2,41*						
4	Счет	4	4	4	12		
5	Сумма	22,3	18,5	20,3	61,1		
6	Среднее	5,575	4,625	5,075	5,091667		
7	Дисперсия	0,009167	0,0025	0,1425	0,206288		
8	*2,07*						
9	Счет	4	4	4	12		
10	Сумма	39,1	31	28	98,1		
11	Среднее	9,775	7,75	7	8,175		
12	Дисперсия	0,0025	0,07	0,1	1,545682		
13	*Итого*						
14	Счет	8	8	8			
15	Сумма	61,4	49,5	48,3			
16	Среднее	7,675	6,1875	6,0375			
17	Дисперсия	5,045	2,82125	1,162679			
18	Дисперсионный анализ						
19	*Источник вариации*	*SS*	*df*	*MS*	*F*	*P-Значение*	*F критическое*
20	Выборка	57,04167	1	57,04167	1047,704	2,0909E-17	4,413873419
21	Столбцы	13,11083	2	6,555417	120,4056	3,8072E-11	3,554557146
22	Взаимодействие	5,180833	2	2,590417	47,57908	6,5205E-08	3,554557146
23	Внутри	0,98	18	0,054444			
24	Итого	76,31333	23				

Рис. 31. Результаты расчета в MS Excel параметров двухфакторного дисперсионного анализа с четырехкратным повторением результатов измерений степени биоразложения в грунте полученных образцов

Анализ данных таблицы (см. рис. 31) свидетельствует о наибольшем и очень сильном влиянии на свойство образца *Y* фактора Z_1 (уровень значимости $p = 2{,}0909 \cdot 10^{-17}$ намного меньше 0,05) и расчетное значение критерия $F = 1047{,}704$ намного больше критического значения распределения Фишера (*F – Критическое* = 4,414).

Также с вероятностью более 0,95 по данным дисперсионного анализа на свойство образца *Y* влияние оказывают изменение фактора Z_1 и одновременное изменение обоих факторов в исследованной факторной области.

В MS Excel нет программы расчета параметров для многофакторного дисперсионного анализа.

В Statistica можно рассчитать параметры для одно-, двух- и многофакторного дисперсионного анализа. Для этого в меню «Анализ» открыть раздел «Дисперсионный анализ» (рис. 32).

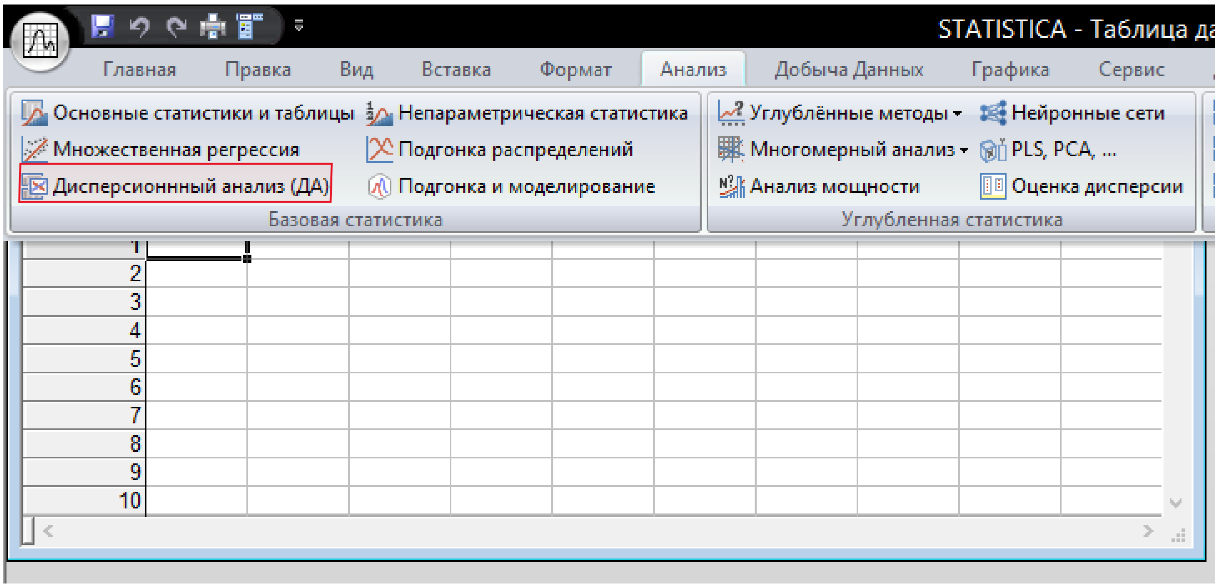


Рис. 32. Экран компьютера после открытия меню «Анализ»

Процедура расчетов в Statistica параметров для многофакторного дисперсионного анализа непростая [10], и она требует большого числа опытов в эксперименте.

Некоторые параметры для многофакторного дисперсионного анализа в MS Excel и Statistica приводятся при расчете параметров для линейного многофакторного регрессионного анализа результатов эксперимента.

4.4. Расчет в программах Microsoft Excel и Statistica значений параметров для регрессионного анализа результатов эксперимента

4.4.1. Расчет в программах Microsoft Excel значений параметров для регрессионного анализа результатов эксперимента

Расчет статистических параметров для линейного регрессионного анализа результатов эксперимента в MS Excel можно провести с помощью инструмента анализа «Регрессия» в меню «Данные».

Для выполнения линейного регрессионного анализа результатов эксперимента выполняют следующие основные этапы [11]:

- выбирают модель уравнения регрессии для реализованного плана эксперимента;
- по полученным результатам эксперимента рассчитывают параметры для выбранной модели уравнения регрессии;
- проверяют для заданной доверительной вероятности статистические гипотезы о значимости рассчитанных параметров выбранной модели уравнения регрессии;
- в случае неудовлетворительной оценки значимости отдельных рассчитанных параметров выбранной модели уравнения регрессии проводят ее корректировку путем исключения из нее статистически незначимых параметров или включения новых параметров.

Самым простым видом моделей уравнения регрессии являются модели парной линейной регрессии с одним или двумя параметрами, представляющие собой следующие линейные функции с кодированными x или натуральными Z значениями независимого входного фактора:

- A $(x; b_1) = b_1x$;
- B $(x; b) = b_0 + b_1x$;
- C $(Z; B_1) = B_1Z$;
- D $(Z; B) = B_0 + B_1Z$.

Включение в регрессионную модель свободного члена B_0 **является обязательным** независимо от оценки его статистических параметров, если в эксперименте проводились опыты при нулевом натуральном значении входного фактора Z.

Для линейного регрессионного анализа результатов однофакторных экспериментов можно использовать не только линейные функции, но и другие функции, линейные относительно их араметров, например, многочлен m-й степени для натурального значения входного фактора:

$$U(Z;\, B) \ = B_0 + B_1 Z + B_2 Z^2 + \ \dots \ + B_m \ Z^m,$$

где Z – натуральное значение одного входного фактора;

B_0, B_1, B_2, …, B_m – параметры регрессионной линейной модели с натуральными значениями одного входного фактора.

Для линейного регрессионного анализа результатов многофакторных Z_i экспериментов можно использовать линейные относительно их параметров многочлены m-й степени для натуральных значений входных факторов, например для двухфакторного эксперимента:

$$U(Z_1;Z_2;B) \ = B_0 + B_1 Z_1 + B_2 Z_2 + B_3 Z_1 Z_2 + B_4 Z_1^2 + B_5 Z_2^2 + B_6 Z_1^3 + B_7 Z_2^3 + \dots,$$

где Z_1 – натуральные значения первого фактора;

Z_2 – натуральные значения второго фактора;

B_0, B_1, B_2, B_3, B_4, B_5, B_6, B_7, … – параметры регрессионной линейной модели.

Если в многофакторном эксперименте проводились опыты **при нулевом натуральном значении хотя бы одного входного фактора,** то в регрессионном анализе результатов эксперимента **необходимо обязательно оставлять в регрессионной модели свободный член *B*** $\boldsymbol{B_0}$ независимо от оценки его статистических параметров.

Представленные в примере 6 результаты двухфакторного эксперимента с четырехкратным повторением каждого опыта были дополнены экспериментальными данными по измерению степени биоразложения в грунте образцов на основе пластифицированного ацетата целлюлозы со степенями ацетилирования целлюлозы 2,19 и 2,29 (рис. 33) с целью применения к ним двухфакторного линейного регрессионного анализа. Для этого выбираем следующую модель уравнения регрессии:

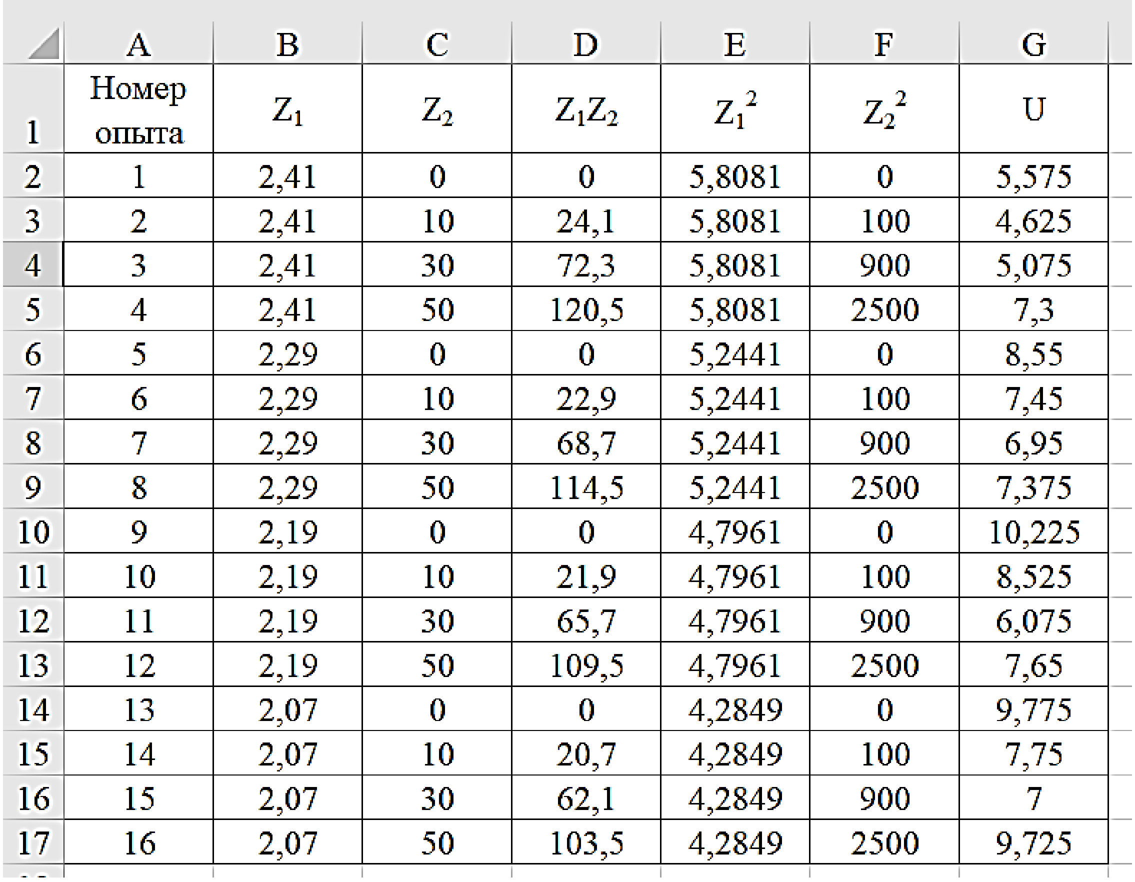
$$U(Z_1;Z_2;B) \ = B_0 + B_1 Z_1 + B_2 Z_2 + B_3 Z_1 Z_2 + B_4 Z_1^2 + B_5 Z_2^2,$$

где U – степень биоразложения в грунте образцов, мас. %;

Z_1 – степень ацетилирования целлюлозы;

Z_2 – содержание древесной муки в образце, мас. %;

B_0, B_1, B_2, B_3, B_4, B_5 – параметры регрессионной линейной модели.



	A	B	C	D	E	F	G
1	Номер опыта	Z_1	Z_2	Z_1Z_2	Z_1^2	Z_2^2	U
2	1	2,41	0	0	5,8081	0	5,575
3	2	2,41	10	24,1	5,8081	100	4,625
4	3	2,41	30	72,3	5,8081	900	5,075
5	4	2,41	50	120,5	5,8081	2500	7,3
6	5	2,29	0	0	5,2441	0	8,55
7	6	2,29	10	22,9	5,2441	100	7,45
8	7	2,29	30	68,7	5,2441	900	6,95
9	8	2,29	50	114,5	5,2441	2500	7,375
10	9	2,19	0	0	4,7961	0	10,225
11	10	2,19	10	21,9	4,7961	100	8,525
12	11	2,19	30	65,7	4,7961	900	6,075
13	12	2,19	50	109,5	4,7961	2500	7,65
14	13	2,07	0	0	4,2849	0	9,775
15	14	2,07	10	20,7	4,2849	100	7,75
16	15	2,07	30	62,1	4,2849	900	7
17	16	2,07	50	103,5	4,2849	2500	9,725

Рис. 33. Таблица исходных данных для расчета параметров регрессионного анализа в MS Excel

Параметры регрессионной линейной модели (число параметров *par* = 6) оценивают влияние на степень биоразложения в грунте образцов:

B_0 – постоянная уравнения регрессии (свободный член);

B_1 – влияние на степень биоразложения в грунте образцов изменения степени ацетилирования целлюлозы в составе пластифицированного ацетата целлюлозы;

B_2 – влияние на степень биоразложения в грунте образцов изменения содержания древесной муки в образце;

B_3 – влияние на степень биоразложения в грунте образцов одновременного изменения степени ацетилирования целлюлозы в составе пластифицированного ацетата целлюлозы и содержания древесной муки в образце;

B_4 – квадратичный эффект влияния на степень биоразложения в грунте образцов изменения степени ацетилирования целлюлозы в составе пластифицированного ацетата целлюлозы;

B_5 – квадратичный эффект влияния на степень биоразложения в грунте образцов изменения содержания древесной муки в образце.

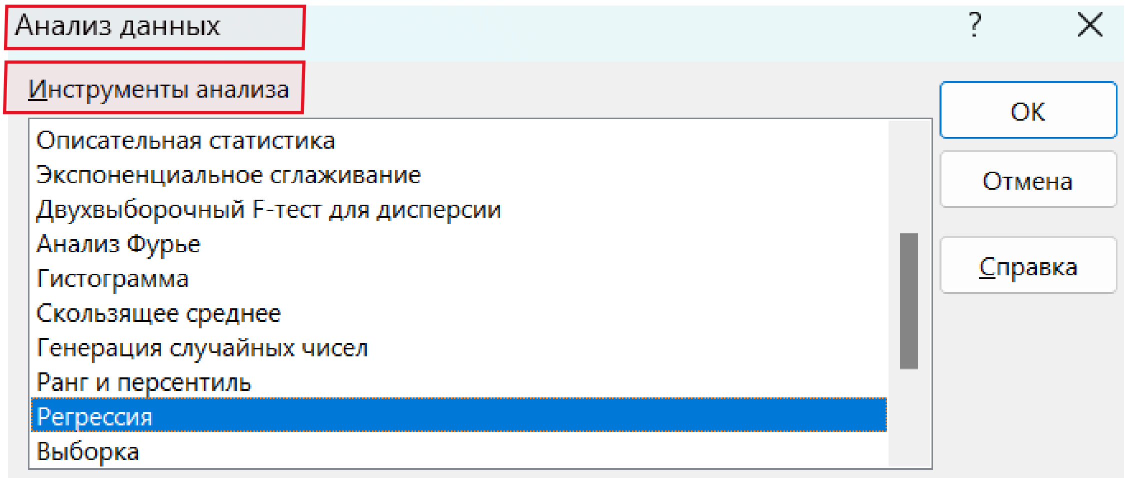
Включение в модель уравнения регрессии свободного члена B_0 является обязательным не только из-за наличия опытов с нулевым значением фактора Z_2, но и потому, что по имеющимся результатам научных исследований при нулевых значениях входных факторов получаемый материал (пластифицированная целлюлоза) имеет положительное значение степени биоразложения в грунте.

Результаты эксперимента по определению закономерности влияния на степень биоразложения в грунте образцов U на основе пластифицированного ацетата целлюлозы степени ацетилирования целлюлозы (Z_1, $Z1$) и содержания в них древесной муки (Z_2, $Z2$) формируем в MS Excel в таблицу (см. рис. 33). В эту таблицу исходных данных были введены также дополнительные переменные $Z1Z2$, $Z1^2$, $Z2^2$ для возможности получения нелинейной модели уравнения регрессии.

Открыв в MS Excel в меню «Данные» раздел «Анализ данных» выбираем инструмент анализа «Регрессия» (рис. 34) и нажимаем кнопку «ОК». В открывшемся окне (рис. 35) в поле ввода «Входной интервал Y» вводим диапазон ячеек со значениями U в таблице исходных данных (см. рис. 33), в поле ввода «Входной интервал X» вводим диапазон ячеек из этой таблицы со значениями всех переменных (Z_1, Z_2, Z_1Z_2, Z_1^2, Z_2^2).

После нажатия кнопки «ОК» в этом окне (см. рис. 34) открывается окно для ввода элементов управления расчетами необходимых статистических параметров для регрессионного анализа результатов эксперимента (см. рис. 35).

В блоке «Входные данные» (см. рис. 35) в квадратике «Метки» нажимаем левую кнопку мыши для учета в расчетах в MS Excel, что в ячейках строки 1 находятся текстовые данные. Если при расчетах нужно применять доверительную вероятность $P \neq 0{,}95$, то нужно в квадратике «Уровень надежности» нажать левую кнопку мыши и заменить в поле ввода число 0,95 на значение требуемой доверительной вероятности. Отсутствие метки в квадратике «Константа – ноль» означает выбор для расчетов параметров уравнения регрессии со свободным членом B_0. В блоке «Параметры вывода» отмечаем место вывода рассчитанных по умолчанию параметров регрессионного анализа. Для вывода дополнительных статистических параметров регрессионного анализа делаем соответствующие отметки в квадратиках блоков «Остатки» и «Нормальная вероятность».

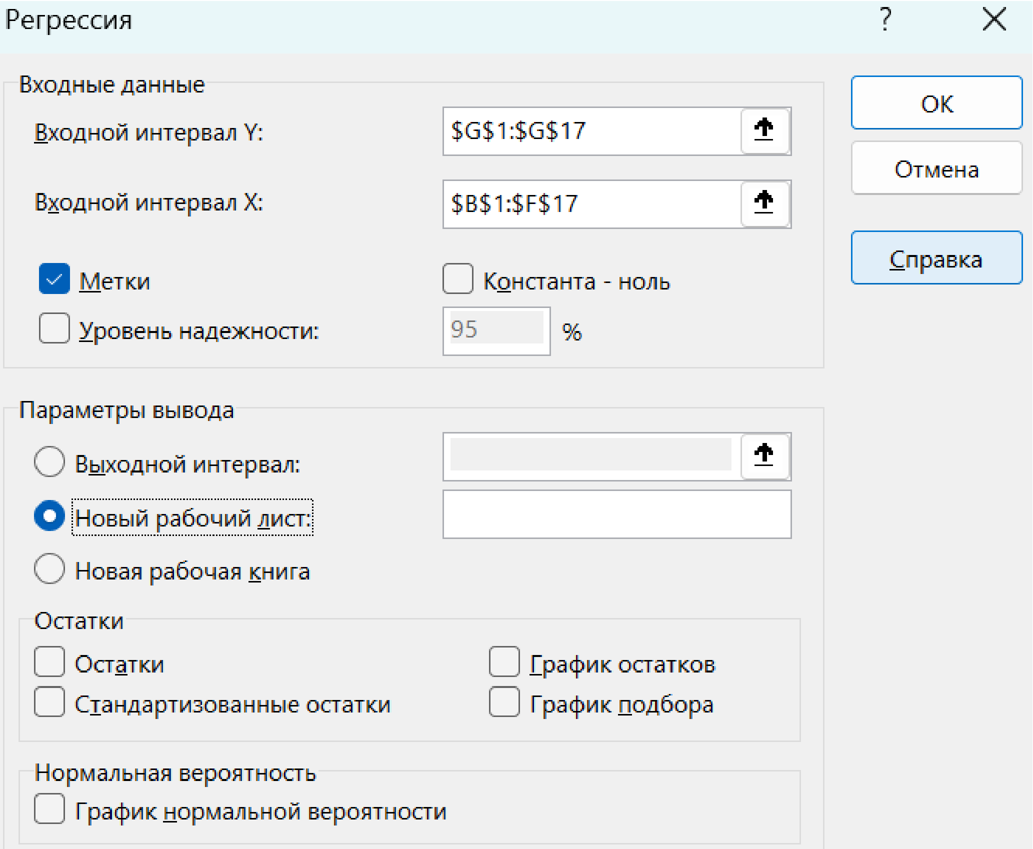


Анализ данных

Инструменты анализа

Описательная статистика
Экспоненциальное сглаживание
Двухвыборочный F-тест для дисперсии
Анализ Фурье
Гистограмма
Скользящее среднее
Генерация случайных чисел
Ранг и персентиль
Регрессия
Выборка

Рис. 34. Окно на экране компьютера для выбора инструмента анализа результатов эксперимента в MS Excel



Регрессия

Входные данные

Входной интервал Y: $G$1:$G$17

Входной интервал X: $B$1:$F$17

☑ Метки ☐ Константа - ноль

☐ Уровень надежности: 95 %

Параметры вывода

○ Выходной интервал:

● Новый рабочий лист:

○ Новая рабочая книга

Остатки

☐ Остатки ☐ График остатков

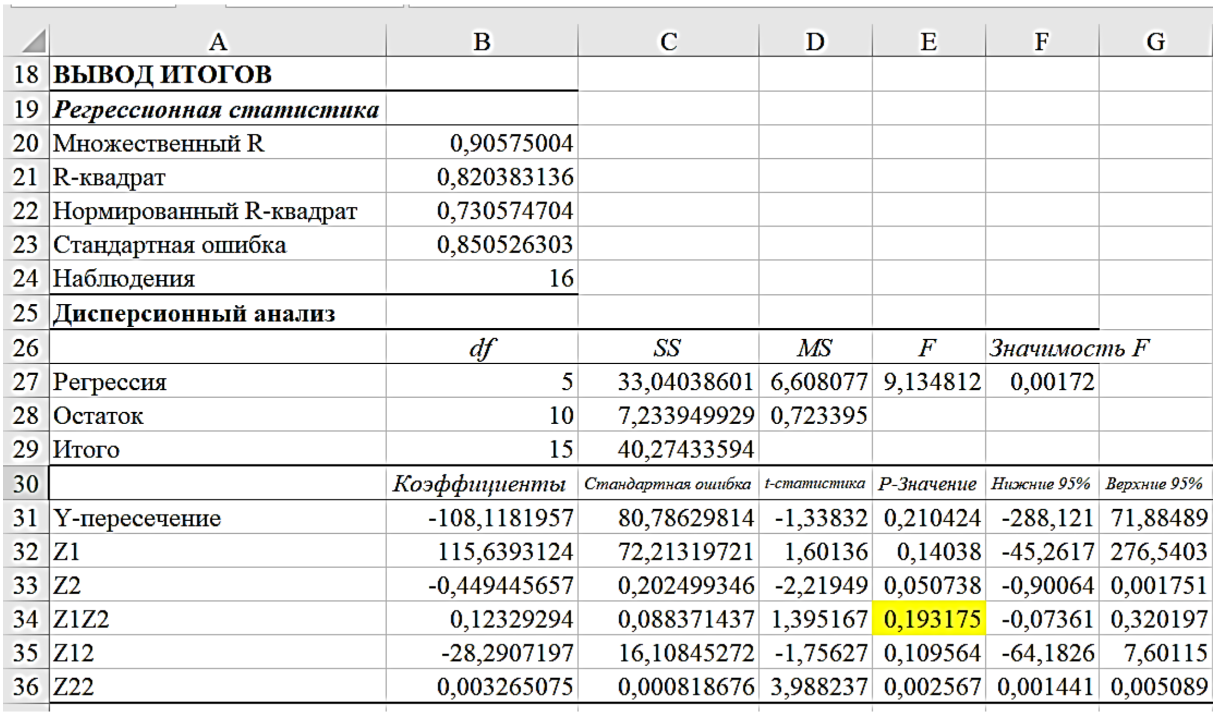
☐ Стандартизованные остатки ☐ График подбора

Нормальная вероятность

☐ График нормальной вероятности

Рис. 35. Окно на экране компьютера для ввода элементов управления расчетами необходимых статистических параметров для регрессионного анализа результатов эксперимента в MS Excel

После выбора элементов управления расчетами необходимых статистических параметров для регрессионного анализа результатов эксперимента и нажатия кнопки «ОК» (см. рис. 35) открывается окно «Вывод итогов» (рис. 36) с результатами выполненных по элементам управления расчетов статистических параметров для регрессионного анализа результатов эксперимента в MS Excel.



	A	B	C	D	E	F	G
18	**ВЫВОД ИТОГОВ**						
19	***Регрессионная статистика***						
20	Множественный R	0,90575004					
21	R-квадрат	0,820383136					
22	Нормированный R-квадрат	0,730574704					
23	Стандартная ошибка	0,850526303					
24	Наблюдения	16					
25	**Дисперсионный анализ**						
26		*df*	*SS*	*MS*	*F*	*Значимость F*	
27	Регрессия	5	33,04038601	6,608077	9,134812	0,00172	
28	Остаток	10	7,233949929	0,723395			
29	Итого	15	40,27433594				
30		*Коэффициенты*	*Стандартная ошибка*	*t-статистика*	*P-Значение*	*Нижние 95%*	*Верхние 95%*
31	Y-пересечение	-108,1181957	80,78629814	-1,33832	0,210424	-288,121	71,88489
32	Z1	115,6393124	72,21319721	1,60136	0,14038	-45,2617	276,5403
33	Z2	-0,449445657	0,202499346	-2,21949	0,050738	-0,90064	0,001751
34	Z1Z2	0,12329294	0,088371437	1,395167	0,193175	-0,07361	0,320197
35	Z12	-28,2907197	16,10845272	-1,75627	0,109564	-64,1826	7,60115
36	Z22	0,003265075	0,000818676	3,988237	0,002567	0,001441	0,005089

Рис. 36. Окно на экране компьютера с итогами выполненных по элементам управления расчетов параметров для регрессионного анализа результатов эксперимента в MS Excel

В окне «Вывод итогов» (см. рис. 36) рассчитанные значения статистических параметров распределены на три блока: «Регрессионная статистика», «Дисперсионный анализ» и блок без названия.

В блоке «Регрессионная статистика» приведены значения следующих статистических параметров регрессионного анализа, алгоритмы расчетов которых в MS Excel описаны в книге Р. Вадзинского [11]:

- множественный *R*;
- *R*-квадрат;
- нормированный *R*-квадрат;
- стандартная ошибка;
- наблюдения.

Статистический параметр «Множественный *R*» (множественный коэффициент линейной корреляции *R* [11]) представляет собой выборочный коэффициент линейной корреляции между фактическими и рассчитанными по полученному уравнению регрессии значениями зависимой переменной U (свойства образца). Значения множественного коэффициента линейной корреляции R находятся в диапазоне $0 \leq R \leq 1$.

Коэффициент детерминации «R-квадрат» (R^2) характеризует долю общего разброса экспериментальных значений свойства образца U от его среднеарифметического значения, рассчитанного по полученному уравнению регрессии. Коэффициент детерминации «R-квадрат» показывает, насколько значимо влияние на свойства образца U параметров уравнения регрессии (B_0, B_1, B_2, B_3, B_4, B_5).

Чем лучше выбранное уравнение регрессии описывает изменение U при изменении значений входных факторов (независимых переменных Z_1 и Z_2) тем ближе значение R^2 к единице.

Статистический параметр «Нормированный R-квадрат» (скорректированный коэффициент детерминации R^2) рассчитывается при упрощении модели регрессии исключением из нее части входных факторов, включенных в первоначальную модель. Эта задача решается подбором подмножества входных факторов, максимизирующим среднее арифметическое значение рассчитываемых значений R^2 [11].

Статистический параметр «Стандартная ошибка» (выборочное стандартное отклонение остатков [11]) характеризует разброс полученных экспериментальных значений U относительно линии полученного уравнения регрессии.

Статистический параметр «Наблюдения» фиксирует число значений зависимой переменной U, для которого проводился расчет параметров для регрессионного анализа (ячейка А24 на рис. 36).

В блоке «Дисперсионный анализ» (см. рис. 36) приведены для строк «Регрессия», «Остаток», «Итого» значения следующих статистических параметров, используемых для регрессионного анализа [11]:

- *df* – число степеней свободы;
- *SS* – сумма квадратов остатков;
- *MS* – средний квадрат;
- *F* – расчетное значение статистики *F*;
- *значимость F* – критическое значение распределения Фишера.

Для строки «Регрессия»:

- *df* = *par* –1;
- *SS* – сумма квадратов остатков, обусловленная регрессией;
- *MS* – средний квадрат регрессии.

Для строки «Остаток»:

- *df* = *par* –1;
- *SS* – сумма квадратов остатков отклонений фактических значений *U* от их расчетных значений по уравнению регрессии;
- *MS* – средний квадрат остатков.

В строке «Итого» приводится число опытов, уменьшенное на единицу.

В блоке без названия для всех членов уравнения регрессии (кроме B_0, названного *Y*-пересечение) приводятся в ячейках 18–22 столбца А названия независимых переменных из строки 1 столбцов В–G таблицы (см. рис. 36). В столбце В – рассчитанные коэффициенты их влияния на свойство *U* (B_0, B_1, B_2, B_3, B_4, B_5 соответственно по строкам 17–20) и в столбцах С–I – статистические параметры для их оценки:

- «*Стандартная ошибка*» – стандартная ошибка рассчитанного параметра уравнения регрессии;
- «*t-статистика*» – рассчитанное значение статистики *T*;
- «*Р-значение*» – значимость *p* параметров (коэффициентов) уравнения регрессии ($p = 1 – P$);
- «*Нижние 95,0 %*» (столбец F, H) – крайний левый член доверительного интервала значений параметров уравнения регрессии для вероятности 0,95;
- «*Верхние 95,0 %*» (столбец G, I) – крайний правый член доверительного интервала значений параметров уравнения регрессии для вероятности 0,95.

Проведем статистический анализ рассчитанных значений параметров для выбранной модели уравнения регрессии.

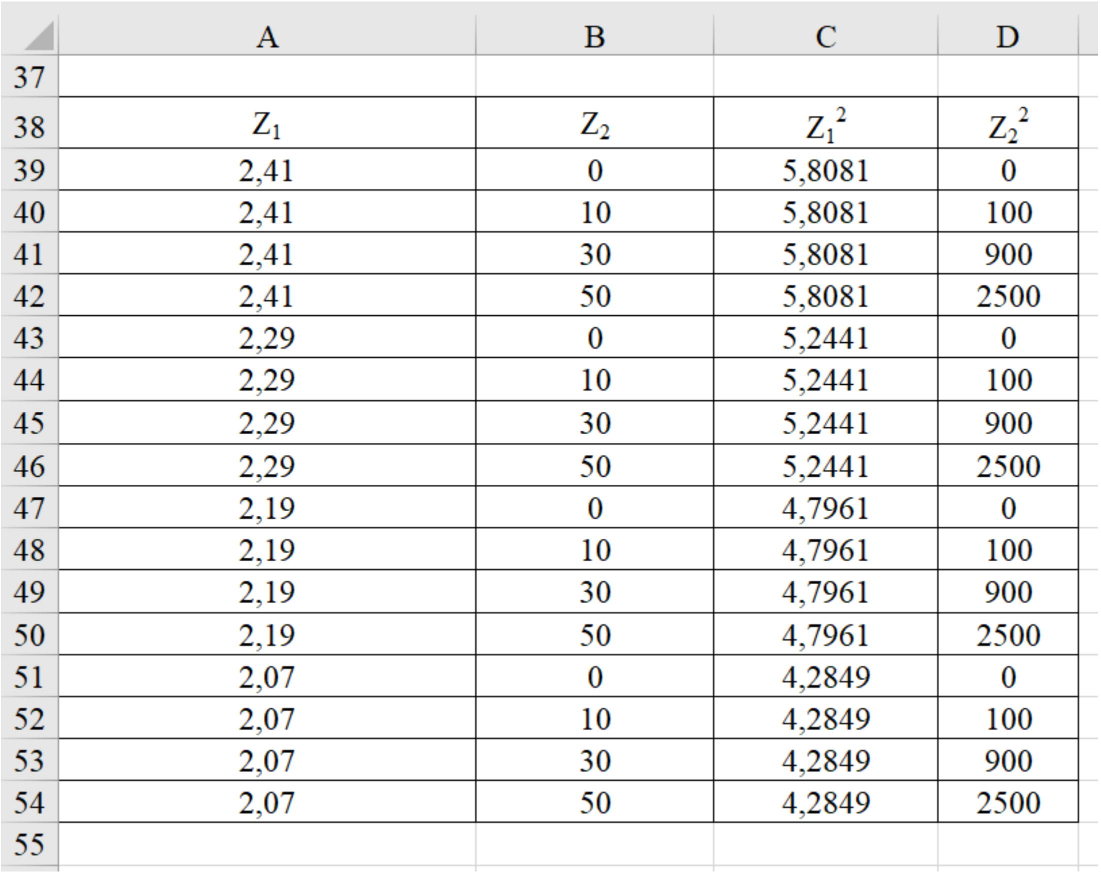
Величина статистического параметра «*Значимость F*» в блоке «Дисперсионный анализ» (см. рис. 36) меньше уровня значимости $p = 0{,}05$, что свидетельствует о достоверности влияния независимых переменных на свойство U с вероятностью $P \geq 0{,}95$ по полученному уравнению регрессии. Но статистические параметры блока «Регрессионная статистика» не очень близки к 1 («Множественный R», «R-квадрат», «Нормированный R-квадрат»).

В блоке без названия (см. рис. 36) значения всех коэффициентов (параметров) выбранной регрессионной модели не выпадают из доверительного интервала для вероятности 0,95 (ячейки F32:I36 в таблице на рис. 36). Все параметры, за исключением B_5, имеют значимость (уровень достоверности их влияния на свойство образца U) больше 0,05 (статистический параметр «*Р-значение*»), т. е. доверительную вероятность $P \leq 0{,}95$ их влияния на свойство образца U. Наименьшую вероятность этого влияния ($P \leq 0{,}81$) оказывает параметр B_3 (ячейка E34, выделенная в таблице на рис. 36). Это свидетельствует о том, что с доверительной вероятностью 0,95 параметр B_3 несущественно (незначимо) отличается от нуля [11].

Для проверки возможности получения уравнения регрессии с параметрами описывающих влияние независимых переменных Z_1 и Z_2 на свойство образца U с доверительной вероятностью не менее 0,95 исключим параметр B_3 из регрессионной модели, которая после этого будет иметь следующий вид:

$$U(Z_1;\, Z_2;\, B) \;= B_0 + B_1 Z_1 + B_2 Z_2 + B_4 Z_1^2 + B_5 Z_2^2 .$$

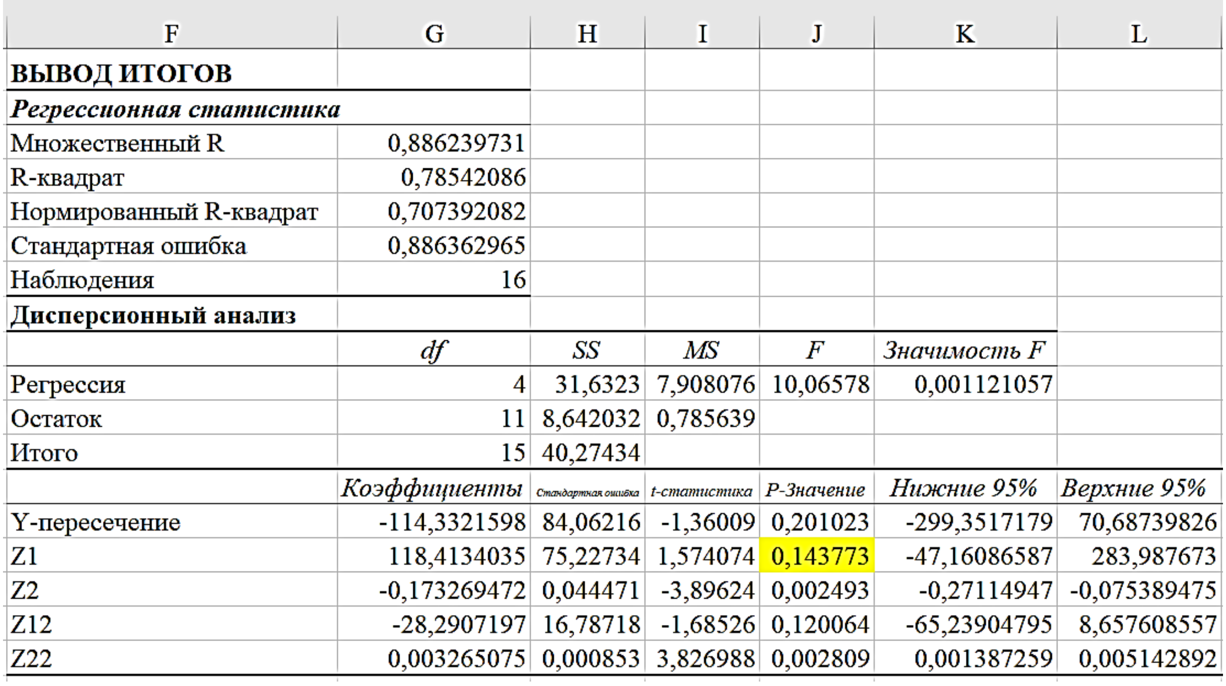
Для проведения расчетов параметров данного уравнения регрессии используем таблицу исходных данных без столбца Z_1Z_2 (рис. 37). В окне для ввода элементов управления расчетами необходимых статистических параметров для регрессионного анализа результатов эксперимента в MS Excel (см. рис. 35) изменим в окне «Входной интервал X» на ячейки A38–D54, выберем новые параметры вывода и нажмем кнопку «ОК».



	A	B	C	D
37				
38	Z_1	Z_2	Z_1^2	Z_2^2
39	2,41	0	5,8081	0
40	2,41	10	5,8081	100
41	2,41	30	5,8081	900
42	2,41	50	5,8081	2500
43	2,29	0	5,2441	0
44	2,29	10	5,2441	100
45	2,29	30	5,2441	900
46	2,29	50	5,2441	2500
47	2,19	0	4,7961	0
48	2,19	10	4,7961	100
49	2,19	30	4,7961	900
50	2,19	50	4,7961	2500
51	2,07	0	4,2849	0
52	2,07	10	4,2849	100
53	2,07	30	4,2849	900
54	2,07	50	4,2849	2500
55				

Рис. 37. Таблица исходных данных без столбца Z_1Z_2

Полученные результаты расчетов в MS Excel параметров уравнения регрессии без B_3 приведены в таблице на рис. 38.



F	G	H	I	J	K	L
ВЫВОД ИТОГОВ						
Регрессионная статистика						
Множественный R	0,886239731					
R-квадрат	0,78542086					
Нормированный R-квадрат	0,707392082					
Стандартная ошибка	0,886362965					
Наблюдения	16					
Дисперсионный анализ						
	df	*SS*	*MS*	*F*	*Значимость F*	
Регрессия	4	31,6323	7,908076	10,06578	0,001121057	
Остаток	11	8,642032	0,785639			
Итого	15	40,27434				
	Коэффициенты	*Стандартная ошибка*	*t-статистика*	*P-Значение*	*Нижние 95%*	*Верхние 95%*
Y-пересечение	-114,3321598	84,06216	-1,36009	0,201023	-299,3517179	70,68739826
Z1	118,4134035	75,22734	1,574074	0,143773	-47,16086587	283,987673
Z2	-0,173269472	0,044471	-3,89624	0,002493	-0,27114947	-0,075389475
Z12	-28,2907197	16,78718	-1,68526	0,120064	-65,23904795	8,657608557
Z22	0,003265075	0,000853	3,826988	0,002809	0,001387259	0,005142892

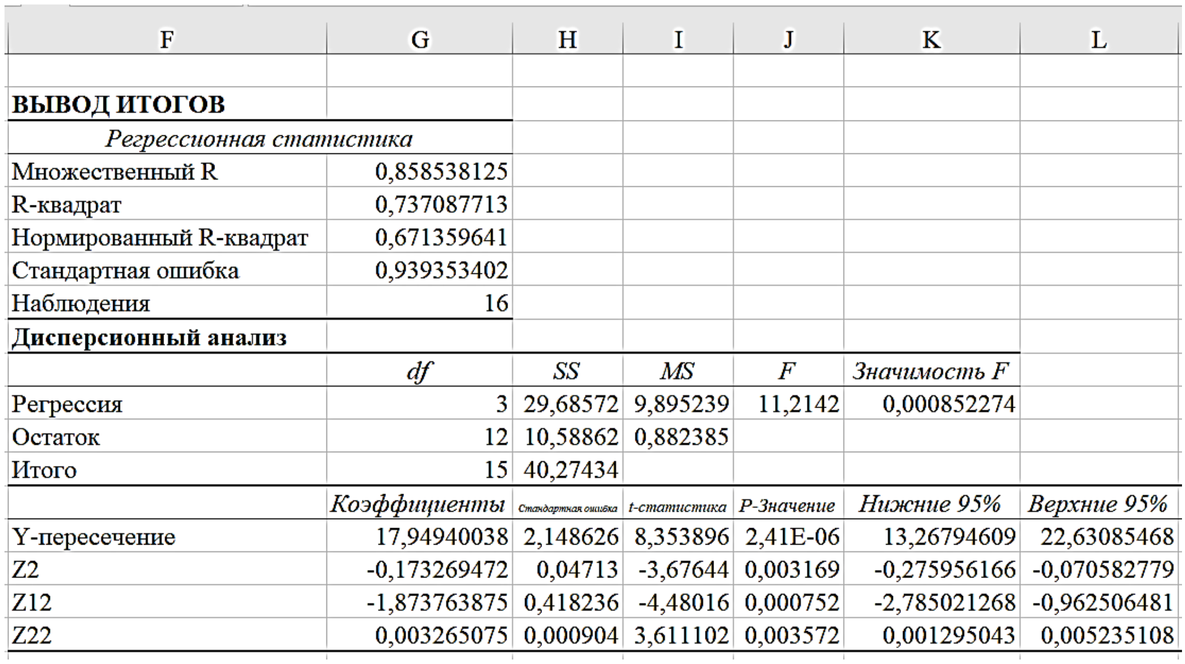
Рис. 38. Окно на экране компьютера с итогами выполненных
по элементам управления расчетов в MS Excel параметров
для регрессионного анализа регрессионной модели без параметра B_3

После исключения из регрессионной модели члена с параметром B_3 в блоке без названия два параметра (B_1 и B_4) имеют «*P-Значение*» на рис. 38 больше 0,05. Из них наибольшее значение имеет параметр B_1, поэтому исключим его из регрессионной модели и, как при исключении параметра B_3, выполним в MS Excel аналогичные процедуры. Результаты расчетов в MS Excel параметров для регрессионного анализа после исключения из регрессионной модели членов с B_1 и B_3 приведены на рис. 39.

Анализируя полученные результаты расчетов параметров для скорректированной регрессионной модели (см. рис. 39), можно считать, что наиболее достоверно с вероятностью не менее 0,95 на свойство образца U влияние факторов Z_1 и Z_2 описывает следующая регрессионная модель:

$$U(Z_1;Z_2;B) = 17{,}94940038 - 0{,}173269472Z_2 - 1{,}873763875Z_1^2 + 0{,}003265075Z_2^2.$$

Основаниями для такого вывода является значение статистического параметра «*Значимость F*» (в блоке «Дисперсионный анализ» на рис. 39) намного меньше 0,05. Все параметры данной регрессионной модели влияют на свойство объекта *U* с вероятностью более 0,95, так как значения статистического параметра «*P-значение*» у них меньше 0,05.



F	G	H	I	J	K	L
ВЫВОД ИТОГОВ						
Регрессионная статистика						
Множественный R	0,858538125					
R-квадрат	0,737087713					
Нормированный R-квадрат	0,671359641					
Стандартная ошибка	0,939353402					
Наблюдения	16					
Дисперсионный анализ						
	df	*SS*	*MS*	*F*	*Значимость F*	
Регрессия	3	29,68572	9,895239	11,2142	0,000852274	
Остаток	12	10,58862	0,882385			
Итого	15	40,27434				
	Коэффициенты	*Стандартная ошибка*	*t-статистика*	*P-Значение*	*Нижние 95%*	*Верхние 95%*
Y-пересечение	17,94940038	2,148626	8,353896	2,41E-06	13,26794609	22,63085468
Z2	-0,173269472	0,04713	-3,67644	0,003169	-0,275956166	-0,070582779
Z12	-1,873763875	0,418236	-4,48016	0,000752	-2,785021268	-0,962506481
Z22	0,003265075	0,000904	3,611102	0,003572	0,001295043	0,005235108

Рис. 39. Окно на экране компьютера с итогами выполненных в MS Excel расчетов параметров для регрессионного анализа (без параметров B_1 и B_3) скорректированной регрессионной модели

Однозначное влияние на свойство *U* по данной регрессионной модели оказывает степень ацетилирования целлюлозы в полимерной фазе образца. Отрицательный знак коэффициента B_4 показывает, что с увеличением степени ацетилирования в полимерной фазе образца его способность к биоразложению в грунте уменьшается.

Влияние на это свойство изменения содержания в образца древесной муки более сложное (рис. 40).

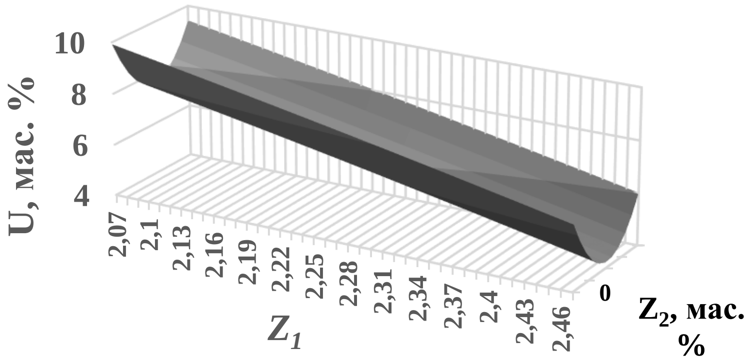
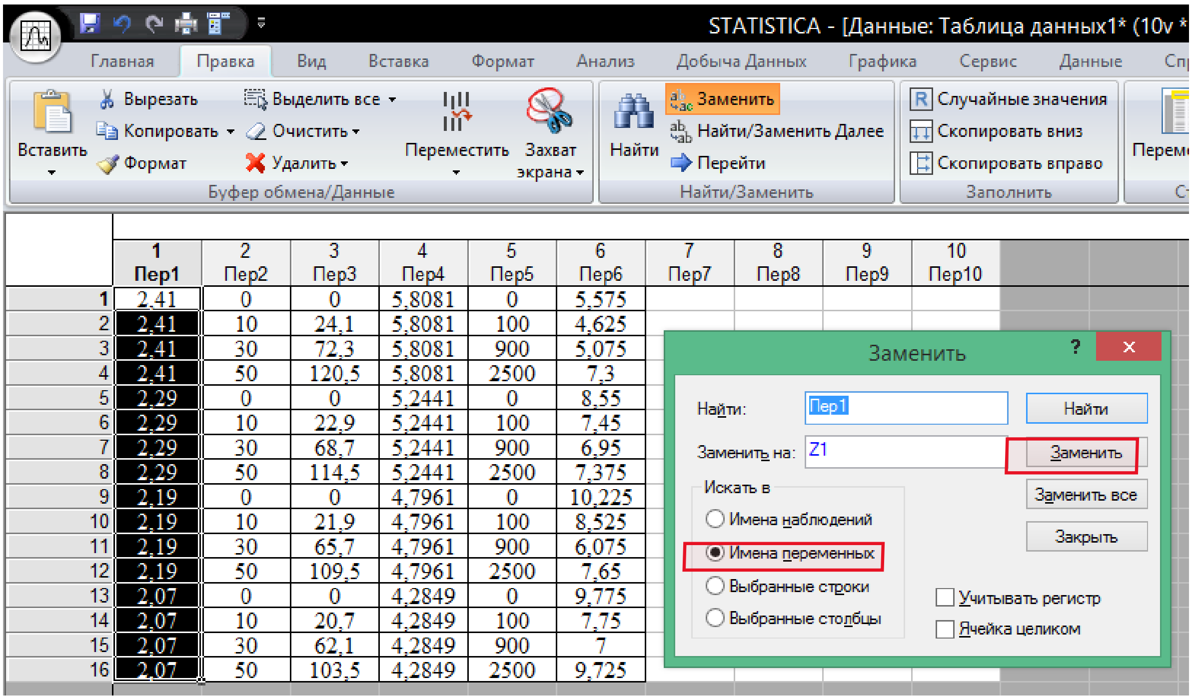


Рис. 40. Влияние на свойство образца *U* независимых входных факторов Z_1 и Z_2

В Statistica применяется такой же математический аппарат расчетов параметров регрессионного анализа, как и в MS Excel, но алгоритм исключения незначимых параметров регрессионной линейной модели автоматизирован и есть возможности расчета дополнительных статистических параметров для анализа результатов экспериментов.

4.4.2. Расчет в программе Statistica значений параметров для регрессионного анализа результатов эксперимента

Для примера и сравнения выполним в Statistica расчеты значений параметров для регрессионного анализа результатов эксперимента (см. рис. 33). Для этого скопируем ячейки B2:G17 в файле MS Excel и вставим в таблицу исходных данных Statistica (рис. 41).



	1 Пер1	2 Пер2	3 Пер3	4 Пер4	5 Пер5	6 Пер6	7 Пер7	8 Пер8	9 Пер9	10 Пер10
1	2,41	0	0	5,8081	0	5,575				
2	2,41	10	24,1	5,8081	100	4,625				
3	2,41	30	72,3	5,8081	900	5,075				
4	2,41	50	120,5	5,8081	2500	7,3				
5	2,29	0	0	5,2441	0	8,55				
6	2,29	10	22,9	5,2441	100	7,45				
7	2,29	30	68,7	5,2441	900	6,95				
8	2,29	50	114,5	5,2441	2500	7,375				
9	2,19	0	0	4,7961	0	10,225				
10	2,19	10	21,9	4,7961	100	8,525				
11	2,19	30	65,7	4,7961	900	6,075				
12	2,19	50	109,5	4,7961	2500	7,65				
13	2,07	0	0	4,2849	0	9,775				
14	2,07	10	20,7	4,2849	100	7,75				
15	2,07	30	62,1	4,2849	900	7				
16	2,07	50	103,5	4,2849	2500	9,725				

Заменить

Найти: Пер1

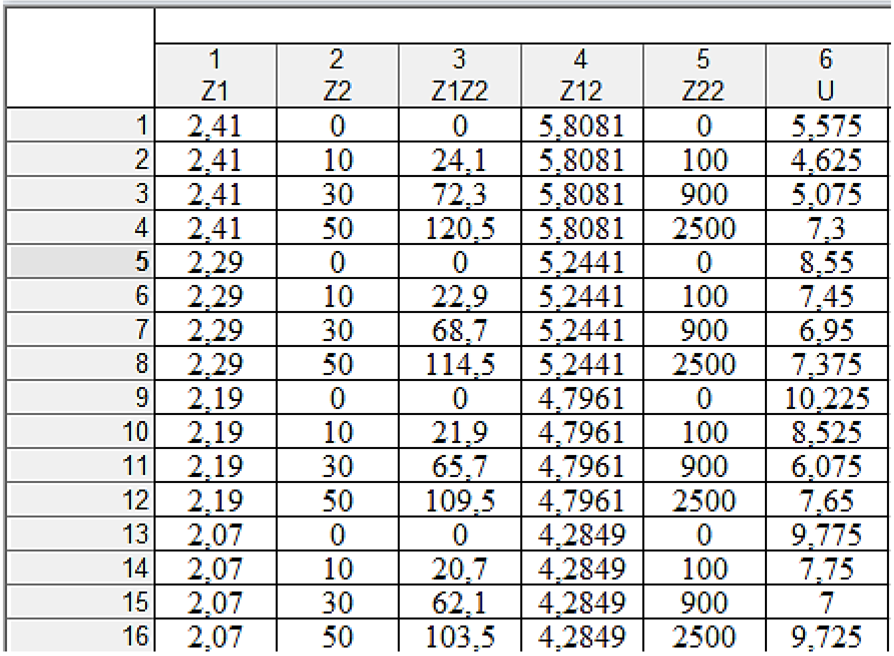
Заменить на: Z1

Искать в
○ Имена наблюдений
● Имена переменных
○ Выбранные строки
○ Выбранные столбцы

☐ Учитывать регистр
☐ Ячейка целиком

Рис. 41. Окно на экране компьютера для изменения названия переменной в Statistica

Для замены названия переменных в главном меню откроем меню «Правка». Далее в таблице исходных данных выделяем столбец 1, открываем раздел «Заменить», в окне «Найти» вводим программное имя первой переменной «Пер1», «Заменить на», вставим новое название переменной «Z1», в блоке «Искать в» выбираем «Имена переменных» и нажмем кнопку «Заменить». Аналогичным образом заменим программные названия переменных «Пер2 – Пер6» на названия остальных переменных для нашей математической регрессионной модели (рис. 42). После этого таблицу исходных данных с новыми названиями переменных сохраняем в меню «Главная» в папке «Документы» как файл с именем «Таблица РА».



	1 Z1	2 Z2	3 Z1Z2	4 Z12	5 Z22	6 U
1	2,41	0	0	5,8081	0	5,575
2	2,41	10	24,1	5,8081	100	4,625
3	2,41	30	72,3	5,8081	900	5,075
4	2,41	50	120,5	5,8081	2500	7,3
5	2,29	0	0	5,2441	0	8,55
6	2,29	10	22,9	5,2441	100	7,45
7	2,29	30	68,7	5,2441	900	6,95
8	2,29	50	114,5	5,2441	2500	7,375
9	2,19	0	0	4,7961	0	10,225
10	2,19	10	21,9	4,7961	100	8,525
11	2,19	30	65,7	4,7961	900	6,075
12	2,19	50	109,5	4,7961	2500	7,65
13	2,07	0	0	4,2849	0	9,775
14	2,07	10	20,7	4,2849	100	7,75
15	2,07	30	62,1	4,2849	900	7
16	2,07	50	103,5	4,2849	2500	9,725

Рис. 42. Окно на экране компьютера после изменения названия всех переменных в таблице исходных данных Statistica

Для выполнения расчетов параметров регрессионного анализа в Statistica в меню «Главная» открываем сохраненный файл в папке «Документы» и затем в меню «Анализ» – раздел «Множественная регрессия» (рис. 43).

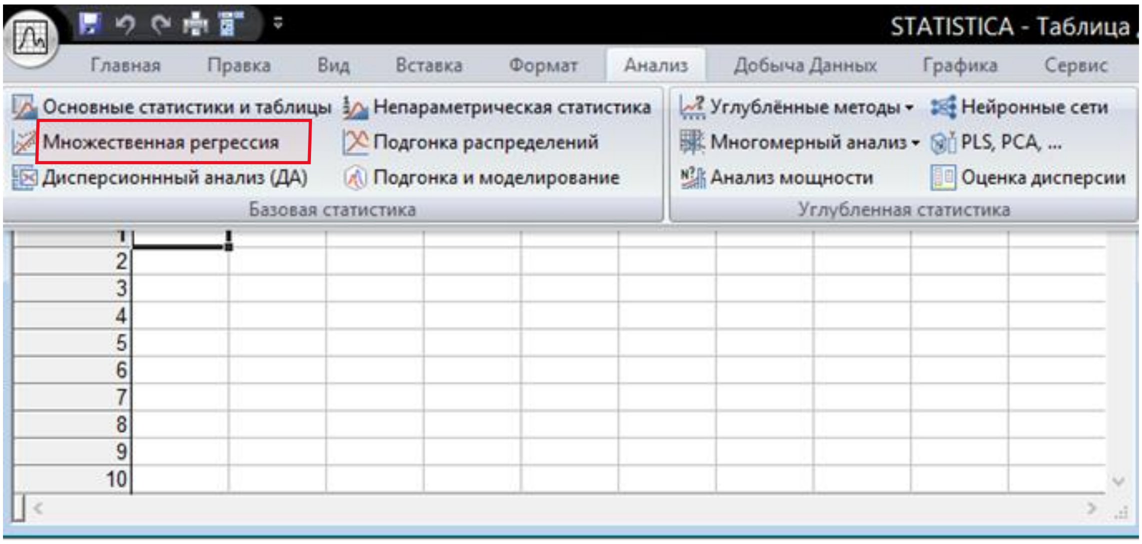
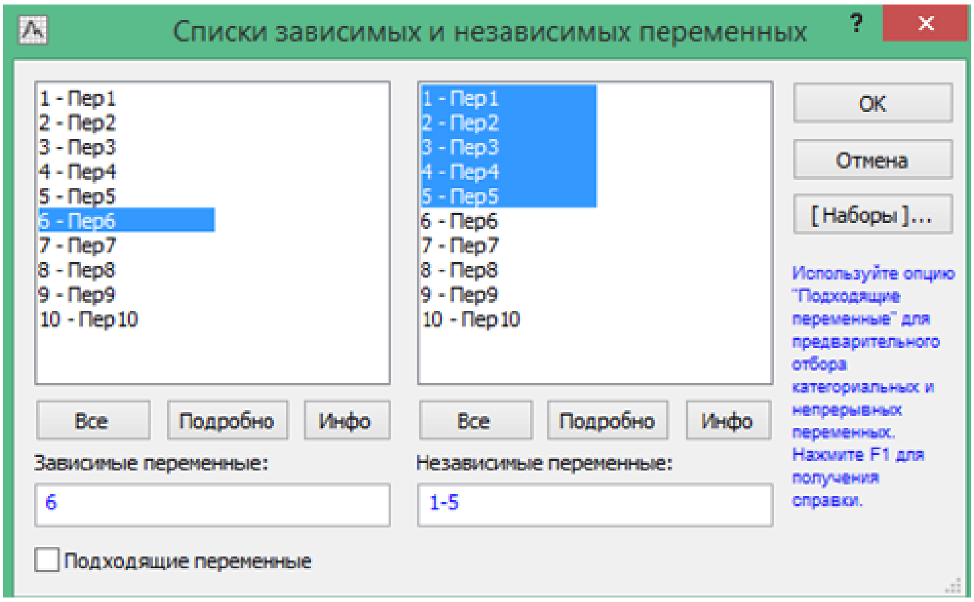


Рис. 43. Разделы меню «Анализ»

В появившемся окне «Множественная регрессия. Таблица» нажимаем кнопку «Переменные», в новом окне «Списки зависимых и независимых переменных» выбираем зависимую переменную 6 (Пер6 – U) и пять независимых переменных 1–5 (рис. 44) и нажимаем кнопку «ОК».



Используйте опцию "Подходящие переменные" для предварительного отбора категориальных и непрерывных переменных. Нажмите F1 для получения справки.

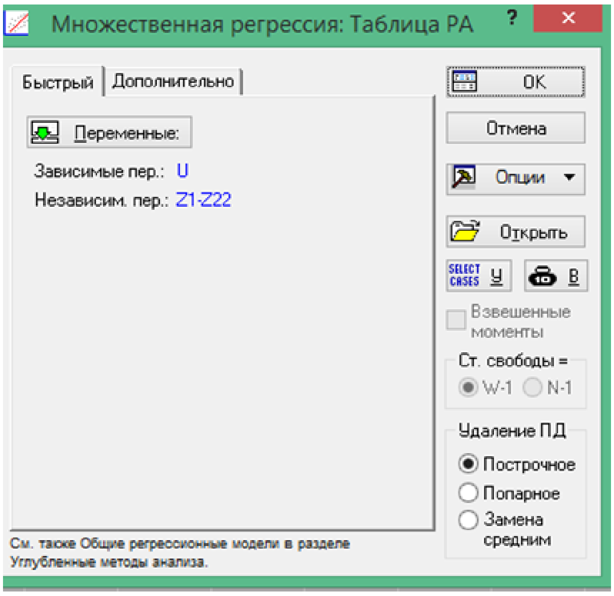
Зависимые переменные: 6

Независимые переменные: 1-5

☐ Подходящие переменные

Рис. 44. Выбор зависимых и независимых переменных

В появившемся окне (рис. 45) нажимаем кнопку «Дополнительно» и в открывшемся окне (рис. 46) выбираем команду «Пошаговая с исключением» и нажимаем кнопку «ОК».



Зависимые пер.: U

Независим. пер.: Z1-Z22

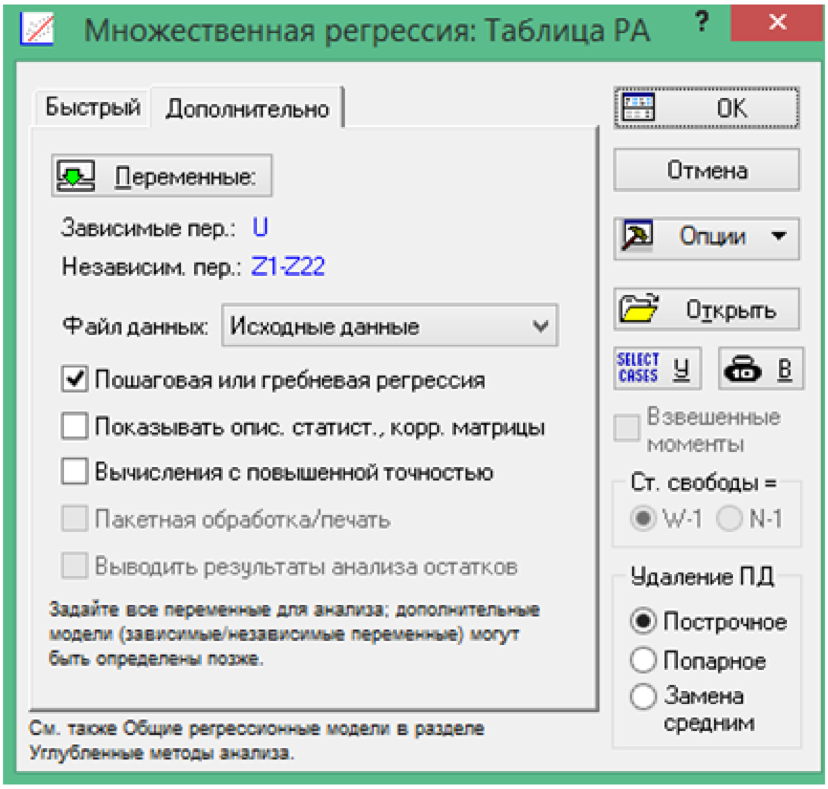
☐ Взвешенные моменты

Ст. свободы = ● W-1 ○ N-1

Удаление ПД
● Построчное
○ Попарное
○ Замена средним

См. также Общие регрессионные модели в разделе Углубленные методы анализа.

Рис. 45. Окно на экране компьютера после открытия в Statistica файла «Таблица РА»



Зависимые пер.: U

Независим. пер.: Z1-Z22

Файл данных: Исходные данные

☑ Пошаговая или гребневая регрессия

☐ Показывать опис. статист., корр. матрицы

☐ Вычисления с повышенной точностью

☐ Пакетная обработка/печать

☐ Выводить результаты анализа остатков

Задайте все переменные для анализа; дополнительные модели (зависимые/независимые переменные) могут быть определены позже.

См. также Общие регрессионные модели в разделе Углубленные методы анализа.

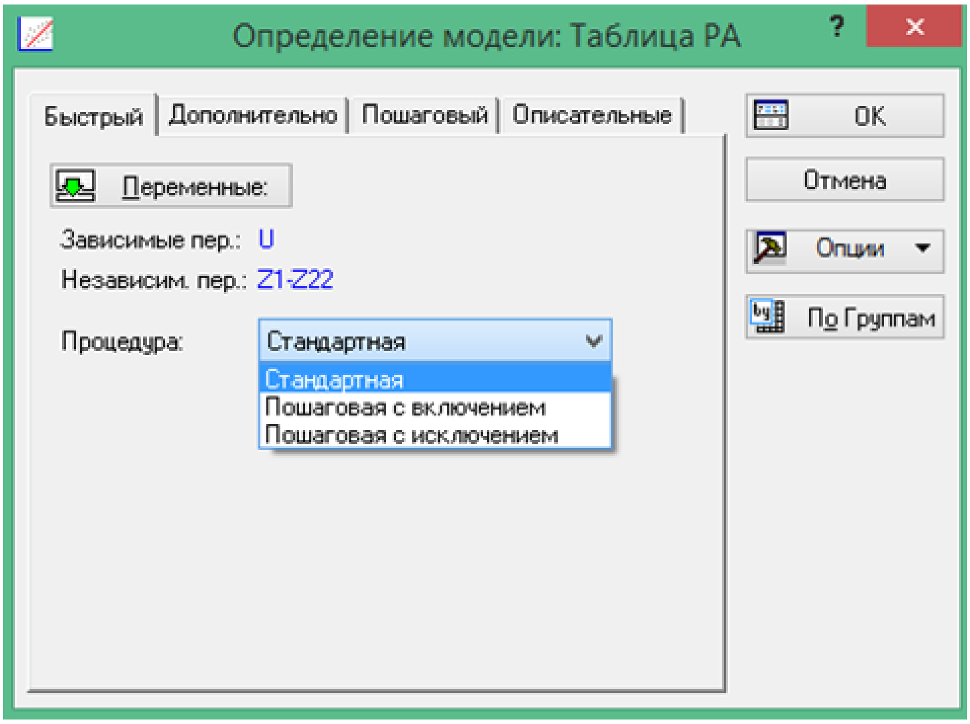
☐ Взвешенные моменты

Ст. свободы = ● W-1 ○ N-1

Удаление ПД: ● Построчное ○ Попарное ○ Замена средним

Рис. 46. Выбор регрессионного анализа

В появившемся окне (рис. 47) выбираем процедуру проведения регрессионного анализа «Пошаговая с исключением» и нажимаем кнопку «Дополнительно».



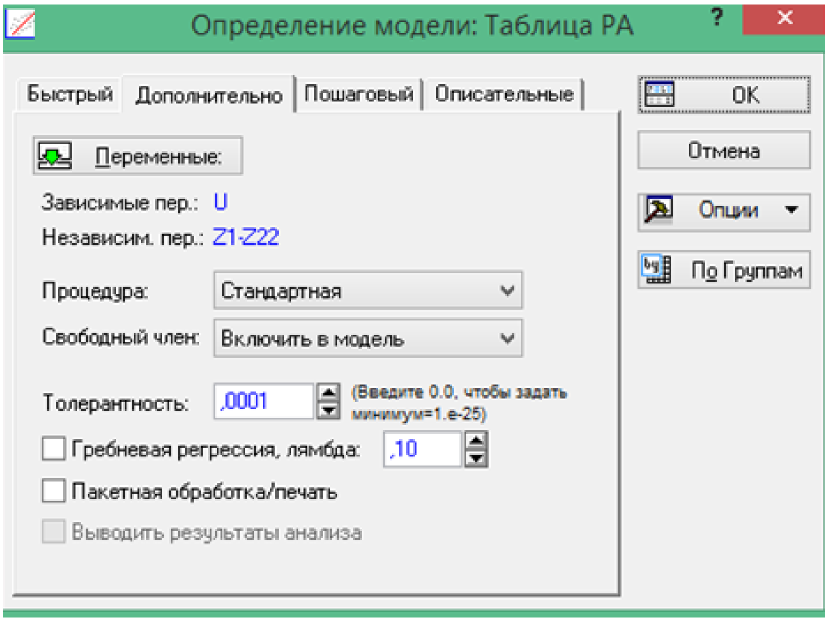
Зависимые пер.: U

Независим. пер.: Z1-Z22

Процедура: Стандартная

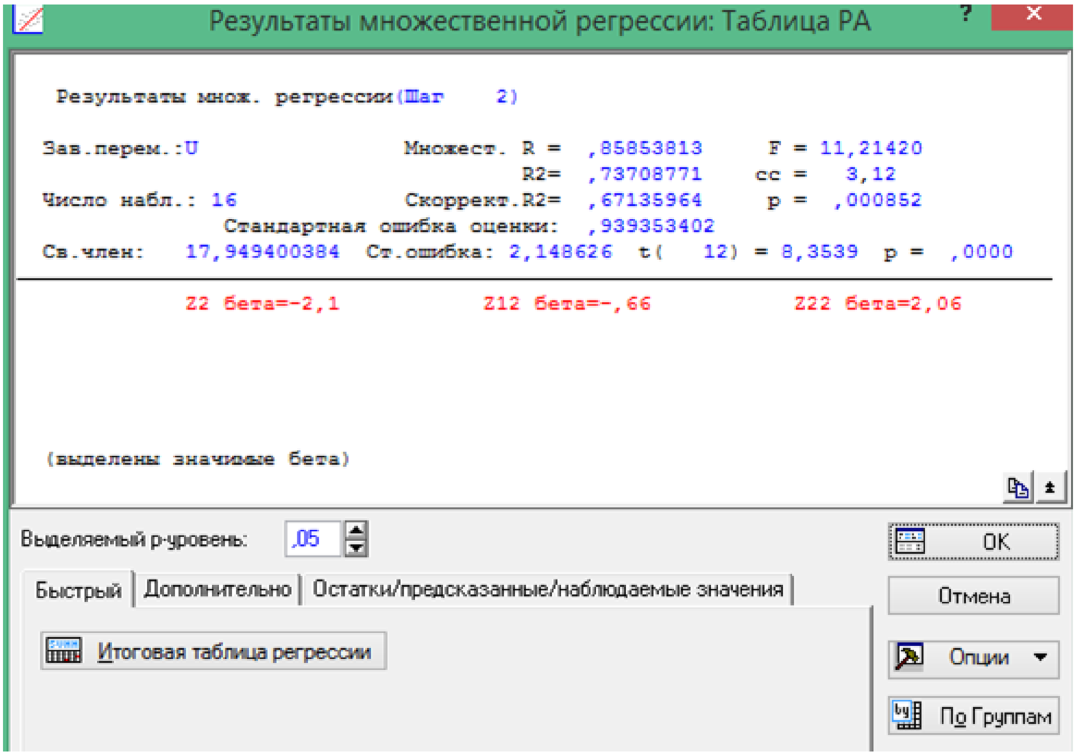
Рис. 47. Выбор процедуры регрессионного анализа

Убедившись, что в появившемся окне (рис. 48) в разделе «Свободный член» стоит команда «Включить в модель», нажимаем кнопку «ОК». В открывшемся окне «Результаты множественной регрессии. Таблица РА» (рис. 49) появляется часть рассчитанных параметров для регрессионного анализа [10].



Зависимые пер.: U
Независим. пер.: Z1-Z22
Процедура: Стандартная
Свободный член: Включить в модель
Толерантность: ,0001 (Введите 0.0, чтобы задать минимум=1.e-25)
☐ Гребневая регрессия, лямбда: ,10
☐ Пакетная обработка/печать
☐ Выводить результаты анализа

Рис. 48. Окно выбора наличия в регрессионной модели свободного члена



```
 Результаты множ. регрессии(Шаг    2)

Зав.перем.:U                Множест. R =  ,85853813     F = 11,21420
                                     R2=  ,73708771    cc =   3,12
Число набл.: 16             Скоррект.R2=  ,67135964     p =  ,000852
              Стандартная ошибка оценки:  ,939353402
Св.член:   17,949400384  Ст.ошибка: 2,148626  t(   12) = 8,3539  p =  ,0000
------------------------------------------------------------------------
           Z2 бета=-2,1           Z12 бета=-,66           Z22 бета=2,06



(выделены значимые бета)
```

Выделяемый p-уровень: ,05

Рис. 49. Результаты множественной регрессии (шаг 2)

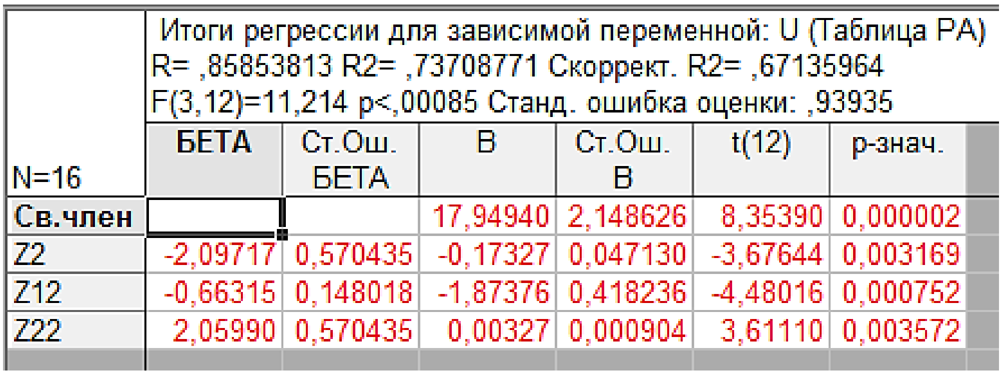
В верхней части окна (см. рис. 49) представлены следующие статистические параметры:

- множест. *R* (коэффициент множественной корреляции);
- *R2* (коэффициент детерминации);
- скоррект. *R2* (скорректированный коэффициент детерминации);
- стандартная ошибка оценки;
- оценка свободного члена B_0;
- ст. ошибка (стандартная ошибка оценки свободного члена B_0);
- *t* (значение *t-критерия*) и *p* для проверки гипотезы о равенстве нулю свободного члена B_0;
- *F* (значение *F-критерия*), *сс* – (число степеней свободы df_1 и df_2), *p* (уровень значимости *p*) для проверки зависимости *U* от всех параметров регрессионной модели.

В средней части окна выделены красным цветом значимые для выбранной доверительной вероятности стандартизованные регрессионные коэффициенты **бета** параметров регрессионной модели. Коэффициенты **бета** позволяют сравнивать вклады каждого параметра регрессионной модели в предсказание значения отклика объекта *U*.

В нижней части окна находятся функциональные кнопки, позволяющие всесторонне просмотреть результаты статистического анализа данных [10]. Если в нижней части окна нажать кнопку «Итоговая таблица регрессии», то появятся дополнительные рассчитанные параметры для регрессионного анализа (рис. 50):

- Ст. Ош. БЕТА (стандартная ошибка оценки коэффициентов бета);
- В (значения коэффициентов *B* параметров регрессионной модели);
- Ст. Ош. В (стандартная ошибка оценки коэффициентов *B*);
- *t* (значение критерия Стьюдента для регрессионной модели);
- *p*-значение (уровень значимости коэффициентов *B*).



Итоги регрессии для зависимой переменной: U (Таблица РА)
R= ,85853813 R2= ,73708771 Скоррект. R2= ,67135964
F(3,12)=11,214 p<,00085 Станд. ошибка оценки: ,93935

N=16	БЕТА	Ст.Ош. БЕТА	B	Ст.Ош. B	t(12)	p-знач.
Св.член			17,94940	2,148626	8,35390	0,000002
Z2	-2,09717	0,570435	-0,17327	0,047130	-3,67644	0,003169
Z12	-0,66315	0,148018	-1,87376	0,418236	-4,48016	0,000752
Z22	2,05990	0,570435	0,00327	0,000904	3,61110	0,003572

Рис. 50. Итоги регрессии для зависимой переменной *U*

Результаты расчетов в MS Excel значений статистических параметров для регрессионного анализа (см. рис. 37) полностью совпадают с результатами расчетов в Statistica (см. рис. 49, 50). Дополнительные статистические параметры, полученные в Statistica, связанные со стандартизованными регрессионными коэффициентами бета, позволяют сделать следующий дополнительный вывод регрессионного анализа. Так, из абсолютных величин значений коэффициентов бета следует расположить параметры регрессионной модели по силе влияния на свойство образца U в следующий ряд: $Z2 > Z22 > Z12$.

Получение регрессионных моделей исследуемого объекта имеет не только важное научное, но и практическое значение, например:

- для управления технологическими процессами производства продукции с использованием исследованного объекта;
- решения производственных задач для выбора оптимальных значений технологических факторов при использовании исследованного объекта, обеспечивающих получение продукции с необходимыми свойствами.

4.5. Решение оптимизационных задач

На обложке книги Б. Я. Курицкого «Поиск оптимальных решений средствами MS Excel 7.0» [6] приведена надпись «Если вы хотите получить наилучший результат, принимайте только оптимальные решения!» С этим выражением нужно согласиться, так как поиск оптимальных решений – одна из важных задач не только в научных исследованиях, но и в производственной деятельности предприятий и организаций различных отраслей экономики [6, 12–16].

Процесс решения оптимизационных задач может быть неформализованным и формализованным. Довольно часто человек принимает неформализованные решения, руководствуясь только своей логикой, опытом и интуицией. Никакой гарантии правильности неформализованных решений нет. Значительно большие гарантии правильности имеют формализованные решения, основанные на научном подходе.

Для формализованных решений оптимизационных задач необходимо иметь критерий оптимизации (целевую функцию), т. е. желательное значение свойства (отклика) объекта.

При этом возможны три вида критерия оптимизации:

- максимальное значение;
- минимальное значение;
- заданное значение.

Наиболее распространенными в производственной технологической деятельности являются задачи:

- оптимального планирования;
- оптимального управления технологическими процессами;
- оптимального проектирования.

Задачи оптимального планирования в большинстве случаев возникают при оценке экономической эффективности производства при распределении ресурсов: финансов, сырья, работников и др. Например, такая задача может быть поставлена для оценки максимальной прибыли при производстве нескольких видов готовой продукции, получаемой из одного сырья.

Для управления технологическими процессами требуется знание сочетания значений технологических факторов, обеспечивающих наилучшее качество готовой продукции, минимальную себестоимость ее изготовления или экологическое воздействие на окружающую среду. Для этих целей решают задачи оптимального управления технологическими процессами.

Задачи оптимального проектирования решают при разработке технологического оборудования, например при создании оборудования с минимальной массой при заданных технических характеристиках.

Оптимизационные задачи – это задачи поиска значений независимых переменных, обеспечивающих наилучшие или заданные значения свойств (отклика) объекта.

В качестве элементов технического языка при решении оптимизационных задач будем использовать термины и определения, описанные в учебнике [17].

При поиске и принятии оптимальных решений выполняют такую последовательность этапов работы.

1. Содержательная постановка и формализация оптимизационной задачи.
2. Получение исходных данных и адекватной математической модели объекта.
3. Формализованное решение оптимизационной задачи.
4. Анализ решений и принятие оптимального решения.

Поиск оптимальных решений рекомендуется начинать с анализа и формулировки основной цели решения оптимизационной задачи.

При постановке оптимизационной задачи необходимо выбрать одну из следующих целей ее решения:

- что будет, если ...?
- что необходимо сделать для того, чтобы ...?

Для достижения первой цели необходимо использовать методы **безусловной оптимизации** (методы вариантного анализа), а для второй цели – методы **условной оптимизации** (методы решения по заказу).

После формулировки цели и метода решения оптимизационной задачи нужно составить ее математическую модель, состоящую из целевой функции (ЦФ), ограничений (ОГР) и граничных условий для независимых переменных (ГРУ).

При формализации оптимизационной задачи составляют и анализируют ее математическую модель, обязательно содержащую в своем составе хотя бы одну целевую функцию, которая должна представлять собой математическую зависимость параметра объекта от значений независимых переменных (входных факторов), влияющих на этот параметр.

Назначения целевой функции (параметра объекта) записывают символами:

max (максимизация);

min (минимизация);

const (заданное значение).

Целевую функцию с ее назначением называют **функцией желательности *G***.

Задача **безусловной оптимизации** представляет собой поиск оптимума целевой функции без всяких дополнительных условий (ограничений и граничных условий). **Математическая модель** такой однокритериальной задачи представляется только функцией желательности:

$$\boldsymbol{G}\,(x_i) \rightarrow \max\ (\min,\ \mathrm{const}),$$

где x_i – независимые переменные (входные факторы), $i = 1, \ldots, n$.

Для задачи **условной оптимизации** в ее математической модели обязательно должны присутствовать условия в виде ограничений и (или) граничных условий. **Математические модели** однокритериальной задачи условной оптимизации могут иметь следующий вид:

$$\left.\begin{aligned} &G(x_i) \to \max\,(\min,\,\text{const}),\\ &g_j(x_i) \le (<;\ =\ ;\ >;\ \ge)\ b_j,\\ &d_i \le x_i \le D_i, \end{aligned}\right\}$$

$$\left.\begin{aligned} &G(x_i) \to \max\,(\min,\,\text{const}),\\ &g_j(x_i) \le (<;\ =\ ;\ >;\ \ge)\ b_j, \end{aligned}\right\}$$

$$\left.\begin{aligned} &G(x_i) \to \max\,(\min,\,\text{const}),\\ &d_i \le x_i \le D_i, \end{aligned}\right\}$$

где $G\ (x_i)$ – целевая функция;

x_i – независимые переменные (входные факторы);

$g_j\ (x_i)$ – левые части ограничений (ОГР) в виде математических зависимостей;

b_j – правые части ограничений в виде чисел;

d_i и D_i (числа) – соответственно левая и правая границы факторной области (ГРУ) – допустимых значений переменных: $x_i\ (i = 1, ..., n)$, $g_j = 1, ..., m$.

Ограничения g_j могут быть в форме равенств или неравенств, переменные x_i – непрерывные или дискретные (в том числе целочисленные).

После составления математической модели оптимизационной задачи проводят ее анализ по следующим параметрам модели:

- вид оптимизационной задачи (одно- или многокритериальная);
- критерии оптимизации и пределы, в которых могут находиться значения критериев оптимизации, ограничений и граничных условий;
- вид значений переменных x_j (детерминированные или случайные, непрерывные или дискретные величины);
- вид отклика объекта y (непрерывная или дискретная величина);
- вид математических зависимостей целевой функции и ограничений (линейные или нелинейные).

Важной характеристикой математической модели задачи оптимизации является ее размерность, определяемая числом переменных n и ограничений m. От соотношения чисел n и m зависит существование решения оптимизационной задачи [17].

Задача оптимизации имеет оптимальное решение, если существует несколько допустимых решений. Допустимым решением (ДР) называется решение, удовлетворяющее всем ограничениям и граничным условиям.

У оптимизационной задачи имеется несколько допустимых решений, если выполняется соотношение $n > m$ или когда все ограничения являются неравенствами.

В зависимости от параметров математической модели оптимизационной задачи проводят ее классификацию для выбора метода решения. Сочетание различных элементов модели образует различные классы задач оптимизации, которые требуют разных методов решения [17].

При классификации оптимизационных задач для выбора метода решения их подразделяют на задачи программирования:

- линейного;
- нелинейного;
- целочисленного;
- стохастического;
- и др.

К **задачам линейного программирования (ЛП)** относят такие оптимизационные задачи, в которых целевая функция и все ограничения являются линейными.

Общий вид математической модели задач ЛП представляется следующим образом:

$$\left.\begin{aligned}
&G\left(x_i\right) = \sum\nolimits_{i=1}^{n} c_i x_i \to \max\left(\min, \mathrm{const}\right),\\
&g_j\left(x_i\right) = \sum\nolimits_{i=1}^{n} a_i x_i \le \ (<;\ =;\ \ge;\ >)\ b_j,\\
&d_i \le x_i \le D_i,\\
&i = 1, \ldots, n;\ j = 1, \ldots, m.
\end{aligned}\right\}$$

К **задачам нелинейного программирования (НЛП)** относят такие оптимизационные задачи, в которых целевая функция или одно из ограничений являются нелинейными. Существуют и другие названия задач оптимизации [17]: стохастического программирования, статичные, динамического программирования, дискретного программирования, целочисленного программирования.

Рассмотрим пример решения оптимизационной задачи с применением MS Excel.

Пример 7. Необходимо спроектировать бак объемом 2000 м^3, имеющий форму прямоугольного параллелепипеда, чтобы на его изготовление пошло как можно меньше материала. Такая задача относится к задачам оптимального проектирования.

Проведем анализ, формализацию и содержательную постановку задачи. Анализ показывает, что задача относится к задачам условной оптимизации (задачам по заказу), так как имеется одно ограничение (заданное значение объема бака, V = 2000 м^3) и три граничных условия (значения длины, ширины и высоты бака должны быть только положительными числами, отличными от нуля).

Выберем в качестве целевой функции площадь бака, т. е. площадь материала для его изготовления (S, м^2), а в качестве переменных – длину (a, м), ширину (b, м) и высоту (h, м) бака. Тогда зависимости S и V от данных переменных будут следующими:

$$S\left(a,\, b,\, h\right) = 2[ab + \left(a + b\right)h];$$
$$V\left(a,\, b,\, h\right) = abh.$$

Математическая модель оптимизационной задачи будет иметь следующий вид:

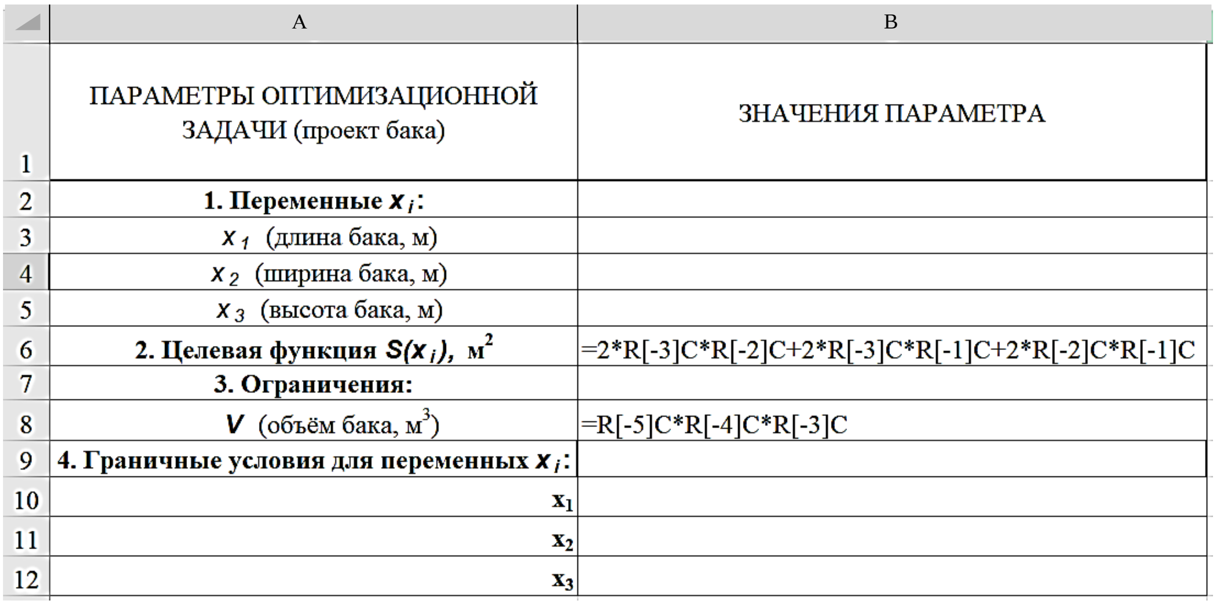
$$\left.\begin{array}{l} S\left(a,\, b,\, h\right) = 2ab + 2ah + 2bh = 2\left[ab + \left(a + b\right)h\right] \rightarrow \min\left(G\ (x_i)\right), \\ V\left(x_i\right) = abh = 2000\ (\text{ОГР}), \\ a,\, b,\, h > 0\ (\textit{ГРУ}). \end{array}\right\}$$

Анализ данной модели показывает следующее:

- задача однокритериальная – одна целевая функция (S);
- задача трехпараметрическая – три детерминированные непрерывные переменные: длина (a, м), ширина (b, м) и высота (h, м) бака;
- отклик объекта S – непрерывная величина;
- задача относится к классу задач нелинейного программирования (S и V – нелинейные зависимости);
- задача имеет множество решений, так как число переменных (n = 3) больше числа ограничений (m = 1);
- для решения задачи можно применить аналитические, графический и численные методы решения.

Применим численные методы решения оптимизационных задач, запрограммированные в MS Excel. Первоначально сформируем в Excel файл с исходными данными для расчетов (рис. 51).

В ячейку B6 вводим формулу расчета площади бака *S* и в ячейку B8 – объема бака с нулевыми значениями переменных x_i. Таблица в файле Excel с исходными данными для расчетов с показом формул приведена на рис. 51.



	A	B
1	ПАРАМЕТРЫ ОПТИМИЗАЦИОННОЙ ЗАДАЧИ (проект бака)	ЗНАЧЕНИЯ ПАРАМЕТРА
2	**1. Переменные x_i:**	
3	x_1 (длина бака, м)	
4	x_2 (ширина бака, м)	
5	x_3 (высота бака, м)	
6	**2. Целевая функция $S(x_i)$, $м^2$**	=2*R[-3]C*R[-2]C+2*R[-3]C*R[-1]C+2*R[-2]C*R[-1]C
7	**3. Ограничения:**	
8	V (объём бака, $м^3$)	=R[-5]C*R[-4]C*R[-3]C
9	**4. Граничные условия для переменных x_i:**	
10	x_1	
11	x_2	
12	x_3	

Рис. 51. Файл в Excel с исходными данными для расчетов при решении оптимизационной задачи для примера 7 с показом формул

Затем в меню «Главная» открываем раздел «Данные» и выбираем надстройку «Поиск решения» (рис. 52). На открывшемся экране «Параметры поиска решения» (рис. 53) в окне «Оптимизировать целевую функцию:» вводим адрес ячейки B6 (см. рис. 51), а в окне «Изменяя ячейки переменных:» вводим диапазон ячеек B3:B5 (см. рис. 51) с нулевыми значениями переменных x_i., а в строке назначения целевой функции «До:» выбираем команду «Минимум».

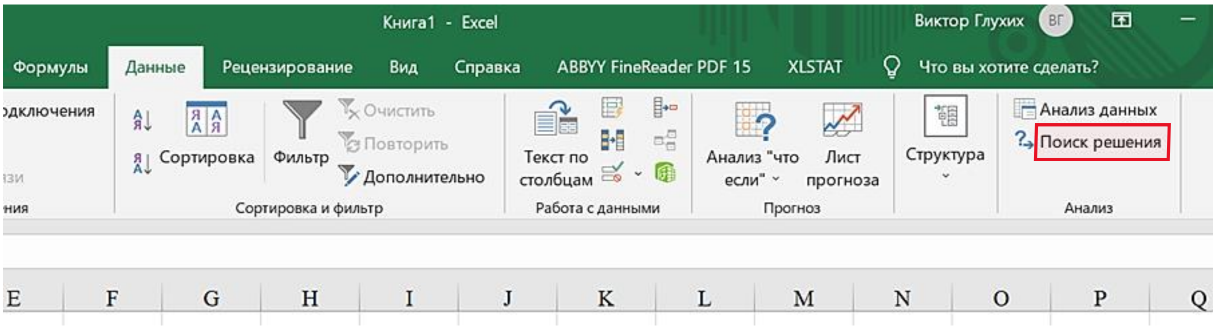
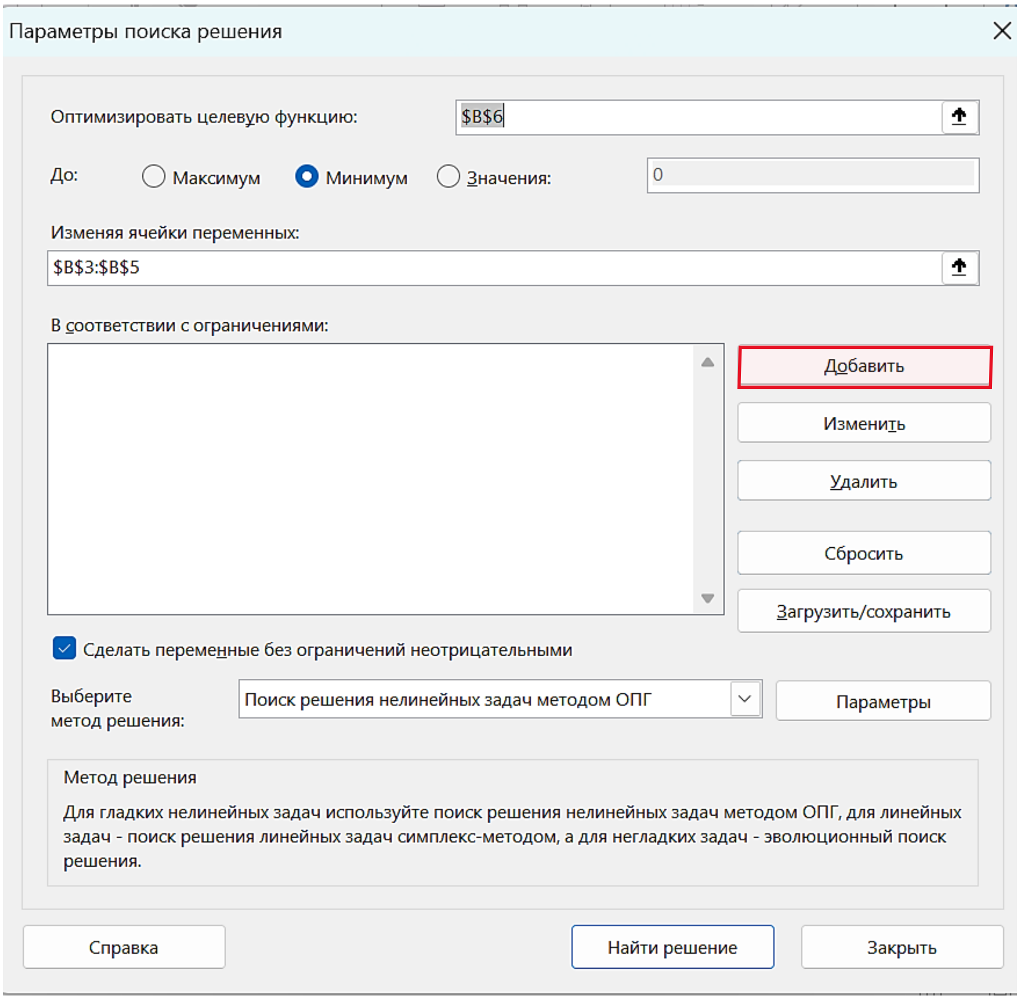


Рис. 52. Выбор надстройки «Поиск решения»



Оптимизировать целевую функцию: $B$6

До: ○ Максимум ● Минимум ○ Значения: 0

Изменяя ячейки переменных: $B$3:$B$5

В соответствии с ограничениями:

☑ Сделать переменные без ограничений неотрицательными

Выберите метод решения: Поиск решения нелинейных задач методом ОПГ

Метод решения

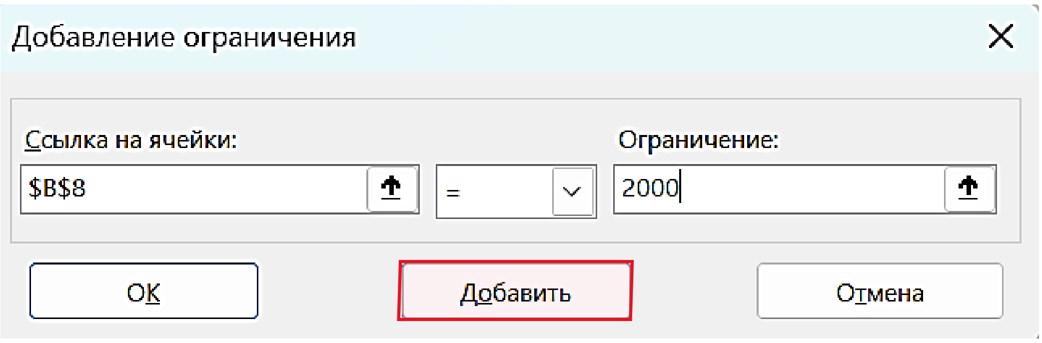
Для гладких нелинейных задач используйте поиск решения нелинейных задач методом ОПГ, для линейных задач - поиск решения линейных задач симплекс-методом, а для негладких задач - эволюционный поиск решения.

Рис. 53. Выбор параметров поиска решения оптимизационной задачи для примера 7

Не изменяем имеющиеся в этом окне (по умолчанию) параметры «Сделать переменные без ограничений неотрицательными» и «Поиск решения нелинейных задач методом ОПГ», так как все линейные размеры бака должны быть только положительными, а целевая функция S (a, b, h) является нелинейной и гладкой.

После этого нажимаем на кнопку «Добавить» и в появившемся новом окне для ввода ограничений (рис. 54) последовательно вводим адрес ячеек (см. рис. 51). Сначала адрес ячейки (B8) с формулой расчета ограничения V и его значение 2000. После этого в этом окне нажимаем кнопку «Добавить» для ввода значений граничных условий для всех независимых переменных x_i: a (B3), b (B4), h (B5).

В процедуре ввода ограничений в надстройке «Поиск решений» нет соотношения « $>$ » левой и правой части ограничений, есть только соотношение « $\geq$ ». Поэтому с данным соотношением в правой части граничных условий для a (B3), b (B4), h (B5) ставим любое число, близкое к нулю, например 0,001. Завершив ввод ограничения и граничных условий в этом окне, нажимаем кнопку «ОК». В появившемся обновленном окне (рис. 55) нажимаем кнопку «Найти решение». Появляется окно с информацией о наличии решения оптимизационной задачи (рис. 56).

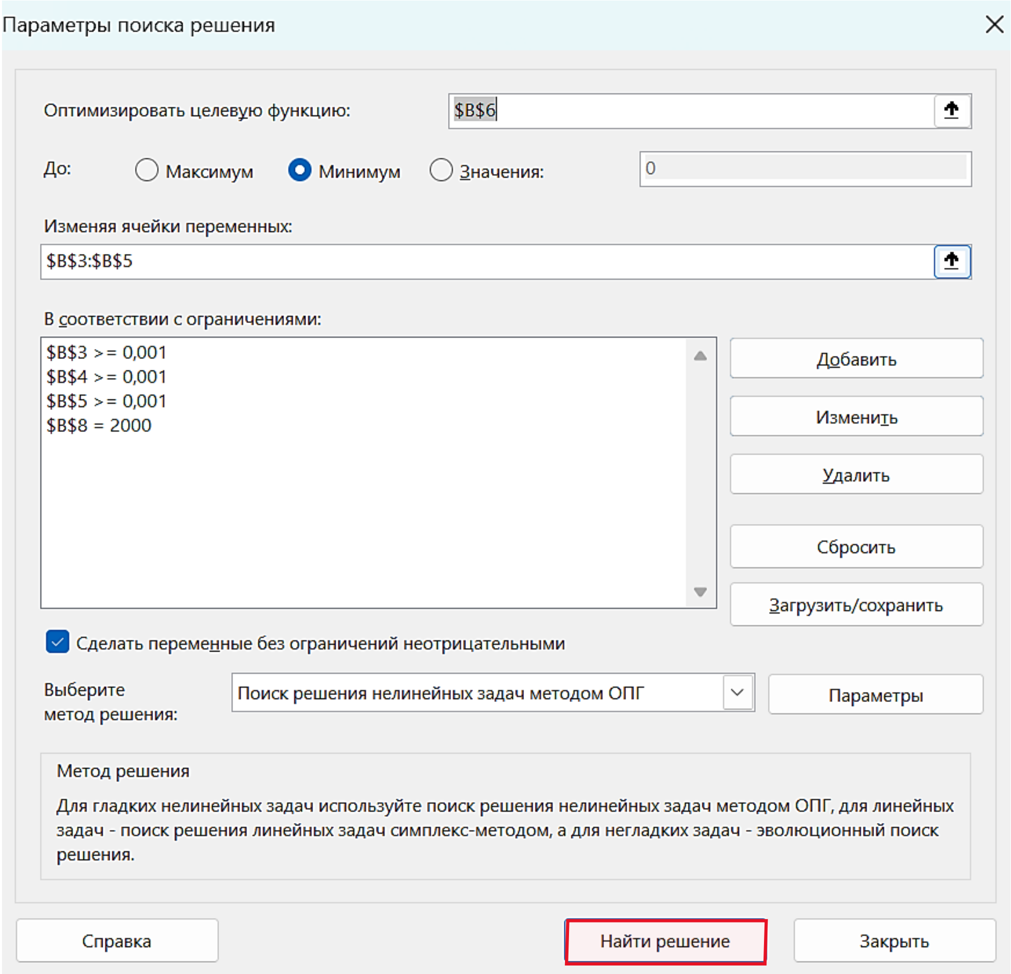

Добавление ограничения

Ссылка на ячейки: $B$8

=

Ограничение: 2000

Рис. 54. Добавление ограничений и граничных условий для примера 7


Параметры поиска решения

Оптимизировать целевую функцию: $B$6

До: ○ Максимум ● Минимум ○ Значения: 0

Изменяя ячейки переменных: $B$3:$B$5

В соответствии с ограничениями:

$B$3 >= 0,001
$B$4 >= 0,001
$B$5 >= 0,001
$B$8 = 2000

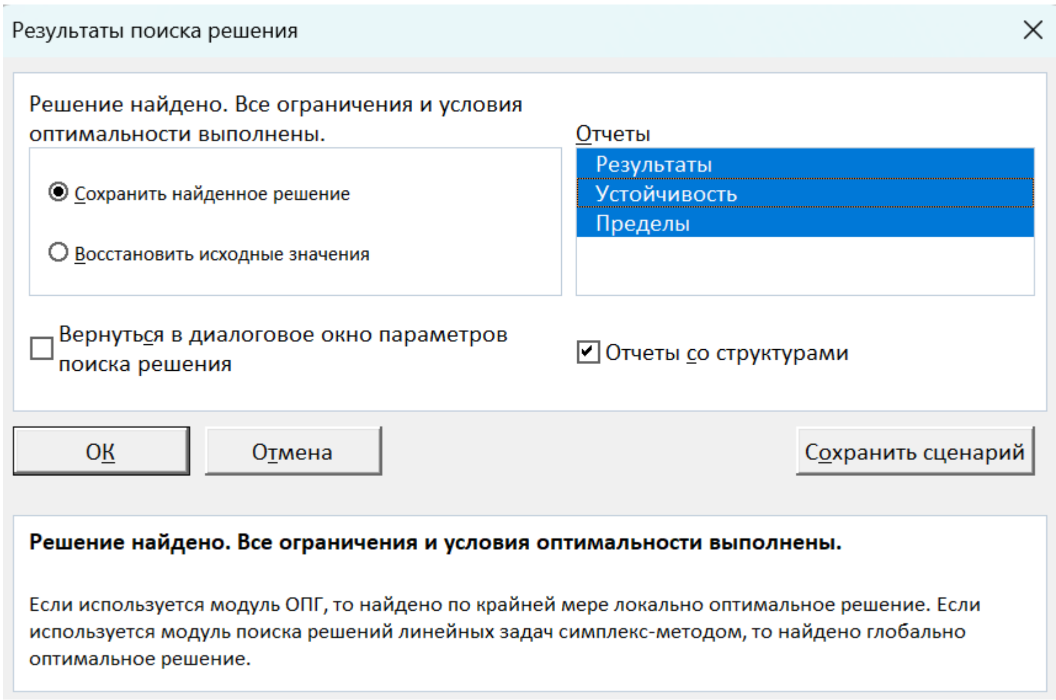
☑ Сделать переменные без ограничений неотрицательными

Выберите метод решения: Поиск решения нелинейных задач методом ОПГ

Метод решения

Для гладких нелинейных задач используйте поиск решения нелинейных задач методом ОПГ, для линейных задач - поиск решения линейных задач симплекс-методом, а для негладких задач - эволюционный поиск решения.

Рис. 55. Параметры поиска решения со всеми ограничениями для примера 7



Решение найдено. Все ограничения и условия оптимальности выполнены.

● Сохранить найденное решение

○ Восстановить исходные значения

☐ Вернуться в диалоговое окно параметров поиска решения

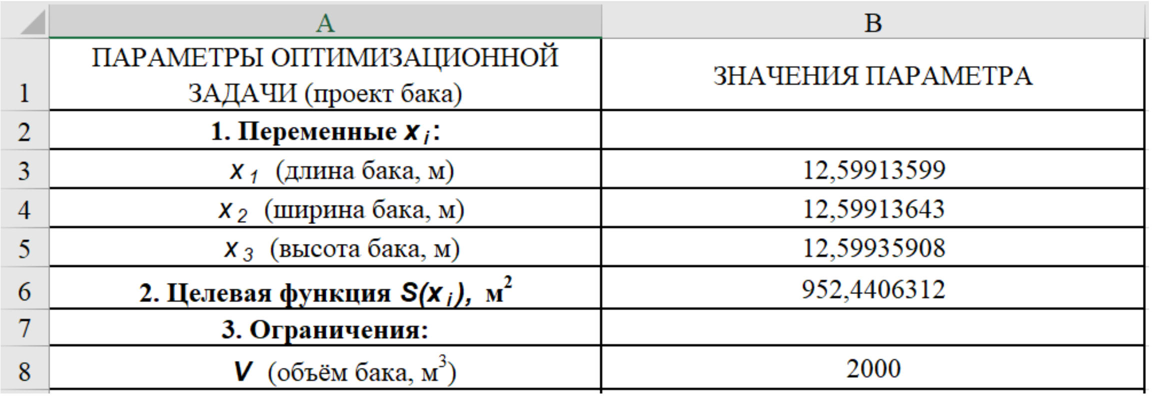
☑ Отчеты со структурами

Решение найдено. Все ограничения и условия оптимальности выполнены.

Если используется модуль ОПГ, то найдено по крайней мере локально оптимальное решение. Если используется модуль поиска решений линейных задач симплекс-методом, то найдено глобально оптимальное решение.

Рис. 56. Информация о наличии решения оптимизационной задачи для примера 7

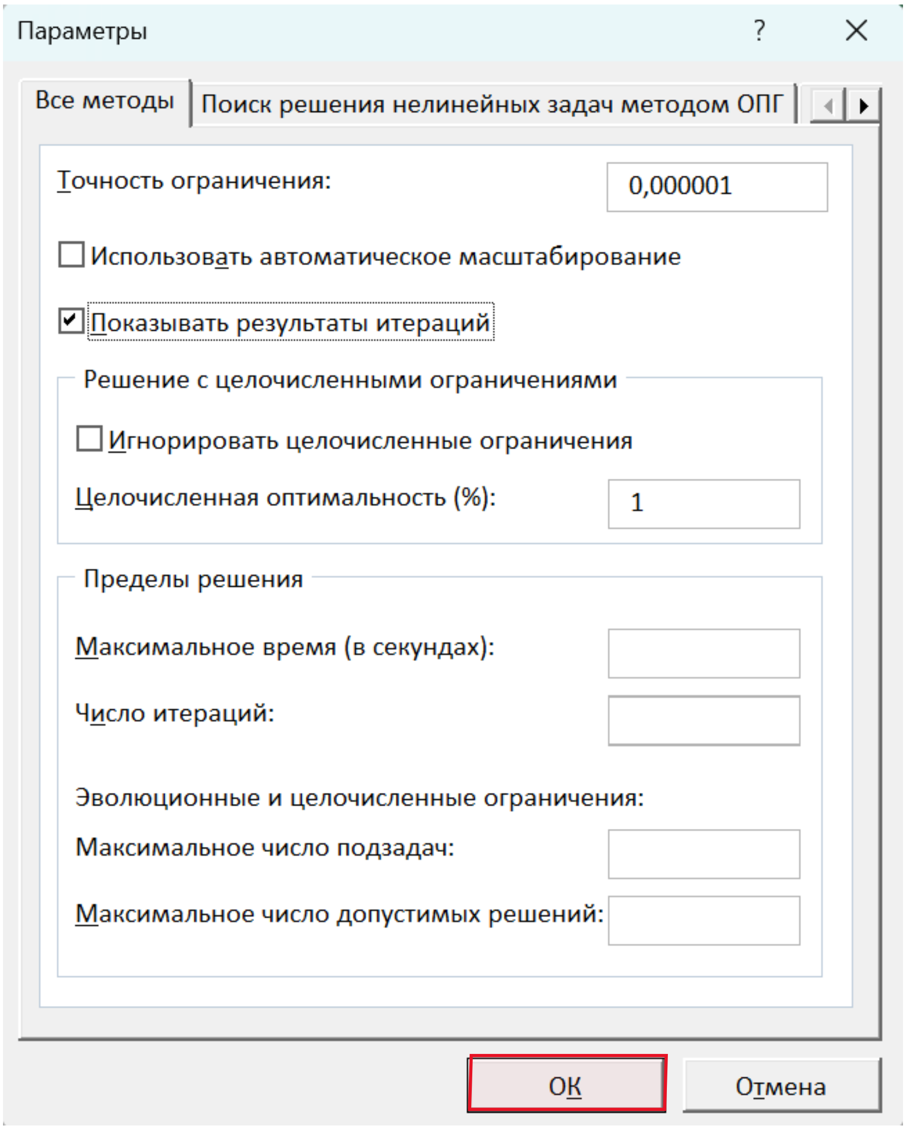
В нижней части окна на рис. 56 выводится сообщение «Если используется модуль ОПГ, то найдено по крайней мере локально оптимальное решение». Для решения данной оптимизационной задачи был использован модуль ОПГ, поэтому необходимо проверить, что найденное решение (рис. 57) является не локальным, а глобальным оптимумом.



	A	B
1	ПАРАМЕТРЫ ОПТИМИЗАЦИОННОЙ ЗАДАЧИ (проект бака)	ЗНАЧЕНИЯ ПАРАМЕТРА
2	**1. Переменные x_i:**	
3	x_1 (длина бака, м)	12,59913599
4	x_2 (ширина бака, м)	12,59913643
5	x_3 (высота бака, м)	12,59935908
6	**2. Целевая функция $S(x_i)$, м^2**	952,4406312
7	**3. Ограничения:**	
8	V (объём бака, м^3)	2000

Рис. 57. Результат решения оптимизационной задачи для примера 7

Для этого в окне (см. рис. 55) необходимо нажать кнопку «Параметры» и в появившемся окне (рис. 58) поставить отметку в квадратике «Показывать результаты итераций» и нажать кнопку «ОК».



Точность ограничения: 0,000001

☐ Использовать автоматическое масштабирование

☑ Показывать результаты итераций

☐ Игнорировать целочисленные ограничения

Целочисленная оптимальность (%): 1

Максимальное время (в секундах):

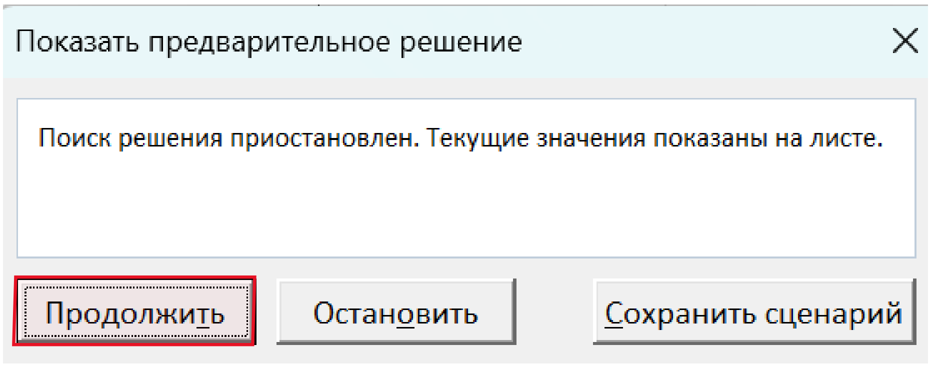
Число итераций:

Максимальное число подзадач:

Максимальное число допустимых решений:

Рис. 58. Окно «Параметры»

В открывшемся окне (рис. 59) нажимая кнопку «Продолжить», просматриваем и фиксируем результаты всех оптимумов до выхода в окно «Результат решения оптимизационной задачи» (см. рис. 57).



Поиск решения приостановлен. Текущие значения показаны на листе.

Рис. 59. Окно «Показать предварительное решение»

Для решаемой оптимизационной задачи для примера 7 из 28 оптимумов глобальным был оптимум, представленный на рис. 57.

Для дополнительного анализа полученного решения оптимизационной задачи можно в окне «Информация о наличии решения оптимизационной задачи» (см. рис. 56) до нажатия кнопки «ОК» сделать отметку в квадратике «Отчеты со структурами» и получить отчеты «Результаты», «Устойчивость» и «Пределы», которые будут представлены на отдельных листах книги с решением оптимизационной задачи (табл. 13–15).

Таблица 13

Отчет о результатах для примера 7

Результат: Решение найдено. Все ограничения и условия оптимальности выполнены					
Модуль поиска решения					
Параметры поиска решения					
Ячейка целевой функции (Минимум)					
Ячейка	Имя	Исходное значение	Окончательное значение		
\$B\$6	2. Целевая функция $S(x_i)$, м^2 ЗНАЧЕНИЯ ПАРАМЕТРА	0	952,441		
Ячейки переменных					
Ячейка	Имя	Исходное значение	Окончательное значение	Целочисленное	
\$B\$3:\$B\$5					
Ограничения					
Ячейка	Имя	Значение ячейки	Формула	Состояние	Допуск
\$B\$8	V (объем бака, м^3) ЗНАЧЕНИЯ ПАРАМЕТРА	2000	\$B\$8=2000	Привязка	0
\$B\$3	$x1$ (длина бака, м) ЗНАЧЕНИЯ ПАРАМЕТРА	12,599	\$B\$3=2000	Без привязки	12,60
\$B\$4	$x2$ (ширина бака, м) ЗНАЧЕНИЯ ПАРАМЕТРА	12,599	\$B\$4=2000	Без привязки	12,60
\$B\$5	$x3$ (высота бака, м) ЗНАЧЕНИЯ ПАРАМЕТРА	12,599	\$B\$5=2000	Без привязки	12,60

Таблица 14

Отчет об устойчивости для примера 7

Ячейки переменных		Окончательное значение	Приведенный градиент
Ячейка	Имя		
	$B$3:$B$5		
	Ограничения	Окончательное значение	Множитель Лагранжа
Ячейка	Имя		
$B$8	V (объем бака, м^3) ЗНАЧЕНИЯ ПАРАМЕТРА	2000	0,32

Таблица 15

Отчет о пределах для примера 7

Ячейка	Целевая функция	Значение				
$B$6	2. Целевая функция $S(x_i)$, м^2 ЗНАЧЕНИЯ ПАРАМЕТРА	952,44				
	Переменная		Нижний предел	Целевая функция	Верхний предел	Целевая функция
Ячейка	Имя	Значение		Результат		Результат
$B$3	x_1 (длина бака, м) ЗНАЧЕНИЯ ПАРАМЕТРА	12,59	12,59	952,44	12,59	952,44
$B$4	x_2 (ширина бака, м) ЗНАЧЕНИЯ ПАРАМЕТРА	12,59	12,59	952,44	12,59	952,44
$B$5	x_3 (высота бака, м) ЗНАЧЕНИЯ ПАРАМЕТРА	12,59	12,59	952,44	12,59	952,44

Отчет о результатах (см. табл. 13) содержит сведения о целевой функции. Во всех разделах таблицы в столбце «Ячейка» приводятся адреса ячеек целевой функции, независимых переменных, ограничения и граничных условий математической модели оптимизационной задачи в таблице исходных данных (см. рис. 51). В столбце «Имя» – название параметров математической модели оптимизационной задачи в таблице исходных данных (см. рис. 51). В первом разделе таблицы «Ячейка целевой функции (Минимум)» в столбце «Исходное значение» приведены значения целевой функции до начала вычислений, а в столбце «Окончательное значение» – найденное значение.

В третьем разделе таблицы «Ограничения» в столбцах приводятся следующие параметры:

- в столбце «Значение ячейки» показываются значения ограничения и граничных условий для полученного оптимального решения;
- в столбце «Формула» – адреса ячеек (см. рис. 51), значения чисел в которых приводят к получению оптимального решения;
- в столбце «Состояние» – используются ли значения независимых переменных для расчета данных параметров (используются – «Привязка», не используются – «Без привязки»);
- в столбце «Допуск» – допустимые отклонения параметра от заданного начального значения.

Отчет об устойчивости (см. табл. 14) состоит из двух разделов: «Ячейки переменных» и «Ограничения».

В обоих разделах таблицы в столбце «Ячейка» приводятся адреса ячеек независимых переменных и ограничения математической модели оптимизационной задачи в таблице исходных данных (см. рис. 51). В столбце «Имя» – название ограничения математической модели оптимизационной задачи в таблице исходных данных (см. рис. 51). Во втором разделе таблицы «Ограничения» приводятся:

- в столбце «Окончательное значение» – значение ограничения при найденном оптимальном значении целевой функции;
- в столбце «Множитель Лагранжа» – изменения целевой функции при изменении правой части ограничения на единицу.

Отчет о пределах (см. табл. 15) состоит из двух разделов: «Целевая функция» и «Переменная».

В обоих разделах таблицы в столбце «Ячейка» приводятся адреса ячеек целевой функции и независимых переменных математической модели оптимизационной задачи в таблице исходных данных (см. рис. 51). В столбце «Имя» – название параметров математической модели оптимизационной задачи в таблице исходных данных (см. рис. 51) и в столбце «Значение» – их величина при решении оптимизационной задачи.

В разделе «Переменная» показывается, в каких пределах могут изменяться переменные, вошедшие в оптимальное решение, при сохранении структуры оптимального решения:

- приводятся значения x_i, для оптимального решения;
- приводятся нижние и верхние пределы изменения значений x_i значения целевой функции при этих предельных значениях переменных.

Пример 8. Необходимо найти оптимальную рецептуру полимерного композита, получаемого методом горячего прессования на основе смеси первичного полиэтилена и вторичного полипропилена, с наполнителем (древесный опил хвойных пород древесины) и лубрикантом (смесь стеариновой кислоты и полиэтиленового воска). Оптимальная рецептура компонентов композита должна обеспечивать ему максимальное значение прочности при изгибе (Y_1, МПа), значение твердости по Бринеллю (Y_2, МПа) не менее 50 МПа и водопоглощение за 30 сут (Y_3, %) не более 15 %. Массовая доля наполнителя в композите должна быть в пределах 40–60 мас. %, а лубриканта – не более 3 мас. %. Такая задача относится к задачам оптимального управления технологическими процессами.

Для данного композита установлены следующие экспериментально-статистические зависимости влияния на свойства этого композита массового содержания в нем компонентов: наполнителя Z_1 и лубриканта Z_2:

$$\begin{aligned}
&Y_1 = 1{,}52Z_1 + 5{,}5Z_2 + 0{,}033Z_1Z_2 - 0{,}02Z_1^2 - 2{,}3Z_2^2;\\
&Y_2 = 3{,}99Z_1 - 6{,}62Z_2 + 0{,}24Z_1Z_2 - 0{,}055Z_1^2 - 0{,}93Z_2^2;\\
&Y_3 = 0{,}029Z_1 + 1{,}42Z_2 - 0{,}05Z_1Z_2 + 0{,}005Z_1^2 + 0{,}11Z_2^2.
\end{aligned}$$

Проведем анализ, формализацию и содержательную постановку задачи. Анализ показывает, что задача относится к задачам условной оптимизации (задачам по заказу), так как имеются два ограничения (значение твердости по Бринеллю и водопоглощение за 30 сут) и два граничных условия (массовая доля наполнителя и лубриканта в композите).

Выберем в качестве целевой функции значение твердости композита по Бринеллю Y_1, а в качестве переменных – содержание в нем наполнителя Z_1 и лубриканта Z_2.

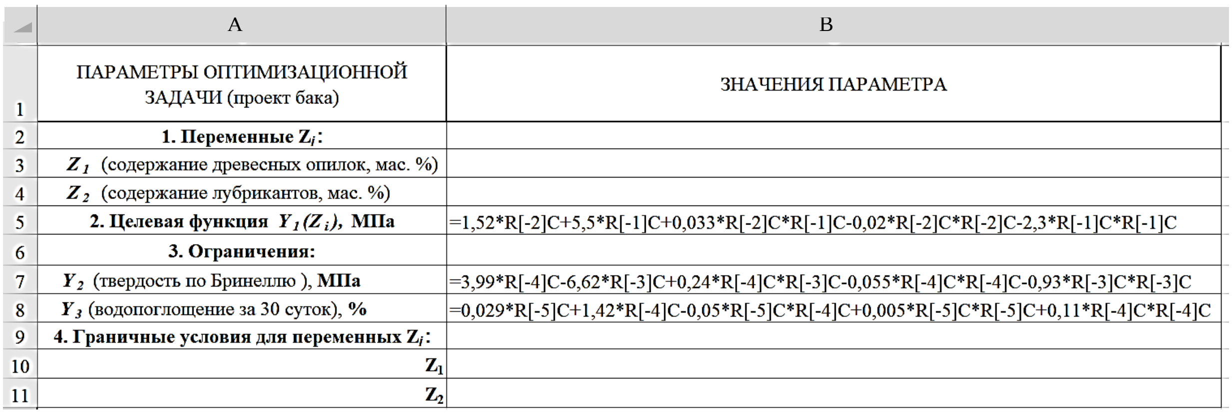
Математическая модель оптимизационной задачи будет иметь следующий вид:

$$\left.\begin{array}{l}
Y_1 = 1{,}52Z_1 + 5{,}5Z_2 + 0{,}033Z_1Z_2 - 0{,}02Z_1^2 - 2{,}3Z_2^2 \to \max(G(Z_i));\\
Y_2 = 3{,}99Z_1 - 6{,}62Z_2 + 0{,}24Z_1Z_2 - 0{,}055Z_1^2 - 0{,}93Z_2^2\left(ОГР\right);\\
Y_3 = 0{,}029Z_1 + 1{,}42Z_2 - 0{,}05Z_1Z_2 + 0{,}005Z_1^2 + 0{,}11Z_2^2\left(ОГР\right);\\
40 \le Z_1 \le 60\left(ГРУ\right);\\
0 \le Z_2 \le 3\left(ГРУ\right).
\end{array}\right\}$$

Анализ данной модели показывает следующее:

- задача однокритериальная – одна целевая функция Y_1;
- задача двухпараметрическая – две детерминированные непрерывные переменные: содержания в композите наполнителя Z_1 и лубриканта Z_2;
- отклик объекта Y_1 – непрерывная величина;
- задача относится к классу задач нелинейного программирования (Y_1, Y_2 и Y_3 – нелинейные зависимости);
- задача имеет множество решений, так как в ограничениях есть неравенства;
- для решения задачи можно применить аналитические, графический и численные методы решения.

Применим численные методы решения оптимизационных задач, запрограммированные в MS Excel. Первоначально сформируем в Excel файл с исходными данными для расчетов с показом формул (рис. 60).

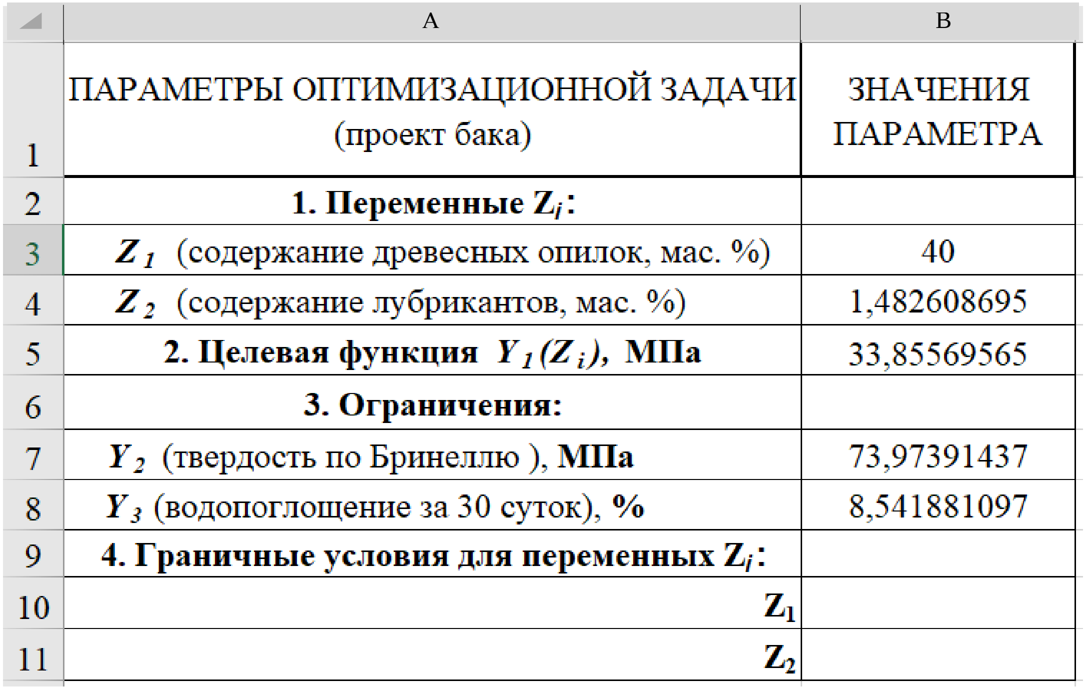


	A	B
1	ПАРАМЕТРЫ ОПТИМИЗАЦИОННОЙ ЗАДАЧИ (проект бака)	ЗНАЧЕНИЯ ПАРАМЕТРА
2	**1. Переменные Z_i:**	
3	Z_1 (содержание древесных опилок, мас. %)	
4	Z_2 (содержание лубрикантов, мас. %)	
5	**2. Целевая функция $Y_1(Z_i)$, МПа**	=1,52*R[-2]C+5,5*R[-1]C+0,033*R[-2]C*R[-1]C-0,02*R[-2]C*R[-2]C-2,3*R[-1]C*R[-1]C
6	**3. Ограничения:**	
7	Y_2 (твердость по Бринеллю), **МПа**	=3,99*R[-4]C-6,62*R[-3]C+0,24*R[-4]C*R[-3]C-0,055*R[-4]C*R[-4]C-0,93*R[-3]C*R[-3]C
8	Y_3 (водопоглощение за 30 суток), **%**	=0,029*R[-5]C+1,42*R[-4]C-0,05*R[-5]C*R[-4]C+0,005*R[-5]C*R[-5]C+0,11*R[-4]C*R[-4]C
9	**4. Граничные условия для переменных Z_i:**	
10	Z_1	
11	Z_2	

Рис. 60. Файл в Excel с исходными данными для расчетов при решении оптимизационной задачи для примера 8 с показом формул

Затем, как для примера 7, выполним все действия в Excel. Полученные результаты поиска глобального оптимума для примера 8 представлены на рис. 61.

Полученные результаты решения оптимизационной задачи для примера 8 были проверены экспериментально. Был получен образец композита с рассчитанными оптимальными значениями содержания в смеси его компонентов древесных опилок (Z_1 = 40 мас. %) и лубриканта (Z_2 = 1,4826 мас. %) и измерены показатели его свойств.



	A	B
1	ПАРАМЕТРЫ ОПТИМИЗАЦИОННОЙ ЗАДАЧИ (проект бака)	ЗНАЧЕНИЯ ПАРАМЕТРА
2	**1. Переменные Z_i:**	
3	Z_1 (содержание древесных опилок, мас. %)	40
4	Z_2 (содержание лубрикантов, мас. %)	1,482608695
5	**2. Целевая функция $Y_1(Z_i)$, МПа**	33,85569565
6	**3. Ограничения:**	
7	Y_2 (твердость по Бринеллю), **МПа**	73,97391437
8	Y_3 (водопоглощение за 30 суток), **%**	8,541881097
9	**4. Граничные условия для переменных Z_i:**	
10	Z_1	
11	Z_2	

Рис. 61. Результат решение оптимизационной задачи
для примера 8

У полученного образца композита с оптимальными значениями содержания в смеси его компонентов фактическая прочность при изгибе составила 36 МПа. По сравнению с расчетным значением (см. рис. 61) это значение было больше на 6,33 %. Значения твердости образца композита по Бринеллю и водопоглощения за 30 сут находились в области соответствующей их граничным условиям.

В программе Statistica решение оптимизационных задач построено не на применении численных методов, а на результатах многостороннего анализа различными методами результатов эксперимента с использованием в меню «Анализ» раздела «Углубленные методы». Программа Statistica позволяет применять метод Тагучи для решения оптимизационных задач в управлении качеством промышленной продукции [10].

В заключение можно отметить, что использование математических методов планирования экспериментов и анализ их результатов с применением компьютерных программ позволяют уменьшить вероятность появления ошибок и ускорить выполнение этих этапов экспериментальной работы.

Для дополнительного анализа полученного решения оптимизационной задачи, аналогичного действиям для примера 7, получаем отчеты «Результаты», «Устойчивость» и «Пределы», которые представлены в табл. 16–18.

Таблица 16

Отчет о результатах для примера 8

Результат: Решение найдено. Все ограничения и условия оптимальности выполнены					
Ячейка целевой функции (Максимум)					
Ячейка	Имя	Исходное значение	Окончательное значение		
R5C2	2. Целевая функция Y1(Zi), МПа ЗНАЧЕНИЯ ПАРАМЕТРА	33,86	33,86		
Ячейки переменных					
Ячейка	Имя	Исходное значение	Окончательное значение	Целочисленное	
R3C2	Z1 (содержание древесных опилок, мас. %) ЗНАЧЕНИЯ ПАРАМЕТРА	40,00	40,00	Продолжить	
R4C2	Z2 (содержание лубрикантов, мас. %) ЗНАЧЕНИЯ ПАРАМЕТРА	1,48	1,48	Продолжить	
Ограничения					
Ячейка	Имя	Значение ячейки	Формула	Состояние	Допуск
R7C2	Y2 (твердость по Бринеллю), МПа ЗНАЧЕНИЯ ПАРАМЕТРА	73,97	R7C2>=50	Без привязки	23,97
R8C2	Y3 (водопоглощение за 30 суток), % ЗНАЧЕНИЯ ПАРАМЕТРА	8,54	R8C2<=15	Без привязки	6,46
R3C2	Z1 (содержание древесных опилок, мас. %) ЗНАЧЕНИЯ ПАРАМЕТРА	40,00	R3C2<=60	Без привязки	20,00
R3C2	Z1 (содержание древесных опилок, мас. %) ЗНАЧЕНИЯ ПАРАМЕТРА	40,00	R3C2>=40	Привязка	0
R4C2	Z2 (содержание лубрикантов, мас. %) ЗНАЧЕНИЯ ПАРАМЕТРА	1,48	R4C2<=3	Без привязки	1,52
R4C2	Z2 (содержание лубрикантов, мас. %) ЗНАЧЕНИЯ ПАРАМЕТРА	1,48	R4C2>=0	Без привязки	1,48

Таблица 17

Отчет об устойчивости для примера 8

Ячейки переменных		Окончательное значение	Приведенный градиент
Ячейка	Имя		
R3C2	Z1 (содержание древесных опилок, мас. %) ЗНАЧЕНИЯ ПАРАМЕТРА	40,00	–0,03
R4C2	Z2 (содержание лубрикантов, мас. %) ЗНАЧЕНИЯ ПАРАМЕТРА	1,48	0
	Ограничения	Окончательное значение	Множитель Лагранжа
Ячейка	Имя		
R7C2	Y2 (твердость по Бринеллю), МПа ЗНАЧЕНИЯ ПАРАМЕТРА	73,97	0
R8C2	Y3 (водопоглощение за 30 сут), % ЗНАЧЕНИЯ ПАРАМЕТРА	8,54	0

Таблица 18

Отчет о пределах для примера 8

Ячейка	Целевая функция	Значение				
R5C2	2. Целевая функция Y1(Zi), МПа ЗНАЧЕНИЯ ПАРАМЕТРА	33,86				
	Переменная		Нижний предел	Целевая функция	Верхний предел	Целевая функция
Ячейка	Имя	Значение		Результат		Результат
R3C2	Z1 (содержание древесных опилок, мас. %) ЗНАЧЕНИЯ ПАРАМЕТРА	40,00	40,00	33,86	55,02	28,88
R4C2	Z2 (содержание лубрикантов, мас. %) ЗНАЧЕНИЯ ПАРАМЕТРА	1,48	0	28,80	3,00	28,56

Пример 9. Для поиска оптимальных технологических параметров получения методом горячего прессования плит толщиной 4 мм из экокомпозитов с полимерной фазой полиэтилена высокой плотности (ПЭНД) и шлифовальной пыли березы был составлен и реализован план эксперимента Плэккетта – Бермана с двукратным повторением опытов со следующими входными факторами Z_i:

Z_1 – содержание шлифовальной пыли березовой фанеры (наполнитель), мас. ч.;

Z_2 – содержание металена F-1108 (компатибилизатор), мас. ч.;

Z_3 – температура горячего прессования, °С;

Z_4 – давление при горячем прессовании, МПа;

Z_5 – давление при холодном прессовании, МПа;

Z_6 – продолжительность горячего прессования, мин;

Z_7 – продолжительность холодного прессования, мин.

Для полученных образцов плит были определены следующие показатели их свойств Y_j:

Y_1 – плотность, кг/м^3;

Y_2 – твердость по Бринеллю, МПа;

Y_3 – пластичность, %;

Y_4 – модуль упругости при сжатии, МПа;

Y_5 – прочность при изгибе, МПа;

Y_6 – ударная вязкость, кДж/м^2;

Y_7 – водопоглощение за сутки, %;

Y_8 – водопоглощение за 30 сут, %;

По результатам регрессионного анализа влияния входных факторов на свойства образцов плит были получены с доверительной вероятностью $P \geq 0{,}90$ следующие уравнения регрессии с коэффициентом детерминации R^2:

$$Y_1,\ \text{кг/м}^3 = 1{,}79Z_1 + 5{,}11Z_3\ (R^2 = 0{,}98);$$
$$Y_2,\ \text{МПа} = 0{,}22Z_1 + 0{,}21Z_3\ (R^2 = 0{,}99);$$
$$Y_3,\ \% = 1{,}68Z_2 + 0{,}21Z_3\ (R^2 = 0{,}99);$$
$$Y_4,\ \text{МПа} = 3{,}13Z_1 + 2{,}51Z_3\ (R^2 = 0{,}99);$$
$$Y_5,\ \text{МПа} = 0{,}18Z_3\ (R^2 = 0{,}99);$$
$$Y_6,\ \text{кДж/м}^2 = 0{,}08Z_3\ (R^2 = 0{,}93);$$
$$Y_7,\ \% = 0{,}02Z_1 - 0{,}06Z_5\ (R^2 = 0{,}96);$$
$$Y_8,\ \% = 0{,}12Z_1\ (R^2 = 0{,}91).$$

Для предполагаемого производства методом горячего прессования по полученным уравнениям регрессии в Excel симплексным методом была найдена оптимальная рецептура (табл. 19) для следующей математической модели оптимизационной задачи:

целевая функция (прочность при изгибе, Y_5) → максимальное значение;

ограничения:

твердость по Бринеллю Y_2 – не менее 50 МПа;

ударная вязкость Y_6 – не менее 10 кДж/м^2;

водопоглощение за 30 сут Y_7 – не более 10 мас. %;

граничные условия:

содержание наполнителя Z_1 – не менее 50 мас. ч.;

содержание компатибилизатора Z_2 – не более 4 мас. ч.;

температура горячего прессования Z_3 – в диапазоне 175–185 °С;

давление при горячем прессовании Z_4 – в диапазоне 15–30 МПа;

давление при холодном прессовании Z_5 – в диапазоне 9,4–12,6 МПа;

продолжительность горячего прессования Z_6 – не менее 10 мин;

продолжительность холодного прессования Z_7 – не менее 5 мин.

В результате расчетов в программе Excel были предложены для экспериментальной проверки следующее оптимальные технологические параметры получения плит методом горячего прессования:

содержание шлифовальной пыли березы – 67,1 мас. ч.;

содержание Металена F-1108 – 0,0 мас. ч.;

температура горячего прессования – 185 °С;

давление при горячем прессовании Z_4 – 15 МПа;

давление при холодном прессовании – 9,9 МПа.

продолжительность горячего прессования Z_6 – 10 мин;

продолжительность холодного прессования Z_7 – 5 мин.

Для большинства физико-механических свойств расчетные значения показателей не отклоняются от фактического результата более 10 %, что свидетельствует о высокой точности разработанной математической модели.

Таблица 19

Значение параметров оптимизационной задачи

Параметры оптимизационной задачи	Значения параметра
1. Переменные Z_i:	
Z1 (содержание шлифовальной пыли березы, мас. %)	67,8
Z2 (содержание металена F-1108, мас. %)	0
Z3 (температура горячего прессования, °С)	185,0
Z4 (давление при горячем прессовании, кгс/м^2)	100,0
Z5 (усилие при холодном прессовании, кН)	60,0
Z6 (продолжительность горячего прессования, мин)	10,0
Z7 (продолжительность холодного прессования, мин)	5,0
2. Целевая функция: ***прочность при изгибе (МПа)***	36,2
3. Ограничения:	
ударная вязкость, кДж/м^2	15,6
твердость по Бринеллю, МПа	55,0
водопоглощение за 24 ч, %	1,0

Окончание табл. 19

Параметры оптимизационной задачи	Значения параметра
4. Граничные условия:	
ударная вязкость, кДж/м^2 $\geq$ 10	
твердость по Бринеллю, МПа $\geq$ 40	
водопоглощение за 24 ч, % < 1,5	
пластичность, % > 0	39,0
модуль упругости при сжатии, МПа > 0	675,7
прочность при изгибе, МПа > 0	36,2
Z1 (содержание шлифовальной пыли березы, мас. %) > 25	
Z1(содержание шлифовальной пыли березы, мас. %) < 67,8	
Z2 (содержание металена F-1108, мас. %) < 1,69	
Z2 (содержание металена F-1108, мас. %) $\geq$ 0	
Z3 (температура горячего прессования, °С) < 185	
Z3 (температура горячего прессования, °С) $\geq$ 175	
Z4 (давление при горячем прессовании, кгс/см^2) $\geq$ 100	
Z4 (давление при горячем прессовании, кгс/см^2) < 200	
Z5 (усилие при холодном прессовании, кН) < 80	
Z5 (усилие при холодном прессовании, кН) $\geq$ 60	
Z6 (продолжительность горячего прессования, мин) < 20	
Z6 (продолжительность горячего прессования, мин) $\geq$ 10	
Z7 (продолжительность холодного прессования, мин) $\geq$ 5	
Z7 (продолжительность холодного прессования, мин) < 10	

В программе Statistica решение оптимизационных задач построено не на применении численных методов, а на результатах многостороннего анализа различными методами результатов эксперимента с использованием в меню «Анализ» раздела «Углубленные методы». Программа Statistica позволяет применять метод Тагучи для решения оптимизационных задач в управлении качеством промышленной продукции [10].

ЗАКЛЮЧЕНИЕ

Целью многих прикладных научных исследований является определение закономерностей предполагаемого влияния изменяемых в эксперименте исследователем факторов (входных факторов) на параметры (свойства) исследуемого объекта. Для этого требуется проведение грамотного расчета погрешностей проводимых измерений, представляющих собой сумму абсолютных значений возможных предельных случайных и систематических ошибок измерений. При этом величина случайных ошибок измерений, зависящая от влияния на свойства объекта нерегулируемых исследователем в эксперименте факторов (случайных факторов), может быть рассчитана при наличия повторных опытов эксперимента. Авторитетные компьютерные программы, к которым относятся MS Excel и Statistica, позволяют с использованием современных математических методов рассчитывать необходимые параметры для оценки случайных ошибок результатов эксперимента.

Планы экспериментов в большинстве случаев являются авторским продуктом интеллектуальной работы исследователей. По современным требованиям к качеству анализа результатов экспериментальных исследований необходимо исключение из них грубых ошибок измерений и, кроме визуального анализа, проведение статистического анализа экспериментальных данных для заданного значения доверительной вероятности. Для статистического анализа экспериментальных данных рекомендуется в плане эксперимента предусматривать повторение его опытов (желательно каждого опыта) для оценки погрешностей проводимых измерений.

Математическое планирование экспериментов является современной тенденцией в проведении научных исследований. Существует очень большое число математических планов эксперимента, при составлении которых используются разнообразные методы математической комбинаторики. Грамотно составленные математические планы экспериментов позволяют получать количественную информацию о влиянии изменяемых исследователем в эксперименте входных факторов на параметры (отклики) объекта. При математическом планировании эксперимента достигаются минимальные ошибки и наибольшая достоверность экспериментально-статистических зависимостей количественных параметров исследованного объекта от величины входных факторов и найденных решений оптимальных задач. В программе Statistica предусмотрено компьютерное составление большого числа популярных математических планов эксперимента. Математическое планирование экспериментов и статистический анализ их результатов с применением авторитетных компьютерных программ позволяют ускорить выполнение научно-исследовательских работ и сократить затраты на их выполнение.

Библиографический список

1. Глухих, В. В. Прикладные научные исследования : учебник / В. В. Глухих. – Екатеринбург : Урал. гос. лесотехн. ун-т, 2016. – 240 с.
2. Семенов, С. А. Планирование и обработка результатов эксперимента : учебное пособие / С. А. Семенов. – 2-е изд., пер. и доп. – Москва : РТУ МИРЭА, 2021. – 48 с. – Текст : электронный // Лань : электронно-библиотечная система. – URL: https://e.lanbook.com/book/176518 (дата обращения: 08.01.2023).
3. Пен, Р. З. Планирование эксперимента в Statgraphics : учебное пособие по дисциплинам «Планирование и организация эксперимента» и «Основы научных исследований». – 2-е изд., доп. / Р. З. Пен. – Красноярск : Красноярский писатель ; СибГТУ, 2012. – 270 с.
4. Пен, Р. З. Статистические методы математического моделирования, анализа и оптимизации технологических процессов / Р. З. Пен, В. Р. Пен. – 3-е изд., стер. – Санкт-Петербург : Лань, 2022. – 308 с. – ISBN 978-5-507-45300-9. – Текст : электронный // Лань : электронно-библиотечная система. – URL: https://e.lanbook.com/book/264239 (дата обращения: 08.01.2023).
5. Ахназарова, С. Л. Методы оптимизации эксперимента в химической технологии / С. Л. Ахназарова, В. В. Кафаров. – Москва : Высшая школа, 1985. – 327 с.
6. Курицкий, Б. Я. Поиск оптимальных решений средствами Excel 7.0 / Б. Я. Курицкий. – Санкт-Петербург, 1997. – 384 с.
7. Дэниел, К. Применение статистики в промышленном эксперименте / К. Дэниел. – Москва : Мир, 1979. – 300 с.
8. Боровиков, В. П. Популярное введение в современный анализ данных в системе STATISTICA. Методология и технология современного анализа данных : учебное пособие / В. П. Боровиков. – Москва : Горячая линия-Телеком, 2018. – 288 с. – ISBN 978-5-9912-0326-5. – Текст: электронный // Лань : электронно-библиотечная система. – URL: https://e.lanbook.com/book/111023 (дата обращения: 08.01.2023).
9. Романенко, В. Н. Книга для начинающего исследователя-химика / В. Н. Романенко, А. Г. Орлов, Г. В. Никитина. – Ленинград : Химия, 1987. – 280 с.
10. Халафян, А. А. STATISTICA 6. Статистический анализ данных : учебник. – 3-е изд. / А. А. Халафян. – Москва : Бином-Пресс, 2008. – 512 с.

11. Вадзинский, Р. Статистические вычисления в среде Excel. Библиотека пользователя / Р. Вадзинский. – Санкт-Петербург : Питер, 2008. – 608 с.

12. Леоненков, А. В. Решение задач оптимизации в среде MS Excel / А. В. Леоненков. – Санкт-Петербург : БХВ-Петербург, 2005. – 704 с.

13. Применение MS Excel и Statistica for Windows для лесотаксационных вычислений и обработки экспериментальных данных методами математической статистики : учебное пособие / Л. В. Стоноженко, А. Н. Югов, В. Н. Карминов, Н. Г. Иванов. – Москва : МГТУ им. Н. Э. Баумана, 2012. – 88 с. – Текст : электронный // Лань : электронно-библиотечная система. – URL: https://e.lanbook.com/book/104645 (дата обращения: 08.01.2023).

14. Бочкарев, В. В. Оптимизация химико-технологических процессов : учебное пособие / В. В. Бочкарев. – Томск : Томский политехнический университет, 2014. – 264 с. – Текст : электронный. – URL: https://e.lanbook.com/books/element.php?pl1_id=62913

15. Пошарников, П. Ф. Моделирование и оптимизация процессов в лесном комплексе : учебное пособие / П. Ф. Пошарников. – Воронеж : ВГЛТУ, 2014. – 271 с. – Текст : электронный. – URL: http://e.lanbook.com/books/element.php?pl1_id=64147 (дата обращения: 08.01.2023).

16. Мухутдинов, А. Р. Основы моделирования и оптимизации материалов и процессов в Microsoft Excel : учебное пособие / А. Р. Мухутдинов, З. Р. Вахидова, М. Р. Файзуллина. – Казань : Изд-во КНИТУ, 2017. – 172 с. – Текст : электронный. – URL: http://biblioclub.ru/index.php?page=book&id=560915 (дата обращения: 08.01.2023).

17. Глухих, В. В. Прикладные научные исследования : учебник / В. В. Глухих. – Екатеринбург : Урал. гос. лесотехн. ун-т, 2016. – 240 с.

Учебное издание

Глухих *Виктор Владимирович*
Шкуро *Алексей Евгеньевич*
Артемов *Артем Вячеславович*
Шишлов *Олег Федорович*
Кривоногов *Павел Сергеевич*

МАТЕМАТИЧЕСКОЕ ПЛАНИРОВАНИЕ ЭКСПЕРИМЕНТОВ И АНАЛИЗ ИХ РЕЗУЛЬТАТОВ С ПРИМЕНЕНИЕМ КОМПЬЮТЕРНЫХ ПРОГРАММ

ISBN 978-5-94984-864-7
9 785949 848647

Редактор Е. Л. Михайлова
Оператор компьютерной верстки Т. В. Упорова

Подписано к публикации 27.04.2023. Дата размещения на сайте 04.05.2023.
Уч.-изд. л. 4,7. Объем 9 Мб.
Тираж 300 экз. (1-й завод 10 экз.)
Заказ № 7627

ФГБОУ ВО «Уральский государственный лесотехнический университет».
620100, г. Екатеринбург, Сибирский тракт, 37.
Редакционно-издательский отдел. Тел.: 8 (343) 221-21-44.

Типография ООО «ИЗДАТЕЛЬСТВО УЧЕБНО-МЕТОДИЧЕСКИЙ ЦЕНТР УПИ».
620062, РФ, Свердловская область, Екатеринбург, ул. Гагарина, 35а, оф. 2.
Тел.: 8 (343) 362-91-16.