МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА
И ПРОДОВОЛЬСТВИЯ РЕСПУБЛИКИ БЕЛАРУСЬ

Учреждение образования
«БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ
ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

Кафедра основ научных
исследований и проектирования

А. Н. Леонов, М. М. Дечко, В. Б. Ловкис

ОСНОВЫ НАУЧНЫХ ИССЛЕДОВАНИЙ И МОДЕЛИРОВАНИЯ

Рекомендовано Учебно-методическим объединением высших учебных заведений Республики Беларусь по образованию в области сельского хозяйства в качестве учебно-методического комплекса для студентов высших учебных заведений, обучающихся по группе специальностей 74 06 Агроинженерия

Минск
БГАТУ
2010

УДК 001.89:[519.2+519.22](07)
ББК 72.4я7
Л47

Рекомендовано научно-методическим советом агромеханического факультета БГАТУ
Протокол № 6 от 21 февраля 2009 г.

Авторы:
доктор технических наук, профессор *А.Н. Леонов*,
кандидат технических наук *М.М. Дечко,*
кандидат технических наук, доцент *В.Б. Ловкис*

Рецензенты:
заведующий кафедрой теории вероятностей и математической статистики БГУ, доктор физико-математических наук, профессор *Н.Н. Труш,*
заведующий кафедрой «Машины и технология обработки металлов давлением» БНТУ, доктор технических наук, профессор *К.Е. Белявин*

Леонов, А.Н.
Л47 Основы научных исследований и моделирования : учебно-методический комплекс / А.Н. Леонов, М.М. Дечко, В.Б. Ловкис. – Минск : БГАТУ, 2010. – 276 с.

ISBN 978-985-519-191-0.

В учебно-методическом комплексе «Основы научных исследований и моделирования» изложены основы математической теории эксперимента, включающей выбор математической модели для описания объектов исследований, построение плана эксперимента, обработку полученных экспериментальных данных методами математической статистики, принятие решения на базе построенной математической модели.

Предназначено для студентов, магистрантов, аспирантов, преподавателей вузов инженерного профиля, а также для научных работников.

УДК 001.89:[519.2+519.22](07)
ББК 72.4я7

ISBN 978-985-519-191-0

СОДЕРЖАНИЕ

ВВЕДЕНИЕ

Развитие науки, сопровождающееся накоплением эмпирических данных, приводит к созданию обобщающих теорий, которые со временем генерируют новые методологические принципы познания.

XIX век явился триумфом науки, основанной на детерминированном подходе. Тем не менее, уже тогда были заложены теоретические основы концептуально нового подхода, основанного на теории вероятностей и математической статистике. От изучения хорошо организованных систем наука перешла к изучению плохо организованных систем. Применение вероятностно-статистических методов является необходимым условием исследования объектов, природа которых описывается стохастическими закономерностями. Концептуальная особенность новой методологии заключается в возможности изучения объектов, характеризующихся большим количеством случайных состояний, путем создания и анализа случайных ситуаций.

Конечная цель любой научной теории – обобщение и сжатие эмпирической информации путем построения математических моделей, позволяющих раскрыть закономерности исследуемого объекта и управлять его поведением. Для изучения плохо организованных систем в настоящее время создана математическая теория эксперимента, представляющая собой новый методологический подход, направленный на моделирование исследуемых объектов методами математической статистики при неполном знании природы объекта. Теория позволяет унифицировать методы экспериментальных исследований, уменьшить субъективизм при интерпретации экспериментальных данных, а также дает возможность объективно сопоставлять результаты различных исследователей. Математическое моделирование объектов, природа которых носит стохастический характер, позволяет количественно оценивать влияние существенных факторов на значимые параметры объектов, прогнозировать и управлять их поведением, решать оптимизационные задачи.

Решение прикладных задач методами математической теории эксперимента включает в себя четыре этапа: выдвижение одной или нескольких конкурирующих математических моделей для адекватного описания исследуемого объекта, разработку эффективного плана эксперимента для оценки параметров выбранной модели, эффективную обработку результатов эксперимента методами математической статистики и принятие решения. Умелое сочетание

этих этапов позволяет эффективно использовать имеющиеся теоретико-эмпирические ресурсы, направленные на получение адекватной модели, способной предсказывать результат с требуемой точностью при минимально возможных временных и материальных затратах.

Предлагаемый учебно-методический комплекс (УМК) содержит основные понятия и положения раздела «Основы научных исследований» учебной дисциплины «Основы научных исследований и моделирования», основная цель которой – дать студентам знания по основным методологическим принципам изучения плохо организованных систем методами математической статистики; изложить теорию и алгоритмы, необходимые для исследования объектов, которым присущи стохастические закономерности; научить решать прикладные задачи, в том числе, в области аграрно-технического сектора народного хозяйства.

УМК задуман, прежде всего, как учебник для решения прикладных задач, отличительная особенность которого заключается в том, что изложенная стратегия научных исследований направлена на нахождение оптимальных условий функционирования изучаемого объекта. Прикладной характер предопределил особенность изложения материала: некоторые понятия и утверждения приводятся в основном тексте без строгих определений и доказательств, некоторые из которых для любознательных читателей вынесены в Приложение 6.

Учебно-методический комплекс состоит из 11 тем, каждая из которых содержит теорию, алгоритмы решения прикладных задач, подробное численное решение базовой задачи, многовариантные тесты, вопросы для самопроверки и задачи трех уровней сложности с ответами для самостоятельного решения.

В основу учебного издания положен курс лекций, подготовленный авторами в БГАТУ по учебной дисциплине «Основы научных исследований и моделирования», являющийся логическим развитием идей и методов, заложенных в 1990–2005 годы профессорами кафедры Н.И. Боханом, И.С. Нагорским, А.М. Дмитриевым по применению методов математической статистики для решения прикладных задач в области агропромышленного комплекса.

Авторы выражают благодарность сотрудникам кафедры теории вероятностей и математической статистики БГУ кандидатам физико-математических наук, доцентам П.М. Лаппо и А.Ю. Харину за внимательное прочтение рукописи и сделанные замечания и пожелания.

Модуль 0
ИСТОРИЧЕСКИЙ АСПЕКТ РАЗВИТИЯ НАУКИ. ОРГАНИЗАЦИОННАЯ СТРУКТУРА НАУКИ В РЕСПУБЛИКЕ БЕЛАРУСЬ И НИРС В БГАТУ. ЭТАПЫ НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЫ (тема 1)

В результате изучения модуля студент должен:

– **знать** приоритетные направления развития науки в Республике Беларусь, в том числе в агропромышленном комплексе (АПК);

– **знать** цели и задачи изучаемой дисциплины и ее структуру;

– **уметь** оформлять результаты научных исследований;

– **иметь** представление об историческом аспекте развития и организационной структуре науки в Республике Беларусь и НИРС в БГАТУ;

– **приобрести** логические и этические навыки ведения научной дискуссии.

0.1. Цель, задачи и структура дисциплины

Цель преподавания дисциплины

Цель преподавания дисциплины «Основы научных исследований и моделирования» – привить теоретические знания и практические навыки, подготовить студентов к самостоятельной научно-исследовательской работе, научить моделированию объектов агропромышленного комплекса и их оптимизации.

Задачи изучения дисциплины

– ознакомить студентов с историческим аспектом развития мировой науки, с организационной структурой науки в РБ и НИРС в БГАТУ, а также с приоритетными научными направлениями, в том числе в агропромышленном комплексе РБ;

– дать студентам знания по основам статистических методов математической теории эксперимента.

В результате изучения дисциплины студенты должны **знать**:

– структуру и очередность выполнении этапов НИР, НИОКР, НИОТР, общие требования и правила оформления НИР;

– методы проведения теоретических и экспериментальных исследований;

– методы оптимизации исследуемых объектов, в том числе в области агропромышленного комплекса.

С учетом полученных знаний и навыков самостоятельной работы при изучении дисциплины студенты должны **уметь**:

– выбирать модель для описания изучаемого объекта;

– составлять план эксперимента для определения параметров выбранной модели;

– разрабатывать методику проведения эксперимента и выбирать необходимое научное и технологическое оборудование;

– проводить эксперимент по выбранному плану;

– рассчитывать параметры модели методами математической статистики (регрессионный анализ);

– проверять полученные математические модели на статистическую значимость и адекватность методами математической статистики (дисперсионный анализ);

– находить оптимальные условия функционирования изучаемого объекта и оценивать точность прогнозирования.

При структурировании содержания раздела «Основы научных исследований» (ОНИ) учебной дисциплины «Основы научных исследований и моделирования» учебный материал разделен на модули (таблица 1).

Таблица 1 – Тематический план раздела ОНИ

Номер и название модуля	Нагрузка на модули	Лекции	Лабораторные	Практические	УСРС
М-0. Исторический аспект развития науки. Организационная структура науки в Республике Беларусь и НИРС в БГАТУ. Этапы научно-исследовательской работы	6	4	–	2	–
М-1. Предварительная обработка экспериментальных данных методами математической статистики	18	8	4	4	2
М-2. Однофакторный эксперимент	18	8	4	4	2
М-3. Многофакторный эксперимент	22	8	6	6	2
М-R. Итоговое занятие	4	2	–	–	2
Всего	68	30	14	16	8

Связь дисциплины «Основы научных исследований и моделирования» с другими учебными дисциплинами

Дисциплина является основополагающей для построения математических моделей, планирования и проведения экспериментов, анализа экспериментальных данных и оптимизации исследуемого объекта. Изучение дисциплины предполагает наличие знаний по некоторым разделам высшей математики, а именно – линейной алгебре, дифференциальному и интегральному исчислению, теории вероятностей и математической статистики.

Учебно-методический комплекс подготовлен на основе лекционного курса для студентов БГАТУ и может быть рекомендован студентам, магистрантам, аспирантам, преподавателям вузов инженерного профиля, а также научным работникам.

0.2. Перечень и содержание лекций

Лекция 1. Важнейшие научно-технические открытия, оказавшие влияние на развитие цивилизации. Глобальные научные проблемы, стоящие перед человечеством в XXI веке. Исторический аспект развития и организационная структура науки в Республике Беларусь. Государственные и общественные научные организации. Основные цели и задачи ГКНТ, НАН Беларуси, ВАК. Отделение аграрных наук в НАН Беларуси. Республиканский фонд поддержки производителей сельскохозяйственной продукции, продовольствия и аграрной науки. Наука, технологии, инновации. Фундаментальные и прикладные науки. Теоретические и экспериментальные методы исследований. Мотивация и психология творчества.

Лекция 2. Образование, популяризация и пропаганда науки. Классификация научных исследований: НИР, НИОКР, НИОТР. Организационные формы НИРС в БГАТУ. Бакалавриат, магистратура, аспирантура, докторантура, соискательство. Вузовские, республиканские и международные студенческие конкурсы и конференции. Выбор и технико-экономическое обоснование научной темы. Этапы научно-исследовательской работы. Поиск, анализ и систематизация априорной научной информации. Приоритетные направления научных исследований в Республике Беларусь, в том числе в агропромышленном комплексе.

0.3. Материалы для самостоятельной работы и практических занятий

Вопросы для самопроверки

1. Назовите важнейшие, на ваш взгляд, научно-технические открытия, повлиявшие на развитие цивилизации.
2. Назовите имена выдающихся белорусских ученых-просветителей.
3. Опишите организационную структуру науки в Республике Беларусь.
4. Назовите основные цели и задачи ГКНТ Республики Беларусь.
5. Назовите основные цели и задачи НАН Беларуси.
6. Назовите основные цели и задачи ВАК Республики Беларусь.
7. Назовите основные цели и задачи отделения аграрных наук в НАН Беларуси.
8. Назовите основные цели и задачи республиканского фонда поддержки производителей сельскохозяйственной продукции, продовольствия и аграрной науки.
9. Опишите организационную структуру НИРС в БГАТУ.
10. Перечислите и прокомментируйте глобальные научные проблемы, стоящие перед человечеством в XXI веке.
11. Назовите приоритетные научные проблемы на ближайшие 15 лет в АПК Республики Беларусь.
12. Охарактеризуйте такие понятия, как образование, популяризация и пропаганда науки.
13. Расскажите о системе высшего образования в Республике Беларусь. Что, на ваш взгляд, может повысить качество подготовки дипломированных специалистов?
14. Охарактеризуйте структуру, содержание и очередность выполнения этапов НИР, НИОКР, НИОТР.
15. Перечислите основные разделы отчета о НИР.
16. Какие требования включаются в ТЗ на НИОКР и НИОТР?
17. Изложите алгоритм формулирования темы курсового и дипломного проектов.
18. Сформулируйте алгоритм написания реферата.
19. Сформулируйте алгоритм написания введения.
20. Сформулируйте алгоритм написания заключения.
21. Сформулируйте несколько правил и приемов, повышающих эффективность ведения научной дискуссии.

Модуль 1
ПРЕДВАРИТЕЛЬНАЯ ОБРАБОТКА ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ МЕТОДАМИ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

В результате изучения модуля студент должен:

– **усвоить** элементы теории вероятностей и математической статистики;

– **уметь** производить предварительную обработку экспериментальных данных при решении прикладных задач, в том числе в области АПК.

1.1. Перечень и содержание лекций

Лекция 1. Детерминированные и стохастические закономерности. Теория вероятностей. Математическая статистика. Предельные абсолютная и относительная погрешности в детерминированном эксперименте. Метод расчета предельных абсолютной и относительной погрешностей в детерминированном эксперименте. Алгоритм решения прикладных задач. Решение базовой задачи.

Лекция 2. Функция распределения плотности вероятностей. Генеральная совокупность. Закон нормального распределения и его параметры: генеральное среднее μ, генеральная дисперсия σ^2, генеральное стандартное отклонение σ. Выборка. Выборочное среднее $\overline{Y}$, выборочная дисперсия S^2, выборочное стандартное отклонение S. Число степеней свободы. Проверка случайных величин выборки на промахи. Критерий Смирнова–Граббса. Проверка соответствия случайных величин выборки нормальному закону распределения.

Лекция 3. Определение доверительного интервала для генерального среднего μ. Критерий Стьюдента. Определение доверительного интервала для генеральной дисперсии σ^2. Критерий Пирсона. Минимальный объем выборки, обеспечивающий оценку генерального среднего μ с допустимой погрешностью. Алгоритм решения прикладных задач (одна выборка). Решение базовой задачи.

Лекция 4. Оценка принадлежности двух выборок с $S_1^2 \neq S_2^2$ к одной $\sigma_1^2 = \sigma_2^2$ или к разным $\sigma_1^2 \neq \sigma_2^2$ генеральным совокупностям. Критерий Фишера. Оценка принадлежности двух выборок с $\overline{Y}_1 \neq \overline{Y}_2$ к одной $\mu_1 = \mu_2$ или к разным $\mu_1 \neq \mu_2$ генеральным совокупностям. Критерий Стьюдента. Проверка трех и более выборочных

дисперсий на однородность. Критерий Кохрена. Алгоритм решения прикладных задач (две выборки и более). Решение базовой задачи.

1.2. Термины и условные обозначения

Детерминированные закономерности исследуемого объекта – закономерности, характеризующиеся тем, что заданной совокупности факторов всегда соответствует только одно значение параметра.

Стохастические закономерности исследуемого объекта – закономерности, характеризующиеся тем, что заданной совокупности факторов соответствуют различные значения параметра, причем каждое значение реализуется с некоторой вероятностью.

Вероятность – мера достоверности случайного состояния исследуемого объекта.

Математическая статистика – наука о математических методах анализа эмпирических данных, отражающих поведение исследуемых объектов, подчиняющихся стохастическим закономерностям.

Фактор – детерминированная величина, влияющая на результат эксперимента.

Параметр – наблюдаемая величина, характеризующая исследуемый объект и зависящая от факторов.

Генеральная совокупность – полная совокупность значений параметра исследуемого объекта при всех возможных комбинациях факторов.

Выборка – совокупность относительно небольшого количества значений параметра исследуемого объекта, входящих в генеральную совокупность.

САО – среднее абсолютное отклонение.

Детерминированный эксперимент:

Y – исследуемый параметр.

dY – дифференциал параметра Y.

ΔY – предельная абсолютная погрешность параметра Y.

$\Delta Y_{пр}$ – предельная абсолютная погрешность параметра Y, обусловленная погрешностью измерительных приборов.

$\Delta Y_{ст}$ – предельная абсолютная погрешность параметра Y, обусловленная стохастической природой исследуемого объекта.

$\Delta Y/Y$ – предельная относительная погрешность параметра Y.

ξ, χ, ζ – истинные значения детерминированных факторов.

x, z, w – экспериментальные значения детерминированных факторов.

Δx, Δz, Δw – предельные абсолютные погрешности факторов x, z, w.

Одна выборка:

μ – генеральное среднее.

σ^2 – генеральная дисперсия.

σ – генеральное стандартное отклонение.

$N(Y, \mu, \sigma)$ – нормальный закон распределения случайных значений параметра Y, характеризуемый генеральным средним μ и генеральным стандартным отклонением σ.

$Y_1, \ldots, Y_i, \ldots, Y_n$ – выборка случайных значений параметра Y.

n – объем выборки.

i – текущий номер случайной величины в выборке, $i = 1, \ldots, n$.

$\overline{Y}$ – выборочное среднее параметра Y.

S^2 – выборочная дисперсия параметра Y.

S – выборочное стандартное отклонение параметра Y.

f – число степеней свободы рассчитываемого параметра.

$Y_{\max}$, $Y_{\min}$ – максимальное (минимальное) значение параметра Y.

$\tau_э$ – экспериментальное значение критерия Смирнова–Граббса.

$\tau_{f,\,p}$ – критическое значение критерия Смирнова–Граббса с числом степеней свободы f при доверительной вероятности p (приложение 1).

$t_э$ – экспериментальное значение критерия Стьюдента.

$t_{f,\,p}$ – критическое значение критерия Стьюдента с числом степеней свободы f при доверительной вероятности p (приложение 2).

$\Delta\overline{Y}$ – доверительный интервал генерального среднего μ, рассчитанный по выборке объемом n и доверительной вероятности p.

$\chi^2_{f,(1-p)/2}$ – нижнее критическое значение критерия Пирсона с числом степеней свободы f при доверительной вероятности p (приложение 3).

$\chi^2_{f,(1+p)/2}$ – верхнее критическое значение критерия Пирсона с числом степеней свободы f при доверительной вероятности p (приложение 3).

$n_{\min}$ – минимальное число опытов для оценки генерального среднего μ с допустимой погрешностью $\Delta\overline{Y}_о$ при доверительной вероятности p (приложение 2).

Две выборки и более:

n_1 – объем первой выборки $Y_{11}, \ldots, Y_{1n_1}$, $i = 1, \ldots, n_1$.

n_2 – объем второй выборки $Y_{21}, \ldots, Y_{2n_2}$, $i = 1, \ldots, n_2$.

μ_1 – генеральное среднее для первой выборки.

σ_1^2 – генеральная дисперсия для первой выборки.

σ_1 – генеральное стандартное отклонение для первой выборки.

μ_2 – генеральное среднее для второй выборки.

σ_2^2 – генеральная дисперсия для второй выборки.

σ_2 – генеральное стандартное отклонение для второй выборки.

$\overline{Y}_1$ – выборочное среднее для первой выборки.

S_1^2 – выборочная дисперсия для первой выборки.

S_1 – выборочное стандартное отклонение для первой выборки.

$\overline{Y}_2$ – выборочное среднее значение для второй выборки.

S_2^2 – выборочная дисперсия для второй выборки.

S_2 – выборочное стандартное отклонение для второй выборки.

$\max(S_1^2, S_2^2)$ – максимальная из двух выборочных дисперсий.

$\min(S_1^2, S_2^2)$ – минимальная из двух выборочных дисперсий.

$F_э$ – экспериментальное значение критерия Фишера.

$F_{f_1,\,f_2,\,p}$ – критическое значение критерия Фишера для сравнения отношения двух дисперсий S_1^2 и S_2^2, в числителе которого стоит бóльшая дисперсия с числом степеней свободы f_1, а в знаменателе - меньшая дисперсия с числом степеней свободы f_2 при доверительной вероятности p (приложение 4).

N – число опытов (число выборок).

j – текущий номер опыта, $j = 1, \ldots, N$.

n – число дублей каждого из N опытов.

i – текущий номер дубля, $i = 1, \ldots, n$.

Y_{ji} – значение параметра Y в j-ом опыте и i-ом дубле.

$\overline{Y}_j$ – выборочное среднее в j-ом опыте.

S_j^2 – выборочная дисперсия в j-ом опыте.

S_j – выборочное стандартное отклонение в j-ом опыте.

$\max S_j^2$ – максимальная выборочная дисперсия в N выборках.

$G_э$ – экспериментальное значение критерия Кохрена.

$G_{f_1,\,f_2,\,p}$ – критическое значение критерия Кохрена для сравнения N дисперсий; в числителе которого стоит наибольшая дисперсия с числом степеней свободы f_1, а в знаменателе сумма всех дисперсий с числом степеней свободы $f_2 = N$ при доверительной вероятности p (приложение 5).

1.3. Обработка данных детерминированного эксперимента (тема 2)

Научное описание исследуемых объектов исторически началось с детерминированного подхода, который основан на следующем постулате: если выполнен некоторый комплекс условий, то вытекающее из них единственно возможное состояние объекта всегда наступает со 100 % вероятностью. Примером детерминированного описания исследуемых объектов является классическая механика, расцвет которой пришелся на XIX век.

По мере накопления знаний о поведении исследуемых объектов были выявлены новые закономерности, называемые стохастическими. Если объект исследования подчиняется стохастическим закономерностям, то даже при строгом соблюдении всех условий вытекающее из них состояние принципиально неоднозначно и может быть предсказано лишь с определенной вероятностью методами математической статистики. Математическая статистика позволяет описывать исследуемый объект в условиях неопределенности, причем, и это необходимо подчеркнуть, изучение объектов, подчиняющихся стохастическим закономерностям, организуется путем создания и анализа случайных экспериментальных ситуаций. В настоящее время статистические методы исследований широко используются в физике, биологии, социологии, промышленном и сельскохозяйственном производстве.

В научной литературе можно встретить понятия «хорошо организованные» (детерминированные) и «плохо организованные» (стохастические) объекты. Особенностью научного познания XX века является переход от изучения «хорошо организованных» объектов к изучению «плохо организованных». Несмотря на концептуальное различие детерминированного и стохастического подходов в описании исследуемых объектов, общим у них является наличие некоторой организации, постоянной в пространстве и времени. Можно предположить, что развитие науки приведет к созданию методов моделирования объектов, внутренняя структура которых со временем усложняется в результате эволюционной самоорганизации, а в дальнейшем и к моделированию «концептуально неорганизованных» систем.

Исследование объектов, описываемых детерминированными закономерностями, включает в себя расчет параметров по известным аналитическим формулам, в которых рассчитываемый параметр

зависит от факторов, определяемых экспериментально. В этом случае детерминированный параметр рассчитывается с некоторой предельной абсолютной погрешностью, обусловленной несовершенством средств измерений. Однако для объекта, подчиняющегося детерминированным закономерностям, использование более точных приборов позволит уменьшить предельную абсолютную погрешность. Например, определение линейных размеров объекта может быть выполнено с различной предельной абсолютной погрешностью: 10 мм (швейный метр), 1 мм (металлическая линейка), 0.1 мм (штангенциркуль), 0.01 мм (микрометр), 0.001 мм (инструментальный оптический микроскоп).

В отличие от объектов, подчиняющихся детерминированным закономерностям, предельная абсолютная погрешность ΔY параметра Y в объектах, подчиняющихся стохастическим закономерностям, зависит как от погрешности измерительного прибора $\Delta Y_{\text{пр}}$, так и от стохастической природы исследуемого объекта $\Delta Y_{\text{ст}}$. Математически это утверждение записывается следующим образом:

$$\Delta Y^2 = \Delta Y_{\text{пр}}^2 + \Delta Y_{\text{ст}}^2 . \qquad (1)$$

Если $\Delta Y_{\text{пр}} >> \Delta Y_{\text{ст}}$, то $\Delta Y \approx \Delta Y_{\text{пр}}$. В этом случае стохастической природой исследуемого объекта можно пренебречь, и он с достаточной точностью описывается только детерминированными закономерностями.

Если $\Delta Y_{\text{ст}} >> \Delta Y_{\text{пр}}$, то $\Delta Y \approx \Delta Y_{\text{ст}}$. В этом случае погрешностью измерительных приборов можно пренебречь, и исследуемый объект с достаточной точностью описывается только стохастическими закономерностями.

1.3.1. Теория

Пусть имеется некоторая аналитическая функция $Y = Y(x, z)$ от 2-х детерминированных факторов x и z. Параметр Y рассчитывается с помощью указанной функции по результатам измерений факторов x и z, которые определяются с предельной абсолютной погрешностью Δx и Δz. Изложим метод определения предельной относительной $\Delta Y/Y$ и предельной абсолютной погрешности ΔY параметра Y.

Предельная абсолютная погрешность измерения каждого фактора однозначно определяется погрешностью прибора (см. уравнение (1)), значение которой приводится в паспорте прибора. Поясним понятие **«предельная абсолютная погрешность»**. Дело в том, что истинное значение измеряемого фактора ξ неизвестно, а прибор позволяет измерить лишь его приближенное значение x (истинные значения впредь будем обозначать греческими буквами, а их экспериментальные значения – латинскими). Разность между истинным и измеренным значением $\xi - x$ составляет абсолютную погрешность определения фактора ξ. Но поскольку ни величина, ни знак абсолютной погрешности неизвестны, то для большей надежности за предельную абсолютную погрешность определения фактора ξ принимают ее самое большое значение – погрешность прибора $\delta \geq 0$. На самом деле абсолютная погрешность определения фактора ξ меньше погрешности прибора $|\xi - x| \leq \delta$, но так как эта величина принципиально непредсказуема, то за величину предельной абсолютной погрешности определения фактора принимают значение, равное погрешности прибора.

Прежде чем вывести несколько формул для определения предельной абсолютной и относительной погрешностей детерминированного параметра Y, напомним, что если $Y = Y(x, z)$, то дифференциал функции двух переменных рассчитывается по уравнению

$$dY(x,z) = Y_x'(x,z)dx + Y_z'(x,z)dz\,, \tag{2}$$

где Y_x' (Y_z') – частная производная параметра Y по фактору x (z).

Пусть параметр Y является суммой двух функций $f(x)$ и $g(z)$, причем факторы x и z определяются с предельными абсолютными погрешностями Δx и Δz. Для оценки абсолютной погрешности параметра Y рассчитаем дифференциал dY, используя уравнение (2):

$$Y(x,z) = f(x) \pm g(z)\,, \tag{3}$$

$$dY = f_x'(x)dx \pm g_z'(z)dz\,. \tag{4}$$

Заменив дифференциалы соответствующими приращениями $dY \approx \Delta Y$, $dx \approx \Delta x$, $dz \approx \Delta z$, получим уравнение для определения предельной абсолютной погрешности параметра Y:

$$\Delta Y = \left|f_x'(x)\right|\Delta x + \left|g_z'(z)\right|\Delta z\,. \tag{5}$$

Обратите внимание на то, что знак «±» в уравнении (4) заменен знаком «+» в уравнении (5), а значения производных взяты по модулю. Это объясняется тем, что оценивается максимально возможная предельная абсолютная погрешность.

Пусть параметр Y является произведением двух функций $f(x)$ и $g(z)$, причем факторы x и z также определяются с предельными абсолютными погрешностями Δx и Δz. Для оценки предельной относительной погрешности параметра Y рассчитаем дифференциал dY, используя уравнение (2):

$$Y(x,z) = f(x) \cdot g(z)\,, \tag{6}$$

$$dY = \left[f(x) \cdot g(z)\right]_x' dx + \left[f(x) \cdot g(z)\right]_z' dz\,, \tag{7}$$

$$dY = f_x'(x)dx \cdot g(z) + f(x) \cdot g_z'(z)dz\,. \tag{8}$$

Разделив уравнение (8) на уравнение (6) и заменив дифференциалы соответствующими приращениями $dY \approx \Delta Y$, $dx \approx \Delta x$, $dz \approx \Delta z$, получим уравнение для определения предельной относительной погрешности параметра Y, который является произведением 2-х функций (объяснение появления модулей см. в тексте после уравнения (5)):

$$\frac{\Delta Y}{Y} = \left|\frac{f_x'(x)}{f(x)}\right|\Delta x + \left|\frac{g_z'(z)}{g(z)}\right|\Delta z\,. \tag{9}$$

Предельная абсолютная погрешность ΔY параметра Y с помощью уравнения (9) рассчитывается следующим способом:

$$\Delta Y = \left(\left|\frac{f_x'(x)}{f(x)}\right|\Delta x + \left|\frac{g_z'(z)}{g(z)}\right|\Delta z\right)Y\,. \tag{10}$$

Используя такую же процедуру, найдем предельную относительную погрешность параметра Y, который является частным двух функций $f(x)$ и $g(z)$, причем факторы x и z также определяются с предельными абсолютными погрешностями Δx и Δz:

$$Y(x,z) = \frac{f(x)}{g(z)}\,, \tag{11}$$

$$dY = \left[\frac{f(x)}{g(z)}\right]_x' dx + \left[\frac{f(x)}{g(z)}\right]_z' dz\,, \tag{12}$$

$$dY = \frac{f_x'(x)dx}{g(z)} - \frac{f(x)g'(z)dz}{g_z^2(z)}, \qquad (13)$$

где
$$\left[\frac{1}{g(z)}\right]' = \left[g^{-1}(z)\right]' = (-1)\cdot g^{-2}(z)g_z'(z) = -\frac{g_z'(z)}{g^2(z)}. \qquad (14)$$

Разделив уравнение (13) на уравнение (11), получим

$$\frac{dY}{Y} = \frac{f_x'(x)dx}{g(z)}\cdot\frac{g(z)}{f(x)} - \frac{f(x)g_z'(z)dz}{g^2(z)}\cdot\frac{g(z)}{f(x)}. \qquad (15)$$

Заменив дифференциалы соответствующими приращениями с учетом пояснений, данных после уравнения (5), получим

$$\frac{\Delta Y}{Y} = \left|\frac{f_x'(x)}{f(x)}\right|\Delta x + \left|\frac{g_z'(z)}{g(z)}\right|\Delta z. \qquad (16)$$

Обратите внимание на то, что уравнения (9) и (16) тождественны, и, следовательно, предельная относительная погрешность параметра Y рассчитывается по одной и той же формуле независимо от того, является ли параметр Y произведением или частным двух функций $f(x)$ и $g(z)$ (уравнения (6), (11)).

После определения предельной абсолютной погрешности результат детерминированного эксперимента следует записать в виде $Y \pm \Delta Y$. Поскольку Y и ΔY расчетные величины, то они могут быть получены с любым количеством знаков, и поэтому важно корректно их округлить. Очевидно, что нет смысла приводить значение параметра Y до четвертого знака после запятой, если предельная абсолютная погрешность ΔY имеет значащую цифру уже в первом знаке после запятой. Поэтому при записи результата будем придерживаться следующих правил округления:

1) если при округлении отбрасываются цифры 0, 1, 2, 3, 4, то сохраняемая цифра остается без изменений. Если при округлении отбрасываются цифры 5 (при условии, что после неё следуют еще другие цифры отличные от нуля), 6, 7, 8, 9, сохраняемая цифра увеличивается на единицу. Если при округлении отбрасывается **только одна** цифра 5 (то есть когда за ней стоят только нули), сохраняемая цифра, если она четная, остается без изменений, а если она нечетная, увеличивается на единицу;

2) предельную абсолютную погрешность округлить до 2-х значащих цифр;

3) если эти две значащие цифры 10, 11, …, 54, то результат округления оставить без изменения;

4) если эти две значащие цифры 55, 56, …, 98, 99, то результат округления повторно округлить до одной значащей цифры, то есть до 60, 60, …, 100, 100. **Последний** ноль, возникший в результате повторного округления, если он стоит в дробной части, в конечном результате не пишется;

5) после округления предельной абсолютной погрешности рассчитываемый параметр округлить до значащих цифр этой погрешности.

Приведем примеры корректной записи результатов расчета:

19.5312 ± 0.0123 → 19.5**31** ± 0.0**12**;

4.3251 ± 0.1252 → 4.**33** ± 0.**13**;

4.325 ± 0.125 → 4.**32** ± 0.**12**;

54.325 ± 1.95 → 5**4.3** ± **2.0**;

19.5552 ± 0.0575 → 19.555 ± 0.058 → 19.5**6** ± 0.0**6**;

0.01989 ± 0.00783 → 0.0199 ± 0.0078 → 0.02**0** ± 0.00**8**;

19.5312 ± 0.0968 → 19.531 ± 0.097 → 19.**53** ± 0.**10**;

1 931.12 ± 98.36 → 1 931 ± 98 → 1 **93**0 ± **10**0;

978 256 ± 98 432 → 978 000 ± 98 000 →**98**0 000 ± **10**0 000.

1.3.2. Алгоритм решения прикладных задач

1. Параметр $Y = f(x) \pm g(z)$, причем факторы x и z определяются с предельной абсолютной погрешностью Δx и Δz. Тогда предельная абсолютная погрешность ΔY находится из следующего уравнения:

$$\Delta Y = \left|f_x'(x)\right|\Delta x + \left|g_z'(z)\right|\Delta z.$$

2. Параметр $Y = f(x)\cdot g(z)$ или $Y = \dfrac{f(x)}{g(z)}$, причем факторы x и z определяются с предельной абсолютной погрешностью Δx и Δz. То-

гда предельная относительная $\Delta Y/Y$ и предельная абсолютная погрешности ΔY находятся из следующих уравнений:

$$\frac{\Delta Y}{Y} = \left|\frac{f'_x(x)}{f(x)}\right|\Delta x + \left|\frac{g'_z(z)}{g(z)}\right|\Delta z ,$$

$$\Delta Y = \left(\left|\frac{f'_x(x)}{f(x)}\right|\Delta x + \left|\frac{g'_z(z)}{g(z)}\right|\Delta z\right)Y .$$

3. Для корректной записи рассчитываемого параметра $Y \pm \Delta Y$ используются следующие правила округления: 1) Если при округлении отбрасываются цифры 0, 1, 2, 3, 4, то сохраняемая цифра остается без изменений. Если при округлении отбрасываются цифры 5 (при условии, что после нее следуют еще другие цифры отличные от нуля), 6, 7, 8, 9, сохраняемая цифра увеличивается на единицу. Если при округлении отбрасывается **только одна** цифра 5 (то есть когда за ней стоят только нули), сохраняемая цифра, если она четная, остается без изменений, а если она нечетная, увеличивается на единицу; 2) предельную абсолютную погрешность округлить до 2-х значащих цифр; 3) если эти две значащие цифры 10, 11, …, 54, то результат округления оставить без изменения; 4) если эти две значащие цифры 55, 56, …, 98, 99, то результат округления повторно округлить до одной значащей цифры, то есть до 60, 60, …, 100, 100. **Последний** ноль, возникший в результате повторного округления, если он стоит после запятой, в конечном результате не пишется; 5) после округления предельной абсолютной погрешности рассчитываемый параметр округлить до значащих цифр этой погрешности.

1.3.3. Базовая задача

На токарном станке отрезали две заготовки ∅40×20 из сталей марок Сталь 20 и Сталь 45. К сожалению, рабочий не пометил заготовки, и они были перепутаны. Знание марки необходимо, так как, например, из Стали 45 путем закалки можно получить детали с большей прочностью и твердостью.

Возникла идея определить марку стали простым и дешевым способом – по плотности. Известно, что Сталь 20 имеет бóльшую плотность, нежели Сталь 45, так как Сталь 20 содержит ≈ 0.2 % уг-

лерода, а Сталь 45 содержит ≈ 0.45 % углерода (из справочника $\rho_{20} = 7.86$ г/см^3, а $\rho_{45} = 7.81$ г/см^3).

Рассчитаем плотность цилиндрической заготовки по формуле

$$\rho = \frac{M}{V} = \frac{4M}{\pi D^2 H}, \tag{17}$$

где ρ – плотность заготовки, г/см^3; M – масса заготовки, г; V – объем заготовки, см^3; D – диаметр заготовки, см; H – высота заготовки, см; $\pi = 3.1415926535897932...$.

Массу заготовок можно определить на весах разного класса точности с погрешностью 10; 1, 0.1; 0.01; 0.001; 0.0001 г. Линейные размеры также можно определить различными инструментами с погрешностью 1, 0.1, 0.01, 0.001 мм. Следует отметить, что чем точнее инструмент, тем он дороже и тем труднее найти его в повседневной практике. Поэтому наша задача – подобрать дешевые и доступные инструменты, которые, тем не менее, позволили бы определить массу и линейные размеры заготовок с такой точностью, чтобы со 100 % достоверностью различить заготовки по плотности и, следовательно, определить марку стали каждой заготовки.

Для того чтобы различить заготовки по плотности со 100 % достоверностью, предельная абсолютная погрешность в определении плотности $\Delta\rho$ каждой заготовки не должна превышать ¼ разницы между справочными значениями плотностей сталей обеих марок:

$$\Delta\rho = \frac{\rho_{20} - \rho_{45}}{4} = \frac{7.86 - 7.81}{4} = \frac{0.05}{4} = 0.0125 \approx 0.012 \text{ г/см}^3 . \tag{18}$$

Округление сделано согласно п. 3 алгоритма 1.3.2.

Обоснуем это утверждение. Определяя плотность Стали 45 с предельной погрешностью 0.012 г/см^3, в худшем случае, отталкиваясь от справочных данных, можно получить значение $\rho_{45} = 7.81 + 0.012 = 7.822$ г/см^3, а для Стали 20 $\rho_{20} = 7.86 - 0.012 = 7.848$ г/см^3. Тогда результат эксперимента запишется так:

$$\begin{cases} \rho_{45} = (7.822 \pm 0.012), \text{ г/см}^3; \\ \rho_{20} = (7.848 \pm 0.012), \text{ г/см}^3. \end{cases} \tag{19}$$

Из уравнения (19) следует, что максимальное значение плотности Стали 45, определенное экспериментально, может быть $\rho_{45,\,max} = 7.822 + 0.012 = 7.834$ г/см^3, а минимальное значение плотности Стали 20, определенное экспериментально, может быть $\rho_{20,\,min} = 7.848 - 0.012 = 7.836$ г/см^3. Таким образом, если предельная абсолютная погрешность определения плотности составит 0.012 г/см^3 (см. уравнение (18)), интервалы экспериментальных значений плотностей заготовок не перекроются. Если же предельная абсолютная погрешность определения плотности заготовок будет больше $\Delta\rho > 0.012$ г/см^3 (уже при $\Delta\rho = 0.013$ г/см^3 – убедитесь самостоятельно), существует некоторая вероятность того, что интервалы экспериментальных значений плотностей заготовок могут перекрыться и с 100 % достоверностью определить марку стали будет невозможно.

Обосновав предельное значение абсолютной погрешности определения плотности, рассчитаем относительную погрешность:

$$\frac{\Delta\rho}{\rho} \approx \frac{0.012}{7.86} \approx \frac{0.012}{7.81} \approx 0.0015\,. \qquad (20)$$

Согласно алгоритму 1.3.2 (п. 2) получаем

$$\frac{\Delta\rho}{\rho} = \frac{\Delta M}{M} + 2\cdot\frac{\Delta D}{D} + \frac{\Delta H}{H} + \frac{\Delta\pi}{\pi}\,. \qquad (21)$$

Первые три слагаемых правой части уравнения (21) имеет смысл выбрать таким образом, чтобы они были равны между собой, то есть

$$\frac{\Delta M}{M} \approx 2\cdot\frac{\Delta D}{D} \approx \frac{\Delta H}{H} \approx \frac{0.0015}{3} = 0.00050\,. \qquad (22)$$

Для расчета предельных абсолютных погрешностей по уравнению (22) можно воспользоваться приблизительными значениями массы, диаметра и высоты заготовок. Приблизительные значения диаметра и высоты каждой заготовки $D \approx 40$ мм, $H \approx 20$ мм взяты из условия задачи $\varnothing 40\times 20$. Приблизительное значение массы каждой заготовки определим расчетным путем по формуле: $M = \pi D^2 H\rho/4 = 3.14\cdot 4^2\cdot 2\cdot 7.835/4 \approx 197$ г (за величину плотности взяли среднее справочное значение $(7.81 + 7.86)/2 = 7.835$)). Тогда из уравнения (22) следует, что

$$\Delta M \approx 0.0005\cdot M = 0.0005\cdot 197\ \text{г} = 0.0985\ \text{г} \approx 0.1\ \text{г}\,; \qquad (23)$$

$$\Delta D \approx \frac{0.0005\cdot D}{2} = \frac{0.0005\cdot 40}{2}\ \text{мм} = 0.01\ \text{мм}\,; \qquad (24)$$

$$\Delta H \approx 0.0005\cdot H = 0.0005\cdot 20\ \text{мм} = 0.01\ \text{мм}\,. \qquad (25)$$

Для того чтобы относительная погрешность $\Delta\pi/\pi$ не увеличивала погрешность определения плотности заготовок (см. уравнение (21)), возьмем число π с относительной погрешностью на два порядка меньше суммарной относительной погрешности (см. уравнение (20)), то есть $\Delta\pi/\pi = 0.000015$.

Тогда с учетом правил округления п. 3 алгоритма 1.3.2 имеем

$$\Delta\pi \approx 0.000015\cdot 3.14 \approx 0.000047.$$

Так как число $\pi = 3.1415926535897932\ldots$, то его приближенными значениями являются 3 (– 0.14), 3.1 (– 0.042), 3.14 (– 0.0016), 3.142 (+ 0.00041), 3.1416 (+ 0.000007), 3.13159 (– 0.0000027) и т. д. (в скобках указана абсолютная погрешность выбранного приближения). Для обеспечения абсолютной погрешности 0.000047 следует взять $\pi \approx 3.1416$. Тогда $\Delta\pi/\pi \approx 0.000007/3.1416 \approx 0.0000022$, что меньше выбранного значения $\Delta\pi/\pi = 0.000015$ в 7 раз. Заметим, что общепринятое значение $\pi = 3.14$ ($\Delta\pi/\pi \approx 0.00051$) и даже 3.142 ($\Delta\pi/\pi \approx 0.00013$) не обеспечило бы требуемую относительную погрешность в определении плотности. Следует, однако, отметить, что число π, взятое с любым приближением, искажает результат расчета плотности на обеих заготовках в одну и ту же сторону и, таким образом, не меняет соотношения значений плотностей для сталей различных марок. Но для чистоты эксперимента число π лучше взять с такой погрешностью, чтобы она дополнительно не искажала результат, тем более сделать это легко.

Проанализировав уравнения (23)–(25), можно сделать вывод, что для определения плотности заготовок с допустимой предельной абсолютной погрешностью массу заготовок необходимо измерять на технических весах (погрешность 0.1 г), а диаметр и длину заготовок - с помощью микрометра (погрешность 0.01 мм). Для расчетов плотности следует использовать $\pi = 3.1416$, так как такое приближение не увеличивает погрешность рассчитываемых параметров. Результаты измерений массы, диаметра и высоты заготовок выбранными инструментами с допустимой погрешностью, промежуточные, а также конечные результаты расчета плотности заготовок приведены в таблице 2.

Таблица 2 – Данные по определению плотности заготовок

Заготовка	M, г	D, см	H, см	$\Delta M/M$	$2\cdot\Delta D/D$	$\Delta H/H$	$\Delta\rho/\rho$	ρ, г/см^3	$\Delta\rho$, г/см^3
№ 1	196.6	4.001	2.001	0.00051	0.00050	0.00050	0.0015	7.8146	0.0118
№ 2	197.1	3.999	2.000	0.00051	0.00050	0.00050	0.0015	7.8463	0.0118

Плотность заготовок рассчитывалась по уравнению (17), предельные относительные и абсолютные погрешности определения плотности - по уравнению (21):

$$\rho_1 = \frac{4M}{\pi D^2 H} = \frac{4\cdot 196.1}{3.1416\cdot 4.001^2\cdot 2.001} = 7.8146\,,$$

$$\rho_2 = \frac{4M}{\pi D^2 H} = \frac{4\cdot 197.2}{3.1416\cdot 3.999^2\cdot 2.000} = 7.8463\,,$$

$$\frac{\Delta\rho_1}{\rho_1} = \frac{\Delta\rho_2}{\rho_2} = 0.00051 + 0.00050 + 0.00050 = 0.00151 \approx 0.0015\,,$$

$$\Delta\rho_1 = \left(\frac{\Delta\rho_1}{\rho_1}\right)\cdot\rho_1 = 0.00151\cdot 7.8146 = 0.0118 \approx 0.012\,,$$

$$\Delta\rho_2 = \left(\frac{\Delta\rho_2}{\rho_2}\right)\cdot\rho_2 = 0.00151\cdot 7.8463 = 0.0118 \approx 0.012\,.$$

Согласно правилам округления (см. п. 3 алгоритма 1.3.2) приведем корректную запись результата:

$$\begin{cases}\rho_1 = 7.815 \pm 0.012\ \text{г}/\text{см}^3;\\ \rho_2 = 7.846 \pm 0.012\ \text{г}/\text{см}^3.\end{cases}$$

Таким образом, со 100 % достоверностью можно утверждать, что материал заготовки № 1 – Сталь 45, а заготовки № 2 – Сталь 20, так как интервалы рассчитанных значений плотности не перекрываются.

1.3.4. Материалы для самостоятельной работы, лабораторных и практических занятий

Тесты

1. Предельная абсолютная ошибка определения параметра Y в общем случае определяется погрешностью измерительных приборов $\Delta Y_{\text{пр}}$ и стохастической природой исследуемого объекта $\Delta Y_{\text{ст}}$ и рассчитывается по формуле:

а) $\pm\Delta Y = \pm\Delta Y_{\text{пр}} \pm \Delta Y_{\text{ст}}$; б) $\Delta Y^2 = \Delta Y_{\text{пр}}^2 + \Delta Y_{\text{ст}}^2$;

в) $\Delta Y = \Delta Y_{\text{пр}} + \Delta Y_{\text{ст}}$; г) $\sqrt{\Delta Y^2} = \sqrt{\Delta Y_{\text{пр}}^2} + \sqrt{\Delta Y_{\text{ст}}^2}$.

2. Факторы x и z в детерминированном эксперименте определены с предельной абсолютной погрешностью Δx и Δz. Если $Y = f(x)\cdot g(z)$, то

а) $\dfrac{\Delta Y}{Y} = \dfrac{\Delta x}{x} + \dfrac{\Delta z}{z}$; б) $\dfrac{\Delta Y}{Y} = \dfrac{f'_x(x)\Delta x}{f(x)} + \dfrac{g'_z(z)\Delta z}{g(z)}$;

в) $\dfrac{\Delta Y}{Y} = \dfrac{f(x)\Delta x}{f'_x(x)} + \dfrac{g(z)\Delta z}{g'_z(z)}$; г) $\dfrac{\Delta Y}{Y} = \dfrac{f'_x(x)\Delta x}{f(x)} \pm \dfrac{g'_z(z)\Delta z}{g(z)}$.

3. Факторы x и z в детерминированном эксперименте определены с предельной абсолютной погрешностью Δx и Δz. Если $Y = f(x)/g(z)$, то

а) $\dfrac{\Delta Y}{Y} = \dfrac{f'_x(x)\Delta x}{f(x)} - \dfrac{g'_z(z)\Delta z}{g(z)}$; б) $\dfrac{\Delta Y}{Y} = \dfrac{\Delta x}{f(x)} + \dfrac{\Delta z}{g(z)}$;

в) $\dfrac{\Delta Y}{Y} = \dfrac{f'_x(x)\Delta x}{f(x)} + \dfrac{g'_z(z)\Delta z}{g(z)}$; г) $\dfrac{\Delta Y}{Y} = \dfrac{f(x)\Delta x}{x} + \dfrac{g(z)\Delta z}{\text{z}}$.

4. В какой формуле округление выполнено корректно:

а) $29.735 \pm 0.0778 \to 29.74 \pm 0.08$;

б) $29.745 \pm 0.0778 \to 29.74 \pm 0.08$;

в) $29.734 \pm 0.0778 \to 29.7 \pm 0.1$;

г) $29.7345 \pm 0.0778 \to 29.74 \pm 0.07$?

Вопросы для самопроверки

1. В чем концептуальное отличие детерминированных закономерностей от стохастических?
2. Какова природа погрешности определяемого параметра Y в детерминированном эксперименте?
3. Что такое предельная абсолютная погрешность?
4. Что такое предельная относительная погрешность?
5. Сформулируйте правила корректного округления рассчитанных параметров в детерминированном эксперименте.

Задачи для самостоятельного решения

Уровень 1

1. Мощность двигателя $W = I \cdot U$, где I – ток, U – напряжение. Рассчитать предельную абсолютную погрешность при определении мощности двигателя, если погрешность определения тока – ΔI, а напряжения – ΔU.

Ответ: $\Delta W = IU\left(\dfrac{\Delta I}{I} + \dfrac{\Delta U}{U}\right)$.

2. Дано $Y = x^5 \cdot e^{2z}$, причем факторы x и z определены с предельной абсолютной погрешностью Δx и Δz. Доказать, что

$$\frac{\Delta Y}{Y} = 5 \cdot \frac{\Delta x}{x} + 2 \cdot \Delta z\ .$$

3. Дано $Y = \ln(\beta x) \cdot \cos(\gamma z)$, причем факторы x и z определены с предельной абсолютной погрешностью Δx и Δz. Доказать, что

$$\frac{\Delta Y}{Y} = \frac{\Delta x}{x \cdot \ln(\beta x)} + \gamma \cdot \mathrm{tg}(\gamma z) \cdot \Delta z\ .$$

4. Дано $Y = \cos x \big/ \sqrt[3]{z}$, причем факторы x и z определены с предельной абсолютной погрешностью Δx и Δz. Доказать, что

$$\frac{\Delta Y}{Y} = \mathrm{tg}x \cdot \Delta x + \frac{1}{3} \cdot \frac{\Delta z}{z}\ .$$

Уровень 2

5. Вольтметром с относительной погрешностью 3 % измерено напряжение электрической сети 225 В. Омметром с относительной погрешностью 4 % измерено сопротивление электродвигателя 50 Ом. Определить мощность электродвигателя и ее абсолютную погрешность.

Ответ: (1060 ± 110) Вт или (1.06 ± 0.11) кВт.

6. Автомобиль массой 4.8 т движется равномерно и прямолинейно со скоростью 72 км/ч. Предельная абсолютная погрешность определения массы автомобиля 150 кг, скорости 1.5 м/с. Определить кинетическую энергию автомобиля и предельную абсолютную погрешность.

Ответ: (960 000 ± 170 000) Дж или (0.96 ± 0.17) МДж.

Уровень 3

7. Дано $Y = f(x) \cdot g(z)\ h(w)$, причем факторы x, z, и w в детерминированном эксперименте определены с предельной абсолютной погрешностью Δx, Δz, Δw. Доказать, что

$$\frac{\Delta Y}{Y} = \frac{f'_x(x)\Delta x}{f(x)} + \frac{g'_z(z)\Delta z}{g(z)} + \frac{h'_w(w)\Delta w}{h(w)}\ .$$

Уровень 4

8. Даны функции: $f(x) = \psi(x) + \xi(x)$ и $g(x) = \psi(x)$, причем $\xi(x)/\psi(x) << 1$ и $\xi'(x)/\psi'(x) << 1$. Рассчитать предельную относительную погрешность параметра Y, который задается следующим уравнением:

$$Y = \frac{f(x) - g(x)}{f(x) + g(x)}\ .$$

Предельная абсолютная погрешность определения фактора Δx.

Ответ: $\dfrac{\Delta Y}{Y} = \left(\dfrac{\xi'(x)}{\xi(x)} + \dfrac{\psi'(x)}{\psi(x)}\right)\Delta x$.

1.4. Обработка данных стохастического эксперимента. Одна выборка (тема 3)

Исследуемый объект, описываемый детерминистическими закономерностями, при заданных значениях факторов всегда находится в единственно возможном состоянии со 100 % вероятностью. В отличие от этого, состояние объекта, описываемого стохастическими закономерностями, в отдельно взятом опыте концептуально непредсказуемо даже при соблюдении одинаковых условий эксперимента. Причем следует отметить, что один и тот же объект может описываться или детерминированными, или стохастическими закономерностями в зависимости от того, какие параметры этого объекта изучаются. Проиллюстрируем приведенное утверждение на примерах экспериментов с монетой.

Первый эксперимент состоит в определении массы монеты путем взвешивания на весах. Пусть масса монеты – $(m_0 \pm \Delta)$ г, где Δ – погрешность используемых весов ($\Delta << m_0$). Сколько бы опытов ни было проведено, в пределах погрешности будем наблюдать одно и то же значение массы, так как никакие случайные факторы в условиях данного эксперимента не могут повлиять на результат измерения. Результаты этого эксперимента позволяют утверждать со 100 % вероятностью, что масса монеты $(m_0 \pm \Delta)$ г, и, следовательно, объект в данном примере описывается детерминированными закономерностями.

Во втором эксперименте будем подбрасывать ту же монету вверх и фиксировать, какой стороной она упала на землю. Результат каждого отдельного наблюдения в таком эксперименте принципиально непредсказуем. Можно только утверждать, что при большом количестве опытов каждая сторона монеты будет выпадать в ≈ 50 % случаев, так как в данном эксперименте наблюдаемый объект может принимать только 2 равновероятных состояния. Следовательно, в данном примере объект описывается стохастическим закономерностям.

В третьем эксперименте будем бросать ту же монету с некоторого расстояния на прямую, проведенную на земле. В каждом опыте будем измерять отклонение центра монеты от прямой (при перелете отклонения будем считать положительными, при недолете – отрицательными). Результат каждого опыта, так же как и в предыдущем примере, непредсказуем, и поэтому объект и в этом эксперименте описывается стохастическим закономерностям. Различие заключается в том, что во втором примере спектр возможных состояний объекта дискретный, а в третьем – непрерывный.

В общем случае для объектов, подчиняющихся стохастическим закономерностям, результат опыта предсказывается только с некой вероятностью, определяемой некоторой функцией плотности вероятности, характеризующей поведение объекта в условиях данного эксперимента. Во втором примере вероятность всех возможных состояний одинакова и равна ½. В третьем примере, из общих соображений, можно ожидать, что отклонения, близкие к нулю, будут наиболее вероятными, а большие отклонения – менее вероятными. Функция плотности вероятности отклонений в экспериментах такого рода будет иметь колоколообразный вид, описываемый законном нормального распределения случайной величины (распределение Гаусса) (рисунок 1).

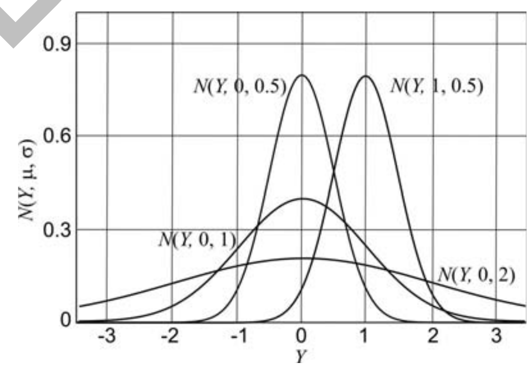

Рисунок 1 – Закон нормального распределения

Теоретическая функция нормального распределения однозначно задается двумя параметрами μ и σ:

$$N(Y, \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(Y-\mu)^2}{2\sigma^2}}. \qquad (26)$$

Число состояний исследуемого объекта, «реализуемое» при бесконечном числе опытов, образует генеральную совокупность. Положение вершины кривой, описываемой уравнением (26), относительно начала координат определяется параметром μ (генеральное среднее), а рассеяние случайной величины Y относительно генерального среднего μ определяется параметром σ (генеральное стандартное отклонение), что наглядно проиллюстрировано на рисунке 1. Параметры Y, μ, σ имеют одинаковую размерность. Величина

σ^2 называется генеральной дисперсией. Нормальный закон распределения обозначают следующим образом: $N(Y, 1, 0.5)$, где N – нормальный закон распределения случайной величины, Y – случайная величина, распределение которой описывается нормальным законом с параметрами: генеральное среднее $\mu = 1$ и генеральное стандартное отклонение $\sigma = 0.5$ (генеральная дисперсия $\sigma^2 = 0.25$).

Нормальный закон распределения случайных величин играет фундаментальную роль в математической статистике, поскольку, во-первых, хорошо описывает распределение случайных отклонений исследуемого параметра, обусловленных множеством малых независимых факторов, носящих случайный характер, а во-вторых, многие функции параметра Y распределены по нормальному закону распределения при объеме выборки $n \geq 30$ даже, если сама случайная величина Y не подчиняется нормальному закону распределения.

В данном учебно-методическом комплексе ограничимся изучением стохастических объектов, для которых случайные значения исследуемого параметра подчиняются нормальному закону распределения.

Уравнение (26) позволяет рассчитать вероятность попадания случайных значений параметра Y в некоторый интервал. Для приведенных ниже интервалов вероятность попадания случайных нормально распределенных величин Y при условии, что $p(-\infty < Y < +\infty) = 1$, такова:

$$p(-\sigma \leq Y \leq \sigma) = 0.6827\,, \tag{27}$$

$$p(-2\sigma \leq Y \leq 2\sigma) = 0.9545\,, \tag{28}$$

$$p(-3\sigma \leq Y \leq 3\sigma) = 0.9973\,. \tag{29}$$

Таким образом, можно считать, что практически все значения параметра Y, подчиняющиеся нормальному закону распределения, находятся в интервал $\mu \pm 3\sigma$ (на самом деле вероятность этого утверждения равна 99.73 %). Однако для анализа технических объектов используют менее жесткое условие: принято считать, что все случайные значения параметра Y, подчиняющиеся нормальному закону распределения, находятся в интервале $\mu \pm 2\sigma$, что недалеко от истины (вероятность этого утверждения равна 95.45 %). Поэтому в дальнейшем при статистической обработке экспериментальных данных будем использовать доверительную вероятность $p = 0.95$.

Экспериментально определить генеральные параметры μ и σ достаточно «просто». Для этого необходимо определить **все** случайные значения генеральной совокупности, а затем рассчитать параметры μ и σ (например, по уравнениям (30)–(32) при $n = \infty$). Очевидно, что такой совет нереализуем на практике. Однако в томто и состоит концептуальная суть математической статистики, что ее методы позволяют оценивать генеральные параметры μ и σ, используя для этого относительно небольшое количество случайных величин, образующих выборку. Однако, «платой» за такую «экономию» будет некоторая погрешность в определении μ и σ, и причем тем большая, чем меньше объем выборки.

1.4.1. Теория

Оценка параметров генеральной совокупности по выборке. Рассмотрим некоторую выборку случайных величин $Y_1,\ldots, Y_n$, принадлежащую генеральной совокупности с параметрами μ и σ. Оценка параметров μ и σ производится по выборочному среднему $\overline{Y}$ и выборочному стандартному отклонениям S:

$$\overline{Y} = \frac{Y_1 + \ldots + Y_n}{n} = \frac{\sum_{i=1}^{n} Y_i}{n}\,; \tag{30}$$

$$S^2(Y) = \frac{(Y_1 - \overline{Y})^2 + \ldots + (Y_n - \overline{Y})^2}{n-1} = \frac{\sum_{i=1}^{n} (Y_i - \overline{Y})^2}{n-1}\,, \quad f = n-1; \tag{31}$$

$$S(Y) = \sqrt{S^2(Y)}\,, \quad f = n-1, \tag{32}$$

где f – число степеней свободы рассчитываемого выборочного параметра равно объему выборки за вычетом числа параметров, рассчитанных по этой же выборки и используемых при расчете.

Отметим несколько полезных свойств выборочного среднего и выборочной дисперсии (доказательства приведены в приложении 6):

$$\overline{CY} = C\overline{Y}\,, \tag{33}$$

$$\overline{Y + Z} = \overline{Y} + \overline{Z}\,, \tag{34}$$

$$S^2(C) = 0, \qquad (35)$$

$$S^2(CY) = C^2 S^2(Y), \qquad (36)$$

$$S^2(Y \pm Z) = S^2(Y) + S^2(Z), \qquad (37)$$

$$S^2(\overline{Y}) = \frac{S^2(Y)}{n} \quad \text{или} \quad S(\overline{Y}) = \frac{S(Y)}{\sqrt{n}}. \qquad (38)$$

Приведем некоторые пояснения и уточнения к формулам (33)–(38). Случайные величины Y, CY, Z относятся к выборкам равного объема n, например, Y_1, …, Y_n. Величины Y и Z независимы друг от друга, что означает равенство нулю коэффициента корреляции между ними (приложение 6, вывод уравнения (37)). Что касается уравнения (38), то в нем параметр $\overline{Y}$ рассматривается как случайная величина, а не константа. Поясним это. Пусть случайная величина Y принадлежит к некоторой генеральной совокупности с параметрами μ и σ. Выборка $Y_1', ..., Y_n'$ характеризуется выборочным средним $\overline{Y'}$. Если на базе той же генеральной совокупности провести еще несколько серий наблюдений, то получим $Y_1'', ..., Y_n''$; $Y_1''', ..., Y_n'''$ и т. п. Все они будут характеризоваться выборочными средними $\overline{Y''}$, $\overline{Y'''}$ и т. п. То есть величина $\overline{Y}$ сама является случайной величиной, которая к тому же подчиняется нормальному закону распределения с генеральным стандартным отклонением в $\sqrt{n}$ раз меньшим, чем генеральное стандартное отклонение случайной величины Y. Поэтому увеличение объема выборки повышает точность определения генерального среднего (полезное для экспериментаторов утверждение).

Проверка выборки случайных величин на промахи. В выборке иногда содержатся промахи (грубые ошибки), которые возникают при ошибочных действиях экспериментатора или незамеченном внешнем воздействии. Естественно, что промахи существенно искажают выборочные параметры $\overline{Y}$, S, а те, в свою очередь, искажают оценку генеральных параметров μ и σ. Поэтому прежде чем приступить к оценке параметров генеральной совокупности необходимо провести статистический анализ экспериментальных данных, направленный на исключение промахов.

Выявление промахов осуществляют с помощью критерия Смирнова–Граббса следующим образом: сначала рассчитывают экспериментальное значение критерия Смирнова–Граббса:

$$\tau_э = \frac{\max\left|Y_i - \overline{Y}\right|}{S(Y)}, \; f = n - 2. \qquad (39)$$

Величина Y_i является промахом, если

$$\tau_{\max} > \tau_{n-2,\,p}, \qquad (40)$$

где $\tau_{n-2,\,p}$ – критическое значение критерия Смирнова–Граббса (приложение 1). Математический смысл критерия Смирнова–Граббса: значение случайной величины, подчиняющейся нормальному закону распределения и лежащее за границей 3σ, при достаточно большом объеме выборки считается промахом. При небольшом объеме выборки это требование становится более жестким, так как оценка генерального среднего и генерального стандартного отклонений по выборке относительно малого объема связано с относительно большой погрешностью. Так, например, для выборки объемом $n = 10$ ($f = 8$) значение, лежащее за границей $2.29 \cdot S$, считается промахом (приложение 1).

После исключения **одного** промаха, оставшаяся выборка вновь подвергается проверке на промахи по такой же процедуре (находят новые выборочное среднее, выборочное стандартное отклонение, экспериментальное и критическое значения критерия Смирнова–Граббса).

Проверка соответствия случайных величин выборки нормальному закону распределения. После исключения промахов необходимо проверить гипотезу о подчинении случайных величин «очищенной» выборки нормальному закону распределения. Дело в том, что весь математический аппарат, применяемый для обработки экспериментальных данных, изложенный в этом учебно-методическом комплексе, основан на предположении принадлежности случайных величин нормальному закону распределения. Для проверки случайных величин выборки на принадлежность их нормальному закону существует несколько критериев. Приведем один из них, достаточно простой, для выборок с объемом $n < 120$:

$$\left|\frac{\text{САО}}{S(Y)} - 0.7979\right| < \frac{0.4}{\sqrt{n}}, \tag{41}$$

$$\text{САО} = \frac{\sum_{i=1}^{n}\left|Y_i - \overline{Y}\right|}{n}, \tag{42}$$

Если выполняется неравенство (41), то с 95 % доверительной вероятностью принимается гипотеза о том, что случайные величины «очищенной» выборки подчиняются закону нормального распределения.

Следует подчеркнуть, что в математической теории эксперимента проверка случайных величин выборки на промахи и на принадлежность их нормальному закону распределения не является необходимой операцией перед проведением дальнейших расчетов. Но в данном учебно-методическом комплексе все математические выкладки будут производиться только со случайными величинами «очищенных» выборок, подчиняющихся нормальному закону распределения.

Оценка генеральных параметров μ и σ по данным выборки. Так как выборка имеет ограниченное количество случайных величин, параметры генеральной совокупности μ и σ оцениваются с некоторой погрешностью, определяемой посредством доверительных интервалов. Доверительный интервал для генерального среднего μ определяется по формуле (см. уравнение (38)):

$$\begin{cases} \Delta\overline{Y} = \left|\mu - \overline{Y}\right| < t_{n-1,\,p} S(\overline{Y}) = t_{n-1,\,p}\dfrac{S(Y)}{\sqrt{n}}; \\ \overline{Y} - t_{n-1,\,p}\dfrac{S(Y)}{\sqrt{n}} < \mu < \overline{Y} + t_{n-1,\,p}\dfrac{S(Y)}{\sqrt{n}}, \end{cases} \tag{43}$$

где $t_{n-1,\,p}$ – критическое значение критерия Стьюдента (приложение 2).

Доверительный интервал для генеральной дисперсии σ^2 задается следующим неравенством:

$$\frac{(n-1)S^2}{\chi^2_{n-1,\,(1-p)/2}} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{n-1,\,(1+p)/2}}, \tag{44}$$

где $\chi^2_{n-1,\,(1-p)/2}$, $\chi^2_{n-1,\,(1+p)/2}$ – нижнее и верхнее критические значения критерия Пирсона (приложение 3).

Определение минимального объема выборки, обеспечивающего требуемую точность в оценке генерального среднего. При проведении научных исследований преследуют две противоречивыми цели: с одной стороны хотят получить максимальную точность (минимальную погрешность) при оценке генерального среднего, а с другой – выполнить как можно меньше опытов. Противоречие заключается в том, что высокая точность оценки генерального среднего требует большого числа опытов (см. уравнения (43)), а стремление уменьшить материальные затраты требует проведения малого количества опытов.

На практике приоритетным требованием является допустимая величина погрешности $\Delta\overline{Y}_{\text{о}} = \left|\,\mu - \overline{Y}\,\right|$. Тогда минимальное количество опытов $n_{\min}$, при котором обеспечивается допустимая погрешность при оценке μ, можно было бы найти из уравнения (43):

$$n_{\min} = \left(t_{n_{\min}-1,\,p}\frac{S(Y)}{\Delta\overline{Y}_{\text{о}}}\right)^2. \tag{45}$$

Однако, с помощью уравнения (45) нельзя рассчитать $n_{\min}$, так как в его правой части содержится параметр $t_{n_{\min}-1,\,p}$, который сам зависит от $n_{\min}$. В связи с этим задача по определению минимального объема выборки $n_{\min}$, обеспечивающего допустимую погрешность $\Delta\overline{Y}_{\text{о}}$, решается в два этапа.

1. На первом этапе по данным исходной выборки объемом n по уравнениям (30)–(32), (43) и приложению 2 рассчитываются параметры $\overline{Y}$, S, $\Delta\overline{Y}$. Объем выборки n на первом этапе определяется интуитивно с учетом минимальных трудозатрат. Если получится, что $\Delta\overline{Y} < \Delta\overline{Y}_{\text{о}}$, то задача решена, интуиция не подвела экспериментатора. Если $\Delta\overline{Y} > \Delta\overline{Y}_{\text{о}}$, то переходят ко второму этапу.

2. На основе формулы (43) рассчитывают отношение $\Delta\overline{Y}_{\text{о}} / S$:

$$\frac{\Delta\overline{Y}_{\text{о}}}{S} = \frac{t_{n_{\min}-1,\,p}}{\sqrt{n_{\min}}}. \tag{46}$$

В приложении 2 в столбце $t_{n-1,\,p}\big/\sqrt{n}$ находят значение, близкое к рассчитанной величине $\Delta\overline{Y}_{о} / S$, а в столбце n – соответствующее ему $n_{\min}$ (округление производится в сторону больших значений n).

Важнейшим параметром, однозначно определяющим минимальное число опытов $n_{\min}$, является отношение $\Delta\overline{Y}_{о} / S$. Чем больше это отношение, тем меньше $n_{\min}$. Если отношение $\Delta\overline{Y}_{о} / S \geq 9$ (трудно представить практическую ценность такого требования), то из приложения 2 следует, что нужно проделать всего 2 опыта (меньше нельзя, так как для $n = 1$ невозможно рассчитать выборочное стандартное отклонение). Если отношение $\Delta\overline{Y}_{о} / S = 0.3$, то из приложения 2 находим, что $n_{\min} = 45$ опытам. Так как $t_{n_{\min}-1,\,p} \to 1.96 \approx 2$ (см. приложение 2), то минимальное количество опытов $n_{\min}$ можно рассчитать по приближенной формуле $n_{\min} = (2S / \Delta\overline{Y}_{о})^2 = (2 \cdot 3.333)^2 = 44.4 \approx 45$ (в данном случае округление до ближайшего большего целого дает число 45). То есть расчеты $n_{\min}$ и по приложению 2, и по приближенной формуле дают один и тот же результат. Для $\Delta\overline{Y}_{о} / S < 0.3$ значения $n_{\min}$, рассчитанные по приложению 2 и по приближенной формуле, совпадают с еще большей точностью. Поэтому

$$n_{\min} = \begin{cases} 2, \quad \text{если } \Delta\overline{Y}_{о} / S \geq 9; \\ \text{из условия } \dfrac{t_{n_{\min}-1,\,p}}{\sqrt{n_{\min}}} = \dfrac{\Delta\overline{Y}_{о}}{S}, \quad \text{если } 0.3 \leq \Delta\overline{Y}_{о} / S < 9; \\ (2S / \Delta\overline{Y}_{о})^2, \quad \text{если } \Delta\overline{Y}_{о} / S < 0.3. \end{cases} \qquad (47)$$

Если $0.3 \leq \Delta\overline{Y}_{о} / S < 9$, то для определения $n_{\min}$ следует воспользоваться данными таблицы приложения 2.

1.4.2. Алгоритм решения прикладных задач

Дана выборка объемом n: $Y_1, \ldots, Y_i, \ldots, Y_n$.

1. Рассчитать выборочное среднее:

$$\overline{Y} = \frac{\sum_{i=1}^{n} Y_i}{n}.$$

2. Рассчитать выборочную дисперсию и выборочное стандартное отклонение:

$$S^2 = \frac{\sum_{i=1}^{n} (Y_i - \overline{Y})^2}{n-1}, \quad f = n - 1,$$

$$S = \sqrt{S^2}, \quad f = n - 1.$$

3. Проверить значения выборки на наличие в ней промахов по критерию Смирнова–Граббса.

3.1. Рассчитать экспериментальное значение $\tau_э$:

$$\tau_э = \max \frac{\left|Y_i - \overline{Y}\right|}{S}, \quad f = n - 2.$$

3.2. Найти по таблице приложения 1 критическое значение критерия Смирнова–Граббса $\tau_{n-2,\,p}$.

3.3. Величина Y_i является промахом, если

$$\tau_э > \tau_{n-2,\,p}.$$

Из выборки исключается **только одно** случайное значение.

4. Проверить соответствие случайных величин выборки закону нормального распределения:

$$\left|\frac{\text{САО}}{S} - 0.7979\right| < \frac{0.4}{\sqrt{n}},$$

$$\text{САО} = \frac{\sum_{i=1}^{n}\left|Y_i - \overline{Y}\right|}{n}.$$

Если выполняется приведенное неравенство, то случайные значения параметра Y_i подчиняются нормальному закону распределения.

5. Рассчитать доверительный интервал при оценке генерального среднего μ по критерию Стьюдента (приложение 2):

$$\overline{Y} - \frac{t_{n-1,\,p} \cdot S(Y)}{\sqrt{n}} < \mu < \overline{Y} + \frac{t_{n-1,\,p} \cdot S(Y)}{\sqrt{n}},$$

где $t_{n-1,\,p}$ – критическое значение критерия Стьюдента (приложение 2).

6. Рассчитать доверительный интервал при оценке генеральной дисперсии σ^2 по критерию Пирсона (приложение 3):

$$\frac{(n-1)S^2}{\chi^2_{n-1,\,(1-p)/2}} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{n-1,\,(1+p)/2}},$$

где $\chi^2_{n-1,\,(1-p)/2}$, $\chi^2_{n-1,\,(1+p)/2}$ – критические значения критерия Пирсона (приложение 3).

7. Рассчитать минимальное число опытов для обеспечения допустимой абсолютной погрешности $\Delta\overline{Y}_{\text{о}}$ при оценке генерального среднего μ:

$$n_{\min} = \begin{cases} 2, \quad \text{если } \Delta\overline{Y}_{\text{о}}/S \ge 9; \\ \text{из условия } \dfrac{t_{n_{\min}-1,\,p}}{\sqrt{n_{\min}}} = \dfrac{\Delta\overline{Y}_{\text{о}}}{S}, \quad \text{если } 0.3 \le \Delta\overline{Y}_{\text{о}}/S < 9; \\ (2S/\Delta\overline{Y}_{\text{о}})^2, \quad \text{если } \Delta\overline{Y}_{\text{о}}/S < 0.3. \end{cases}$$

Если $0.3 \le \Delta\overline{Y}_{\text{о}}/S < 9$, то для определения $n_{\min}$ следует воспользоваться данными таблицы приложения 2.

1.4.3. Базовая задача

В вольерах птицефабрики содержится приблизительно 10 000 цыплят одного возраста. Поставлена задача: оценить суммарный вес цыплят с предельной относительной погрешностью 1 %.

Сформулируем задачу в терминах математической статистики. Генеральная совокупность состоит из ≈ 10 000 цыплят. Необходимо определить генеральные параметры μ и σ, причем μ с предельной относительной погрешностью 1 %. Для оценки μ с допустимой погрешностью необходимо рассчитать минимальный объем выборки $n_{\min}$.

Определение массы цыплят – операция нетрудоемкая, поэтому примем объем исходной выборки $n_1 = 12$. Экспериментальные данные и результаты предварительной обработки сведены в таблицу 3.

Обработка данных исходной выборки, $n_1 = 12$ (см. пункты алгоритма 1.4.2).

1. Выборочное среднее (п. 1): $\overline{m_1} = \dfrac{\sum_{i=1}^{n_1} m_{1i}}{n_1} = \dfrac{\sum_{i=1}^{12} m_{1i}}{12} =$

$$= \frac{565+860+882+893+893+902+909+909+923+937+945+1215}{12} = 902.8\,.$$

2. Выборочное стандартное отклонение (п. 2):

$$S_1^2 = \frac{\sum_{i=1}^{n_1}(m_{1i}-\overline{m_1})^2}{n_1-1} = \frac{\sum_{i=1}^{12}(m_{1i}-\overline{m_1})^2}{12-1} =$$

$$= \frac{(565-902.8)^2+(860-902.8)^2+(882-902.8)^2+(893-902.8)^2+(893-902.8)^2}{11} +$$

$$+ \frac{(902-902.8)^2+(909-902.8)^2+(909-902.8)^2+(923-902.8)^2+(937-902.8)^2}{11} +$$

$$+ \frac{(945-902.8)^2+(1215-902.8)^2}{11} = \frac{217470}{11} = 19770\,,$$

$$S_1 = \sqrt{S_1^2} = \sqrt{19770} = 140.6\,.$$

3. Проверка на промахи (п. 3):

$$\tau_{\text{э}1} = \frac{\max\left|m_{1i}-\overline{m_1}\right|}{S_1} = \frac{|565-902.8|}{140.6} = 2.403\,, \quad f_1 = n_1 - 2 = 10,$$

так как $m_{1i} = m_{1,1} = 565$ образует с $\overline{m_1} = 902.8$ максимальную разность по модулю.

4. Критическое значение критерия Смирнова–Граббса $\tau_{10,\,0.95} = 2.387$ (приложение 1).

5. В выборке m_{1i} значение $m_{1,1} = 565$ – промах, так как для него $\tau_{\text{э}1} = 2.403 > \tau_{10,\,0.95} = 2.387$. После исключения выявленного промаха из исходной выборки оставшиеся $n_2 = 11$ значений повторно проверяются на наличие в них промахов.

6. Выборочное среднее (п. 1): $\overline{m_2} = \frac{\sum_{i=1}^{n_2} m_{2i}}{n_2} = \frac{\sum_{i=1}^{11} m_{2i}}{11} =$

$$= \frac{860+882+893+893+902+909+909+923+937+945+1215}{11} = 933.5\,.$$

7. Выборочное стандартное отклонение (п. 2):

$$S_2^2 = \frac{\sum_{i=1}^{n_2}(m_{2i}-\overline{m_2})^2}{n_2-1} = \frac{\sum_{i=1}^{11}(m_{2i}-\overline{m_2})^2}{11-1} =$$

$$= \frac{(860-933.5)^2+(882-933.5)^2+(893-933.5)^2+(893-933.5)^2}{10} +$$

$$+ \frac{(902-933.5)^2+(909-933.5)^2+(909-933.5)^2+(923-933.5)^2}{10} +$$

$$+ \frac{(937-933.5)^2+(945-933.5)^2+(1215-933.5)^2}{10} = \frac{93025}{10} = 9302.5\,,$$

$$S_2 = \sqrt{S_2^2} = \sqrt{9302.5} = 96.4\,.$$

8. Проверка на промахи (п. 3):

$$\tau_{э2} = \frac{\max\left|m_{2i}-\overline{m_2}\right|}{S_2} = \frac{\left|1215-933.4\right|}{96.4} = 2.921\,,\ f_2 = n_2 - 2 = 9,$$

так как $m_{2i} = m_{2,\,11} = 1215$ образует с $\overline{m_2} = 933.5$ максимальную разность по модулю.

9. Критическое значение критерия Смирнова–Граббса $\tau_{9,\,0.95} = 2.343$ (приложение 1).

10. В выборке m_{2i} значение $m_{2,\,11} = 1215$ – промах, так как для него $\tau_{э2} = 2.921 > \tau_{9,\,0.95} = 2.343$. После исключения промаха из промежуточной выборки оставшиеся $n_3 = 10$ значений повторно проверяются на наличие в них промахов.

Таблица 3 – Экспериментальные данные и результаты предварительной обработки

Исходная выборка $n_1 = 12$				Промежуточная выборка $n_2 = 11$				Конечная выборка $n_3 = 10$				
n_{1i}	Масса цыплят	Квадрат отклонения от среднего	τ-критерий	n_{2i}	Масса цыплят	Квадрат отклонения от среднего	τ-критерий	n_{3i}	Масса цыплят	Квадрат отклонения от среднего	τ-критерий	$\lvert m_{3i} - \overline{m}_3 \rvert$
	m_{1i}, г	$(m_{1i} - \overline{m}_1)^2$	τ_{1i}		m_{2i}, г	$(m_{2i} - \overline{m}_2)^2$	τ_{2i}		m_{3i}, г	$(m_{1i} - \overline{m}_3)^2$	τ_{3i}	
1	565	114109	2.403		—	—	—		—	—	—	—
2	860	1832	0.304	1	860	5388	0.761	1	860	2052	1.776	45.3
3	882	433	0.148	2	882	2642	0.533	2	882	543	0.914	23.3
4	893	96	0.070	3	893	1632	0.419	3	893	151	0.482	12.3
5	893	96	0.070	4	893	1632	0.419	4	893	151	0.482	12.3
6	902	1	0.006	5	902	986	0.326	5	902	11	0.129	3.3
7	909	38	0.044	6	909	595	0.253	6	909	14	0.145	3.7
8	909	38	0.044	7	909	595	0.253	7	909	14	0.145	3.7
9	923	408	0.144	8	923	108	0.108	8	923	313	0.694	17.7
10	937	1170	0.243	9	937	13	0.037	9	937	1005	1.243	31.7
11	945	1781	0.300	10	945	135	0.120	10	945	1576	1.557	39.7
12	1215	97469	2.220	11	1215	79299	2.921		—	—	—	—
	$\overline{m}_1$	S_1^2	$\tau_{11,\,0.95}$		$\overline{m}_2$	S_2^2	$\tau_{10,\,0.95}$		$\overline{m}_3$	S_3^2	$\tau_{9,\,0.95}$	САО
	902.8	19770	2.387		933.5	9302.5	2.343		905.3	647.8	2.294	19.3
		S_1	$m_{1,\,1}$ промах			S_2	$m_{2,\,11}$ промах			S_3	Нет промахов	
		140.6				96.4				25.5		

Обработка данных конечной выборки, n_3 = 10.

11. Выборочное среднее (п. 1): $\overline{m_3} = \dfrac{\sum_{i=1}^{n_3} m_{3i}}{n_3} = \dfrac{\sum_{i=1}^{10} m_{3i}}{10} =$

$$= \frac{860+882+893+893+902+909+909+923+937+945}{10} = \frac{9053}{10} = 905.3 .$$

2. Выборочное стандартное отклонение (п. 2):

$$S_1^2 = \frac{\sum_{i=1}^{n_3}(m_{1i}-\overline{m_3})^2}{n_3-1} = \frac{\sum_{i=1}^{10}(m_{1i}-\overline{m_3})^2}{10-1} =$$

$$= \frac{(860-905.3)^2+(882-905.3)^2+(893-905.3)^2+(893-905.3)^2+(902-905.3)^2}{9} +$$

$$= \frac{(909-905.3)^2+(909-905.3)^2+(923-905.3)^2+(937-905.3)^2+(945-905.3)^2}{9} =$$

$$= \frac{5830}{9} = 647.8 ,$$

$$S_3 = \sqrt{S_3^2} = \sqrt{647.8} = 25.5 .$$

13. Проверка на промахи (п. 3):

$$\tau_{э3} = \frac{\max\left|m_{3i}-\overline{m_3}\right|}{S_3} = \frac{\left|860-905.3\right|}{25.5} = 1.776 , \quad f_3 = n_3 - 1 = 8,$$

так как $m_{3i} = m_{3,\,1} = 860$ образует с $\overline{m_3} = 905.3$ максимальную разность по модулю.

14. Критическое значение критерия Смирнова–Граббса $\tau_{8,\,0.95} = 2.294$ (приложение 1).

15. В выборке m_{3i} промахов нет, так как даже для $m_{3,\,1} = 860$ $\tau_{э3} = 1.776$ меньше критического значения $\tau_{8,\,0.95} = 2.294$.

16. Проверка выборки m_{3i} на принадлежность ее значений к нормальному закону распределению (п. 4). Случайные значения выборки m_{3i} принадлежат нормальному закону распределения, так как

$$\left|\frac{\text{САО}_3}{S_3} - 0.7979\right| < \frac{0.4}{\sqrt{n_3}} \rightarrow \left|\frac{19.3}{25.5} - 0.7979\right| < \frac{0.4}{\sqrt{10}} \rightarrow 0.0410 < 0.126,$$

$$\text{САО} = \frac{\sum_{i=1}^{n_3}\left|m_{3j}-\overline{m_3}\right|}{10} = \frac{\sum_{i=1}^{10}\left|m_{3j}-\overline{m_3}\right|}{10} =$$

$$= \frac{|860-905.3|+|882-905.3|+|893-905.3|+|893-905.3|+|902-905.3|}{10} +$$

$$+ \frac{|909-905.3|+|909-905.3|+|923-905.3|+|937-905.3|+|945-905.3|}{10} =$$

$$= \frac{193}{10} = 19.3 .$$

17. Доверительный интервал генерального среднего μ (п. 5):

$$\Delta\overline{m_3} = t_{n_3-1,\,0.95}\frac{S_3}{\sqrt{n_3}} = t_{9,\,0.95}\frac{S_3}{\sqrt{n_3}} = 2.262\cdot\frac{25.5}{\sqrt{10}} = 18.2 \approx 18$$

887 г < μ < 923 г.

где $t_{9,\,0.95} = 2.262$ – критическое значение критерия Стьюдента (приложение 2). Правила округления см. п. 3 алгоритма 1.3.2.

18. Доверительный интервал генеральной дисперсии и генерального стандартного отклонения (п. 6):

$$\frac{(n_3-1)S_3^2}{\chi^2_{n_3-1,\,(1-p)/2}} < \sigma_3^2 < \frac{(n_3-1)S_3^2}{\chi^2_{n_3-1,\,(1+p)/2}} \rightarrow \frac{9\cdot 647.8}{19.023} < \sigma_3^2 < \frac{9\cdot 647.8}{2.700} \rightarrow$$

$$\rightarrow 306.5 < \sigma_3^2 < 2159.3 \rightarrow 17.5 < \sigma_3 < 46.5 ,$$

где $\chi^2_{9,\,0.975} = 2.700$, $\chi^2_{9,\,0.025} = 19.023$ – критические значения критерия Пирсона (приложение 3).

19. Относительная погрешность определения генерального среднего μ по данным «очищенной» конечной выборки равна

$$\frac{\Delta \overline{Y}_3}{\overline{Y}_3} = \frac{18.2}{905.3} = 0.020 \ \ (2\,\%)\,.$$

20. Расчет $n_{\min}$ (п. 7). По условию задачи допустимая относительная погрешность генерального среднего 1 %. Потому предельная абсолютная погрешность $\Delta \overline{Y}_\text{o} = 0.01 \cdot \overline{Y}_3$ = 0.01 · 905.3 = 9.05.

$$\frac{t_{n_{\min}-1}}{\sqrt{n_{\min}}} = \frac{\Delta \overline{Y}_3}{S_3} = \frac{9.05}{25.5} = 0.355\,.$$

Из приложения 2 находим, что $n_{\min}$ = 33.

Резюме. Для оценки μ с предельной относительной погрешностью 1 % объем выборки должен быть равен 33. Учитывая, что 10 измерений уже сделано, следует определить массу еще 23 цыплят (при условии, что среди них не будет промахов). Допустим, что по 33 измерениям получим $\overline{m_4}$ = 906.2 г. Тогда масса 10 000 цыплят равна M = 906.2 · 10 000 = 9 062 000 г = 9.062 т с абсолютной погрешностью ± 0.09062 т (1 %) или M = 9.06 ± 0.09 т (см. правила округления, п. 3 алгоритма 1.3.2). Зная массу всех цыплят, можно спланировать, например, количество тары и автомобилей для транспортировки, кроме того можно рассчитать экономические показатели, например, прибыль.

1.4.4. Материалы для самостоятельной работы, лабораторных и практических занятий

Тесты

Укажите правильные формулы.

1. Выборочное среднее:

$$\text{а) } \overline{Y} = \frac{\sum_{i=1}^{n} Y_i}{n-1}; \quad \text{б) } \overline{Y} = \frac{Y_1 + \ldots + Y_n}{n}; \quad \text{в) } \overline{Y} = \frac{\sum_{i=1}^{n} Y_i}{n}; \quad \text{г) } \overline{Y} = \frac{\sum_{i=1}^{n} |Y_i|}{n}\,.$$

2. Выборочная дисперсия:

$$\text{а) } S^2 = \frac{\sum_{i=1}^{n} (Y_i - \overline{Y})^2}{n}; \qquad \text{б) } S^2 = \frac{\sum_{i=1}^{n} (Y_i - \overline{Y})^2}{n-1};$$

$$\text{в) } S^2 = \frac{\sum_{i=1}^{n} Y_i^2}{n}; \qquad \text{г) } S^2 = \frac{\sum_{i=1}^{n} \left|Y_i - \overline{Y}\right|}{n-1}\,.$$

3. Критерий Смирнова–Граббса:

$$\text{а) } \frac{\left|Y_{\min} - Y_{\max}\right|}{S} < \tau_{n-1,\,0.95}; \qquad \text{б) } \frac{\left(Y_{\max} - \overline{Y}\right)}{S} < \tau_{n-2,\,0.95};$$

$$\text{в) } \frac{\max\left|Y_i - \overline{Y}\right|}{S} < \tau_{n-2,\,0.95}; \qquad \text{г) } \frac{\max\left|Y_i - \overline{Y}\right|}{S} < \tau_{n-1,\,0.95}\,.$$

4. Критерий принадлежности случайных величин выборки к $N(Y, \mu, \sigma)$:

$$\text{а) } \left|\frac{\sum_{i=1}^{n}\left|Y_i - \overline{Y}\right|}{nS} - 0.7979\right| < \frac{2}{5\sqrt{n}}; \quad \text{б) } \left|\frac{\sum_{i=1}^{n}\left|Y_i - \overline{Y}\right|}{S^2} - 0.7979\right| < \frac{0.4}{\sqrt{n}};$$

$$\text{в) } \left|\frac{\sum_{i=1}^{n}(Y_i - \overline{Y})^2}{S} - 0.7979\right| < \frac{0.4}{\sqrt{n}}; \quad \text{г) } \left|\frac{\sum_{i=1}^{n}\left|Y_i - \overline{Y}\right|}{nS} - 0.7979\right| < 0.4 \cdot \sqrt{n}\,.$$

5. Доверительный интервал генерального среднего $\overline{Y} - \Delta\overline{Y} < \mu < \overline{Y} + \Delta\overline{Y}$:

$$\text{а) } \Delta\overline{Y} = t_{n-1,\,0.95} \cdot S(\overline{Y}); \qquad \text{б) } \Delta\overline{Y} = \frac{t_{n-1,\,0.95} \cdot S(Y)}{\sqrt{n}};$$

$$\text{в) } \Delta\overline{Y} = \frac{t_{n-1,\,0.95} \cdot S(\overline{Y})}{\sqrt{n}}; \qquad \text{г) } \Delta\overline{Y} = \frac{t_{n,\,0.95} \cdot S}{\sqrt{n}}\,.$$

6. Доверительный интервал генерального стандартного отклонения:

а) $\sqrt{\dfrac{(n-1)}{\chi^2_{n-1,\,0.025}}} < \sigma < \sqrt{\dfrac{(n-1)}{\chi^2_{n-1,\,0.975}}}$; б) $\dfrac{S\sqrt{n-1}}{\sqrt{\chi^2_{n-1,\,0.025}}} < \sigma < \dfrac{S\sqrt{n-1}}{\sqrt{\chi^2_{n-1,\,0.975}}}$;

в) $\sqrt{\dfrac{n}{\chi^2_{n-1,\,0.025}}} < \dfrac{\sigma^2}{S^2} < \sqrt{\dfrac{n}{\chi^2_{n-1,\,0.975}}}$; г) $\dfrac{(n-1)}{\chi^2_{n-1,\,0.025}} < \dfrac{\sigma}{S} < \dfrac{(n-1)}{\chi^2_{n-1,\,0.975}}$.

7. Минимальный объем выборки для оценки генерального среднего с допустимой предельной абсолютной погрешностью ΔY_{o} можно найти из следующего выражения:

а) $\dfrac{t_{n_{\min}-1,\,0.95}}{n_{\min}} = \dfrac{\Delta Y_{\text{o}}}{S}$; б) $\dfrac{t_{n_{\min}-1,\,0.95}}{\sqrt{n_{\min}}} = \dfrac{\Delta Y_{\text{o}}}{S}$;

в) $\dfrac{t_{n_{\min}-1,\,0.95}}{\sqrt{n_{\min}}} = \dfrac{S}{\Delta Y_{\text{o}}}$; г) $\dfrac{t_{n_{\min}-1,\,0.95}}{\sqrt{n_{\min}-1}} = \dfrac{\Delta Y_{\text{o}}}{S}$.

8. Какая часть случайных величин, подчиняющихся нормальному закону распределения, попадает в диапазон $-2\sigma \le \delta \le 2\sigma$

а) 68.27 %; б) 95.44 %; в) 99.27 %; г) 95 %?

Вопросы для самопроверки

1. Какими параметрами однозначно определяется нормальный закон распределение случайных величин?
2. В чем различие между понятиями «генеральная совокупность» и «выборка»?
3. Напишите формулу для расчета выборочного среднего.
4. Напишите формулу для расчета выборочной дисперсии.
5. Сформулируйте алгоритм отсева промахов по критерию Смирнова–Граббса.
6. Сформулируйте критерий принадлежности случайных величин выборки к нормальному закону распределения.
7. Напишите формулу для расчета доверительного интервала генерального среднего.

8. Напишите формулу для расчета доверительного интервала генеральной дисперсии.
9. Сформулируйте алгоритм расчета минимального количества опытов, необходимых для определения генерального среднего с предельной абсолютной погрешностью не более ΔY_{o}.

Задачи для самостоятельного решения

Уровень 1

1. Дана выборка: 5.2, 5.0, 5.4, 3.1, 6.4, 5.8, 4.4, 4.6, 4.6, 6.1. Рассчитать выборочные среднее, дисперсию и стандартное отклонение.

Ответ: $\overline{Y} = 5.1$, $S^2 = 0.92$, $S = 0.96$.

2. Есть ли в выборке задачи 1 промахи? Чему равно критическое значение критерия Смирнова–Граббса для этой выборки при доверительной вероятности $p = 0.95$?

Ответ: В этой выборке промахов нет. Критерий $\tau_{9,\,0.95} = 2.294$.

3. Проверить, выполняется ли для случайных величин выборки, приведенной в задаче 1, закон нормального распределения при доверительной вероятности $p = 0.95$?

Ответ: Да, выполняется.

4. Рассчитать доверительный интервал генерального среднего μ по выборке задачи 1 при доверительной вероятности $p = 0.95$.

Ответ: $4.4 < \mu < 5.8$.

5. Рассчитать доверительный интервал генерального стандартного отклонения σ по выборке задачи 1 при доверительной вероятности $p = 0.95$.

Ответ: $0.7 < \sigma < 1.8$.

Уровень 2

6. Найти минимальный объем выборки для оценки среднего генеральной совокупности μ с погрешностью $\overline{Y}_{\text{o}} = 0.25$ при доверительной вероятности $p = 0.95$. Исходная выборка представлена в задаче 1.

Ответ: $n_{\min} = 60$.

7. Есть ли в выборке 5.2, 5.0, 5.4, 2.2, 6.4, 5.8, 4.4, 4.6, 4.6, 6.1. промахи при доверительной вероятности $p = 0.95$?

Ответ: Значение 2.2 является промахом.

8. Проверить, выполняется ли для случайных величин выборки 5.2, 5.0, 5.4, 2.9, 6.5, 6.4, 4.4, 4.6, 4.6, 7.1 закон нормального распределения при доверительной вероятности $p = 0.95$?

Ответ: Промахов нет, но закон нормального распределения при доверительной вероятности $p = 0.95$ не выполняется.

Уровень 3

9. Проверить, выполняется ли для случайных величин выборки 5.2, 5.0, 5.2, 1.6, 6.3, 6.4, 4.4, 5.0, 4.6, 7.1 закон нормального распределения при доверительной вероятности $p = 0.95$? Рассчитать выборочные среднее, дисперсию и стандартное отклонение, а также минимальный объем выборки для оценки генерального среднего μ с погрешностью $\Delta \overline{Y}_{\text{o}} = 0.25$ при доверительной вероятности $p = 0.95$.

Ответ: Рассматривать выполнение закона нормального распределения для заданной выборки некорректно, так как значение 1.6 является промахом. После его исключения значения «очищенной» выборки 5.2, 5.0, 5.2, 6.3, 6.4, 4.4, 5.0, 4.6, 7.1 подчиняются нормальному закону распределения. Для нее $\overline{Y} = 5.5$, $S^2 = 0.84$, $S = 0.92$, $4.8 < \mu < 6.2$; $0.6 < \sigma < 1.8$; $n_{\min} = 54$.

10. На станке-автомате изготовлена партия втулок 10 000 шт. с номинальным диаметром 68 мм. Контроль диаметра выполнен на 8 втулках: 67.85, 67.89, 67.98, 67.99, 68.02, 68.06, 68.12, 68.14. Для расчета рентабельности крупносерийного производства генеральное среднее μ требуется оценить с абсолютной погрешностью не более 0.025 мм. Определить минимальный объем выборки $n_{\min}$, обеспечивающий допустимую погрешность оценки μ.

Ответ: $n_{\min} = 67$.

1.5. Обработка данных стохастического эксперимента. Две выборки и более (тема 4)

1.5.1. Теория

В разделе 1.4 приведен математический аппарат предварительной обработки экспериментальных данных для одной выборки. На практике часто возникает необходимость оценить, принадлежат ли две выборки, полученные разными способами, к одной генеральной совокупности. Напоминаем, что здесь и в дальнейшем речь идет только об «очищенных» выборках, то есть таких, в которых нет промахов и случайные величины которых подчиняются нормальному закону распределения (см. раздел 1.4.1).

В общем случае можно считать, что данные двух выборок ($\overline{Y}_1$, S_1^2, n_1) и ($\overline{Y}_2$, S_2^2, n_2) относятся к одной генеральной совокупности, если $\mu_1 = \mu_2$ и $\sigma_1 = \sigma_2$. Проверка указанной гипотезы производится в два этапа. Сначала необходимо убедиться в однородности дисперсий обеих выборок, то есть в том, что $\sigma_1 = \sigma_2$.

Однородность выборочных дисперсий проверяется по критерию Фишера. Для этого рассчитывают экспериментальное значение критерия Фишера $F_э$, которое сравнивают с критическим значением:

$$F_э = \frac{\max(S_1^2, S_2^2)}{\min(S_1^2, S_2^2)}. \tag{48}$$

При $S_1^2 > S_2^2$ условие $\sigma_1 = \sigma_2$ выполняется в том случае, если

$$F_э = \frac{S_1^2}{S_2^2} < F_{n_1-1,\, n_2-1,\, p}, \tag{49}$$

где $F_{n_1-1,\, n_2-1,\, p}$ – критическое значение критерия Фишера (приложение 4).

При $S_1^2 < S_2^2$ условие $\sigma_1 = \sigma_2$ выполняется в том случае, если

$$F_э = \frac{S_2^2}{S_1^2} < F_{n_2-1,\, n_1-1,\, p}, \tag{50}$$

где $F_{n_2-1,\, n_1-1,\, p}$ – критическое значение критерия Фишера (приложение 4).

Если неравенства (49)–(50) не выполняются, тогда различие между выборочными дисперсиями S_1^2 и S_2^2 существенно: $\sigma_1^2 > \sigma_2^2$ при $S_1^2 > S_2^2$ и $\sigma_1^2 < \sigma_2^2$ при $S_1^2 < S_2^2$.

Если выборочные дисперсии однородны $\sigma_1^2 = \sigma_2^2$, то в этом случае (**и только в этом случае**) можно сравнивать выборочные средние $\overline{Y}_1$ и $\overline{Y}_2$. Выборочные средние существенно не различаются, то есть $\mu_1 = \mu_2$, если выполняется критерий Стьюдента:

$$t_э = \frac{\left|\overline{Y}_1 - \overline{Y}_2\right|}{\sqrt{(n_1-1)\cdot S_1^2 + (n_2-1)\cdot S_2^2}} \sqrt{\frac{n_1 n_2 (n_1+n_2-2)}{n_1+n_2}} < t_{n_1+n_2-2,\, p}, \qquad (51)$$

где $t_{n_1+n_2-2,\, p}$ – критическое значение критерия Стьюдента (приложение 2).

Если неравенство (51) выполняется, то различие между $\overline{Y}_1$ и $\overline{Y}_2$ несущественно и $\mu_1 = \mu_2$. Если неравенство (51) не выполняется, то различие между $\overline{Y}_1$ и $\overline{Y}_2$ существенно: $\mu_1 > \mu_2$ при $\overline{Y}_1 > \overline{Y}_2$ или $\mu_1 < \mu_2$ при $\overline{Y}_1 < \overline{Y}_2$. Еще раз подчеркнем, что сравнение выборочных средних возможно только при условии однородности дисперсий $\sigma_1 = \sigma_2$. Если $n_1 = n_2 = n$, то неравенство (51) приобретает более простой вид (сравни с уравнением (43)):

$$t_э = \frac{\left|\overline{Y}_1 - \overline{Y}_2\right|}{\sqrt{S_1^2 + S_2^2}} \sqrt{n} < t_{2n-2,\, p}. \qquad (52)$$

Если данные двух выборок ($\overline{Y}_1$, S_1^2, n_1) и ($\overline{Y}_2$, S_2^2, n_2) относятся к одной генеральной совокупности $\mu_1 = \mu_2$ и $\sigma_1 = \sigma_2$, то их данные можно объединить для дальнейшего анализа.

При математическом моделировании исследуемых объектов возникает необходимость сравнения дисперсий трех и более выборок. В случае, когда все выборки имеют одинаковый объем n, проверка дисперсий на однородность производится по критерию Кохрена. Проверка N дисперсий с n дублями в каждой выборке на однородность осуществляется в несколько этапов (таблица 4).

Таблица 4 – Массив экспериментальных данных

N	n					$\overline{Y}_j$	S_j^2
	Y_{j1}	…	Y_{j1}	…	Y_{j1}		
1	Y_{11}	…	Y_{1i}	…	Y_{1n}	$\overline{Y}_1$	S_1^2
…	…	…	…	…	…	…	…
j	Y_{j1}	…	Y_{ji}	…	Y_{jn}	$\overline{Y}_j$	S_j^2
…	…	…	…	…	…	…	…
N	Y_{N1}	…	Y_{Ni}	…	Y_{Nn}	$\overline{Y}_N$	S_N^2
							$\sum_{j=1}^{N} S_j^2$

1. Рассчитывают выборочное среднее $\overline{Y}_j$ в каждой выборке:

$$\overline{Y}_j = \frac{\sum_{i=1}^{n} Y_{ji}}{n}, \quad j = 1, \ldots, N. \qquad (53)$$

2. Рассчитывают выборочную дисперсию S_j^2 в каждой выборке:

$$S_j^2 = \frac{\sum_{i=1}^{n} (Y_{ji} - \overline{Y}_j)^2}{n-1}, \quad j = 1, \ldots, N, \quad f = n-1. \qquad (54)$$

3. Рассчитывают экспериментальное значение критерия Кохрена $G_э$:

$$G_э = \frac{\max(S_j^2)}{\sum_{j=1}^{N} S_j^2} \qquad (55)$$

4. Находят критическое значение критерия Кохрена $G_{n-1,\, N,\, p}$ (приложение 5).

5. Если $G_э < G_{n-1,\, N,\, p}$, то все выборочные дисперсии однородны $\sigma_1^2 = \ldots = \sigma_n^2$. Если $G_э > G_{n-1,\, N,\, p}$, то наибольшая выборочная дисперсия статистически значимо превосходит все остальные. Для дальнейшей корректной обработки экспериментальных данных необходимо повторить опыт, в котором обнаружено максимальное значение дисперсии.

Напоминаем, что построению математических моделей должно предшествовать «очистка» всех выборок: проверка их на промахи, на принадлежность нормальному закону распределения, на однородность дисперсий всех опытов.

1.5.2. Алгоритм решения прикладных задач

Даны две выборки и более, $j = 1, \ldots, N$. Объем каждой выборки n_j.

1. Рассчитать выборочное среднее в каждой выборке:

$$\overline{Y}_j = \frac{\sum_{i=1}^{n_j} Y_{ji}}{n_j}.$$

2. Рассчитать выборочную дисперсию и выборочное стандартное отклонение в каждой выборке:

$$S_j^2 = \frac{\sum_{i=1}^{n_j} (Y_{ji} - \overline{Y}_j)^2}{n_j - 1}, \quad f_j = n_j - 1,$$

$$S_j = \sqrt{S_j^2}, \quad f_j = n_j - 1.$$

3. Проверить значения всех выборок на наличие в них промахов по критерию Смирнова–Граббса.

3.1. Рассчитать экспериментальное значение критерия Смирнова–Граббса $\tau_{эj}$ для каждой выборки

$$\tau_{эj} = \frac{\max\left|Y_{ji} - \overline{Y}_j\right|}{S_j}, \quad f_j = n_j - 2.$$

3.2. Найти по таблице приложения 1 критическое значение критерия Смирнова–Граббса $\tau_{n_j-2,\,p}$ для каждой выборки;

3.3. Величина Y_{ji} является промахом, если

$$\tau_{эj} > \tau_{n_j-2,\,p}.$$

Из выборки исключается только **одно** случайное значение. После исключения одного промаха проверку оставшейся выборки на промахи повторить.

4. Проверить соответствие случайных величин каждой выборки закону нормального распределения

$$\left|\frac{\text{САО}_j}{S_j} - 0.7979\right| < \frac{0.4}{\sqrt{n_j}},$$

$$\text{САО}_j = \frac{\sum_{i=1}^{n_j}\left|Y_{ji} - \overline{Y}_j\right|}{n_j}.$$

Если выполняется приведенное выше неравенство, то случайные значения параметра Y_{ji} подчиняются нормальному закону распределения.

5. Две выборки ($\overline{Y}_1, S_1^2, n_1$) и ($\overline{Y}_2, S_2^2, n_2$) при $S_1^2 \neq S_2^2$ обладают однородными дисперсиями, то есть $\sigma_1^2 = \sigma_2^2$, если

$$\text{при } S_1^2 > S_2^2 \text{ имеем } F_э = \frac{S_1^2}{S_2^2} < F_{n_1-1,\,n_2-1,\,p},$$

$$\text{при } S_1^2 < S_2^2 \text{ имеем } F_э = \frac{S_2^2}{S_1^2} < F_{n_2-1,\,n_1-1,\,p},$$

где $F_{n_1-1,\,n_2-1,\,p}$, $F_{n_2-1,\,n_1-1,\,p}$ – критические значения критерия Фишера (приложение 4).

В противном случае:

$$\text{при } S_1^2 > S_2^2 \text{ и } F_э = \frac{S_1^2}{S_2^2} > F_{n_1-1,\,n_2-1,\,p} \;\rightarrow\; \sigma_1^2 > \sigma_2^2,$$

$$\text{при } S_1^2 < S_2^2 \text{ и } F_э = \frac{S_1^2}{S_2^2} > F_{n_1-1,\,n_2-1,\,p} \;\rightarrow\; \sigma_1^2 < \sigma_2^2.$$

6. Две выборки ($\overline{Y}_1, S_1^2, n_1$) и ($\overline{Y}_2, S_2^2, n_2$) с $\overline{Y}_1 \neq \overline{Y}_2$, но однородными дисперсиями, то есть $\sigma_1^2 = \sigma_2^2$, обладают одинаковыми генеральными средними, то есть $\mu_1 = \mu_2$, если

$$t_э = \frac{\left|\overline{Y}_1 - \overline{Y}_2\right|}{\sqrt{(n_1-1)\cdot S_1^2 + (n_2-1)\cdot S_2^2}} \sqrt{\frac{n_1 n_2 (n_1+n_2-2)}{n_1+n_2}} < t_{n_1+n_2-2,\, p},$$

где $t_{n_1+n_2-2,\, p}$ – критическое значение критерия Стьюдента (приложение 2).

7. Для трех и более выборок одинакового объема n ($\overline{Y}_1, S_1^2, n$), …, ($\overline{Y}_j, S_j^2, n$), …, ($\overline{Y}_N, S_N^2, n$) проверить дисперсии всех выборок на однородность по критерию Кохрена, то есть $\sigma_1^2 = ... = \sigma_j^2 = ... = \sigma_N^2$ (приложение 5):

$$G_э = \frac{\max(S_j^2)}{\sum\limits_{j=1}^{N} S_j^2} < G_{n-1,\, N,\, p}.$$

Предварительная обработка экспериментальных данных каждой выборки (нахождение выборочного среднего и выборочной дисперсии, отсев промахов, проверка значений выборки на принадлежность к нормальному закону распределения, проверка всех выборочных дисперсий на однородность) должна предшествовать построению математических моделей методами математической статистики.

1.5.3. Базовая задача

Идет уборка урожая пшеницы комбайном, который движется со скоростью 15 км/ч. Бригадир чувствует, что в условиях дождливой погоды, при которой происходит уборка урожая, снижение скорости комбайна до 12 км/ч позволит уменьшить потери зерна. Для экспериментальной проверки этой гипотезы часть поля разбили на 16 равных делянок. Данные по урожайности пшеницы на 8 делянках Y_1 (выборка $n_1 = 8$) получены при скорости движения комбайна 15 км/ч, данные по урожайности пшеницы на других 8 делянках Y_2 (выборка $n_2 = 8$) при 12 км/ч. Результаты по сбору зерна с делянок при разных скоростях комбайна с соответствующей предварительной обработкой представлены в таблице 5 (ссылки на алгоритм 1.5.2).

Таблица 5 – Данные сбора пшеницы

N	Y_1, ц/га	Y_2, ц/га
1	37.4	40.2
2	39.6	38.0
3	37.4	39.9
4	36.9	38.0
5	38.9	38.3
6	36.7	40.8
7	38.5	39.6
8	32.6	38.2
$\overline{Y}$	37.25	39.13
S^2	4.557	1.265
S	2.135	1.125

На первый взгляд, бригадир оказался прав, так как при скорости комбайна 12 км/ч сбор зерна на 39.13 – 37.25 = 1.88 ≈ 1.9 ц/га больше, чем при скорости комбайна 15 км/ч (больше на 5 %).

Статистические расчеты подтвердили гипотезу бригадира. Выборочные дисперсии существенно не различаются, так как экспериментальный критерий Фишера $F_э$ меньше критического $F_{n_1-1,\, n_2-1,\, p}$ (п. 5 и приложение 4):

$$F_э = \frac{S_1^2}{S_1^2} = \frac{4.557}{1.265} = 3.602 < \mathrm{F}_{7,\,7,\,0.95} = 3.787.$$

Однородность выборочных дисперсий позволяет проверить гипотезу о существенном различии выборочных средних (п. 6 и приложение 2):

$$t_э = \frac{\left|\overline{Y}_1 - \overline{Y}_2\right|}{\sqrt{(n_1-1)\cdot S_1^2 + (n_2-1)\cdot S_2^2}} \sqrt{\frac{n_1 n_2 (n_1+n_2-2)}{n_1+n_2}} =$$

$$= \frac{(39.13-37.25)}{\sqrt{7\cdot 4.557 + 7\cdot 1.265}} \cdot \sqrt{\frac{8\cdot 8\cdot 14}{16}} = 2.204 > t_{14,\,0.95} = 2.145.$$

Так как экспериментальное значение критерия Стьюдента $t_э$ больше критического значения $t_{14,\,0.95}$, то выходит, что бригадир прав: снижение скорости комбайна с 15 до 12 км/ч уменьшает потери зерна при сборе урожая пшеницы, поскольку генеральное среднее μ_2 значимо больше генерального среднего μ_1.

Однако сделанный вывод является ошибочным, так как при обработке экспериментальных данных авторы **сознательно** (из методико-педагогических соображений) нарушили одно из основных требований теории эксперимента: проверять статистические гипотезы только по «очищенным» выборкам (отсев возможных промахов (п. 3), проверка выполнения закона нормального распределения (п. 4)). Несоблюдение этих требований может привести к ошибкам, что и случилось.

Проведем обработку экспериментальных данных заново. В таблице 6 приведены те же результаты, что и в таблице 5, но с проверкой на промахи и на принадлежность данных каждой выборки к нормальному закону распределения.

Таблица 6 – Данные сбора пшеницы

N	Y_1, ц/га	τ_1	Y_1,* ц/га	τ_1*	Y_2, ц/га	τ_2
1	37.4	0.070	37.4	0.466	40.2	0.951
2	39.6	1.101	39.6	1.543	38.0	1.004
3	37.4	0.070	37.4	0.466	39.9	0.684
4	36.9	0.164	36.9	0.922	38.0	1.004
5	38.9	0.773	38.9	0.904	38.3	0.738
6	36.7	0.258	36.7	1.105	40.8	1.484
7	38.5	0.586	38.5	0.537	39.6	0.418
8	**32.6**	**2.178**	–	–	38.2	0.827
	$\overline{Y_1}$	37.25	$\overline{Y_1^*}$	37.91	$\overline{Y_2}$	39.13
	S_1^2	4.557	S_1^{*2}	1.20	S_2^2	1.27
	S_1	2.135	S_1^*	1.10	S_2	1.13
	$\tau_{7,\,0.95}$	**2.172**	$\tau_{6,\,0.95}$	**2.093**	$\tau_{7,\,0.95}$	**2.172**
			CAO_1*	0.849	CAO_2	0.889
			$F_э$	1.056	$F_{7,\,6,\,0.95}$	4.207
			$t_э$	2.122	$t_{13,\,0.95}$	2.160

Проверка на промахи данных выборки Y_1 ($n_1 = 8$) показала (таблица 6), что значение $Y_{1,\,8} = 32.6$ является промахом (п. 3 и приложение 1), так как:

$$\tau_{э1} = \frac{\max\left|Y_{1i} - \overline{Y}_1\right|}{S_1} = \frac{|32.6 - 37.25|}{2.135} = \frac{4.65}{2.135} = 2.178 > \tau_{6,\,0.95} = 2.172 .$$

После исключения промаха из выборки Y_1, очищенная выборка Y_1^* ($n_1^* = 7$) промахов не имеет, так как критерий Смирнова–Граббса для всех ее значений меньше критического $\tau_{5,\,0.95} = 2.093$ (приложение 1). Выборка Y_2 ($n_2 = 8$) изначально не имеет промахов, так как критерий Смирнова–Граббса для всех ее значений меньше критического $\tau_{6,\,0.95} = 2.172$ (приложение 1). Проверка случайных значений обеих выборок Y_1^* и Y_2 показала, что они обе подчиняются нормальному закону распределения (п. 4):

$$\left|\frac{CAO_1^*}{S_1^*} - 0.7979\right| = \left|\frac{0.849}{1.095} - 0.7979\right| = 0.0226 < \frac{0.4}{\sqrt{n_1^*}} = \frac{0.4}{\sqrt{7}} = 0.151 ,$$

$$\left|\frac{CAO_2}{S_2} - 0.7979\right| = \left|\frac{0.889}{1.125} - 0.79797\right| = 0.0077 < \frac{0.4}{\sqrt{n_2}} = \frac{0.4}{\sqrt{8}} = 0.141 ,$$

$$CAO_1^* = \frac{\sum_{i=1}^{7}\left|Y_{1i} - \overline{Y_1^*}\right|}{7} = \frac{|37.4 - 37.91| + |39.6 - 37.91| + |37.4 - 37.91|}{7} +$$

$$+ \frac{|36.9 - 37.91| + |38.9 - 37.91| + |36.7 - 37.91| + |38.5 - 37.91|}{7} = 0.849 ,$$

$$CAO_2 = \frac{\sum_{i=1}^{8}\left|Y_{2i} - \overline{Y_2}\right|}{8} = 0.889 \text{ (проверить самостоятельно).}$$

Только после того как убедились, что обе выборки Y_1^* ($n_1^* = 7$) и Y_2 ($n_2 = 8$) не имеют промахов и подчиняются нормальному закону распределения, можно приступить к сравнению их выборочных дисперсий и средних. Выборочные дисперсии S_1^{*2} и S_2^2, имеющие степени свободы $f_1^* = 6$ и $f_2 = 7$, различаются несущественно (п. 5 и приложение 4), так как

$$F_э = \frac{S_2^2}{S_1^{*2}} = \frac{1.265}{1.198} = 1.056 < F_{7,\,6,\,0{,}95} = 4.207 .$$

Поскольку $\sigma_1^* = \sigma_2$, то можно сравнить выборочные средние $\overline{Y_1^*}, \overline{Y_2}$ (п. 6 и приложение 2):

$$t_э = \frac{\left|\overline{Y_1^*} - \overline{Y_2}\right|}{\sqrt{(n_1^* - 1)\cdot S_1^{*2} + (n_2 - 1)\cdot S_2^2}} \sqrt{\frac{n_1^* n_2 (n_1^* + n_2 - 2)}{n_1^* + n_2}} =$$

$$= \frac{|37.91 - 39.13|}{\sqrt{6\cdot 1.198 + 7\cdot 1.265}} \cdot \sqrt{\frac{7\cdot 8\cdot 13}{15}} = 2.122 < t_{13,\,0.95} = 2.160\,.$$

После сравнения «очищенных» выборок оказалось, что различие в выборочных средних случайно (несущественно), то есть $\mu_2 = \mu_1$. Поэтому нет никакого смысла снижать скорость комбайна, так как это приведет к уменьшению производительности при уборке урожая, а потери зерна все равно останутся такими же (интуитивное предположение бригадира не подтвердилось).

1.5.4. Материалы для самостоятельной работы, лабораторных и практических занятий

Тесты

1. Различие двух выборочных дисперсий несущественно, то есть $\sigma_1^2 = \sigma_2^2$, при $S_1^2 > S_2^2$, если

а) $\dfrac{S_2^2}{S_1^2} < F_{n_2-1,\, n_1-1,\, 0.95}$; б) $\dfrac{S_1^2}{S_2^2} < F_{n_1-1,\, n_2-1,\, 0.95}$;

в) $\dfrac{S_1}{S_2} < F_{n_1-1,\, n_2-1,\, 0.95}$; г) $\dfrac{S_1^2}{S_2^2} < F_{n_1,\, n_2,\, 0.95}$.

2. Различие двух выборочных средних, для которых $\sigma_1^2 = \sigma_2^2$, несущественно, то есть $\mu_1 = \mu_2$ при $\overline{Y}_1 \neq \overline{Y}_2$, если

а) $\dfrac{\left|\overline{Y}_1 - \overline{Y}_2\right|}{\sqrt{n_1 \cdot S_1^2 + n_2 \cdot S_2^2}} \cdot \sqrt{\dfrac{n_1 \cdot n_2 \cdot (n_1 + n_2 - 2)}{n_1 + n_2}} < t_{n_1+n_2-2,\, p}$;

б) $\dfrac{\left|\overline{Y}_1 - \overline{Y}_2\right|}{\sqrt{(n_1 - 1)\cdot S_1^2 + (n_2 - 1)\cdot S_2^2}} \cdot \sqrt{\dfrac{n_1 \cdot n_2 \cdot (n_1 + n_2 - 2)}{n_1 + n_2}} < t_{n_1+n_2-2,\, p}$;

в) $\dfrac{\left|\overline{Y}_1 - \overline{Y}_2\right|}{\sqrt{(n_1 - 1)\cdot S_1 + (n_2 - 1)\cdot S_2}} \cdot \sqrt{\dfrac{n_1 \cdot n_2 \cdot (n_1 + n_2 - 2)}{n_1 + n_2}} < t_{n_1+n_2-2,\, p}$;

г) $\dfrac{\left|\overline{Y}_1 - \overline{Y}_2\right|}{\sqrt{(n_1 - 1)\cdot S_1^2 + (n_2 - 1)\cdot S_2^2}} \cdot \sqrt{\dfrac{n_1 \cdot n_2 \cdot (n_1 + n_2 - 2)}{n_1 + n_2}} < t_{n_1+n_2,\, p}$.

3. Выборочные дисперсии N выборок $S_1^2, S_2^2, ..., S_N^2$ однородны (объем каждой выборки равен n), то есть $\sigma_1^2 = ... = \sigma_j^2 = ... = \sigma_N^2$, если

а) $\dfrac{\min (S_j^2)}{\sum_{j=1}^{N} S_j^2} < G_{n-1,\, N,\, p}$; б) $\dfrac{\max (S_j^2)}{\sum_{j=1}^{N} S_j^2} < G_{n-1,\, N,\, p}$;

в) $\dfrac{\max (S_j^2)}{\sum_{j=1}^{N} S_j^2} < G_{N,\, n-1,\, p}$; г) $\dfrac{\sum_{j=1}^{N} S_j^2}{\max (S_j^2)} < G_{n-1,\, N,\, p}$.

Вопросы для самопроверки

1. Сформулируйте алгоритм определения существенного различия между выборочными дисперсиями двух выборок. Как найти критического значения критерия Фишера?
2. Нужен ли одинаковый объем выборок $n_1 = n_2$ для сравнения двух выборочных дисперсий?
3. Приведите формулу для определения существенного различия между двумя выборочными средними. Как найти критическое значение критерия Стьюдента?
4. Сформулируйте алгоритм проверки выборочных дисперсий на однородность для трех и более выборок. Как найти критическое значение критерия Кохрена?

5. Нужен ли одинаковый объем всех выборок при сравнении выборочных дисперсий по критерию Кохрену?

Задачи для самостоятельного решения

Уровень 1

1. Для сравнения точности двух станков произвели измерения отклонений диаметров втулок от номинального размера для двух партий, обработанных на этих станках: $n_1 = 25$, $S_1^2 = 65.68$ мкм^2; $n_2 = 30$, $S_2^2 = 30.36$ мкм^2. Можно ли утверждать, что точность обработки втулок (разброс размеров), изготовленных на втором станке, выше, чем на первом?

Ответ: Второй станок обеспечивает более высокую точность обработки втулок (меньший разброс по диаметру), чем первый.

2. Для сравнения двух крупных партий жемчуга по качеству (чем меньше разброс по диаметру жемчужин, тем выше качество и стоимость) произвели измерения отклонений диаметров от номинального размера в двух выборках: $n_1 = 18$, $S_1^2 = 0.015$ мм^2; $n_2 = 22$, $S_2^2 = 0.025$ мм^2. Можно ли утверждать, что качество обеих партий жемчуга одинаково?

Ответ: Качество жемчуга в обеих партиях одинаковое $\sigma_1 = \sigma_2$.

Уровень 2

3. При производстве кирпичей возникло подозрение, что кирпичи, обожженные в двух разных печах, различаются по плотности. Выборочный контроль дал следующие результаты: $n_1 = 14$, $\overline{Y}_1 = 2.43$ г/см^3, $S_1^2 = 16.4$ (г/см^3)^2; $n_2 = 10$, $\overline{Y}_2 = 4.90$ г/см^3, $S_2^2 = 22.5$ (г/см^3)^2. Существенно ли различие кирпичей по плотности?

Ответ: Статистическая обработка экспериментальных данных показала, что $\sigma_1 = \sigma_2$ и $\mu_1 = \mu_2$. Наблюдаемое различие по плотности случайно.

4. Выполнены 6 опытов ($N = 6$) каждый с 5 дублями ($n = 5$). Можно считать, что дисперсии во всех опытах однородны?

N	Y_{j1}	Y_{j2}	Y_{j3}	Y_{j4}	Y_{j5}
1	5.0	4.6	6.3	4.3	3.6
2	4.7	5.4	6.2	7.2	6.2
3	7.2	5.5	6.5	7.5	6.6
4	9.1	8.5	9.1	7.6	9.7
5	10.6	10.6	11.4	8.8	8.9
6	10.5	9.3	9.3	10.8	8.0

Ответ: Да, дисперсии всех шести опытов однородны.

5. Выполнены 6 опытов ($N = 6$) каждый с 5 дублями ($n = 5$). Можно ли считать, что дисперсии во всех опытах однородны?

N	Y_{j1}	Y_{j2}	Y_{j3}	Y_{j4}	Y_{j5}
1	4.0	5.5	5.0	7.1	3.8
2	6.9	6.9	6.4	6.7	6.5
3	6.2	7.9	7.7	7.9	7.1
4	8.2	9.0	7.5	9.1	8.5
5	10.0	9.3	9.9	9.2	9.9
6	10.1	11.4	9.7	10.1	11.4

Ответ: Нет, дисперсии неоднородны.

6. Для обработки деталей применялись сверла, изготовленные из инструментальных сталей различных марок. Работоспособность сверл оценивалась временем их работы в минутах до переточки. Получены следующие результаты: $\overline{Y}_1 = 26.7$ мин, $S_1^2 = 48.2$ мин^2, $n_1 = 14$; $\overline{Y}_2 = 15.0$ мин, $S_2^2 = 53.3$ мин^2, $n_2 = 16$. Можно ли считать различие выборочных средних значений случайным или это различие столь существенно, что можно констатировать бóльшую стойкость сверл первой партии, нежели второй?

Ответ: Значимое различие в качестве изготовления сверл обеих партий отсутствует (генеральные дисперсии $\sigma_1 = \sigma_2$), а различие в значениях средней стойкости существенно $\mu_1 > \mu_2$.

Уровень 3

6. Даны две выборки: 17.4, 16.3, 17.1, 15.2, 14.1, 14.8, 14.5, 12.9, 17.6, 17.4, 16.6, 16.3 и 14.6, 16.2, 16.2, 17.9, 15.4, 16.6, 18.3, 18.4, 19.5, 16.2, 15.3, 17.7, 21.2, 15.9. Найти выборочные параметры обеих выборок $\overline{Y}$ и S^2 и сделать вывод об аналогичных генеральных параметрах.

Ответ: В обеих выборках промахов нет, и обе они подчиняются нормальному закону распределения. Параметры выборок: $\overline{Y_1} = 15.85$, $S_1^2 = 2.32$, $n_1 = 12$; $\overline{Y_2} = 17.10$, $S_1^2 = 3.32$, $n_2 = 14$. Несмотря на то, что выборочные параметры различаются, генеральные параметры существенно не различаются: $\mu_1 = \mu_2$, $\sigma_1^2 = \sigma_2^2$.

7. Даны две выборки: 17.2, 15.1, 17.2, 16.3, 14.2, 16.8, 16.5, 16.0, 15.2, 11.0, 12.9, 16.9 и 21.9, 20.8, 19.1, 18.0, 17.0, 19.1, 19.9, 20.0, 17.4, 20.3, 16.3, 19.4, 20.4, 18.2. Найти выборочные параметры обеих выборок $\overline{Y}$ и S^2 и сделать вывод о генеральных параметрах.

Ответ: В обеих выборках промахов нет, и они подчиняются нормальному закону распределения. Параметры выборок: $\overline{Y_1} = 15.44$, $S_1^2 = 3.66$, $n_1 = 12$; $\overline{Y_2} = 19.13$, $S_1^2 = 2.50$, $n_2 = 14$. Выборочные дисперсии однородны: $\sigma_1^2 = \sigma_2^2$. Выборочные средние различаются существенно $\mu_1 < \mu_2$.

8. Даны две выборки: 17.4, 16.3, 17.1, 15.2, 14.1, 14.8, 14.5, 12.9, 17.6, 17.4, 16.6, 16.3 и 12.7, 18.8, 12.3, 15.3, 14.1, 13.3, 13.1, 9.8, 17.7, 14.8, 19.2, 12.7, 19.3, 10.9. Найти выборочные параметры обеих выборок $\overline{Y}$ и S^2 и сделать вывод об аналогичных генеральных параметрах.

Ответ: В обеих выборках промахов нет, и обе они подчиняются нормальному закону распределения. Параметры выборок: $\overline{Y_1} = 15.85$, $S_1^2 = 2.32$, $n_1 = 12$; $\overline{Y_2} = 14.57$, $S_1^2 = 9.57$, $n_2 = 14$. Различие выборочных дисперсий существенно $\sigma_1^2 < \sigma_2^2$. Выборочные средние сравнивать с помощью приведенных в алгоритме методов нельзя.

Модуль 2
ОДНОФАКТОРНЫЙ ЭКСПЕРИМЕНТ

В результате изучения модуля студент должен:
– **уметь** планировать однофакторный эксперимент;
– **уметь** обрабатывать экспериментальные данные однофакторного эксперимента методами математической статистики и моделировать исследуемый объект уравнениями регрессии первого и второго порядка;
– **уметь** проверять полученные модели на статистическую значимость (значимость коэффициентов регрессии, адекватность уравнения регрессии);
– **находить** оптимальные значения фактора на основе полученного уравнения регрессии второго порядка и оценивать погрешность рассчитываемого параметра;
– **приобрести** навыки самостоятельной работы.

2.1. Перечень и содержание лекций

Лекция 1. Параметр. Требования к параметру. Дублирование измерений. Фактор. Требования к фактору. Натуральные и нормированные значения факторов. Формулы перехода от натуральных значений фактора к нормированным и обратно. Уравнения регрессии *m* – порядка. Метод наименьших квадратов (МНК) – универсальный способ расчета коэффициентов уравнения регрессии. Преимущества моделирования исследуемых объектов в нормированных значениях факторов. Симметричность и ортогональность однофакторных планов. Равномерный симметричный план (РСП). Примеры построения планов однофакторного эксперимента. Алгоритм построения планов однофакторного эксперимента.

Лекция 2. Ортогональность факторов. Диагонализация системы уравнений, полученной МНК, – необходимое условие для независимой оценки коэффициентов уравнения регрессии. Расчет суммы квадратов столбцов факторов для РСП. Элементы регрессионного анализа. Расчет коэффициентов уравнения регрессии первого порядка. Дисперсия воспроизводимости. Элементы дисперсионного анализа. Дисперсии значимости коэффициентов регрессии, доверительные интервалы и проверка коэффициентов уравнения регрессии первого порядка на значимость по критерию Стьюдента. Остаточная сумма квадратов. Дисперсия адекватности. Проверка уравнения регрессии первого порядка на адекватность по критерию Фишера. Точность прогнозирования по уравнению регрессии первого порядка. Принятие решения на основе уравнения регрессии первого порядка. Алгоритм решения прикладных задач с помощью уравнений регрессии первого порядка. Решение базовой задачи.

Лекция 3. Неадекватность уравнения регрессии первого порядка. Моделирование исследуемого объекта уравнением регрессии второго порядка. Ортогонализация квадратичного члена – необходимое условие диагонализации системы уравнений для расчета коэффициентов уравнения регрессии второго порядка, полученной МНК. Расчет величины суммы квадратов ортогонализированного квадратичного члена для РСП. Расчет коэффициентов уравнения регрессии второго порядка. Дисперсия воспроизводимости. Дисперсии значимости коэффициентов, доверительные интервалы и проверка коэффициентов уравнения регрессии второго порядка на значимость по критерию Стьюдента.

Лекция 4. Остаточная сумма квадратов. Дисперсия адекватности. Проверка уравнения регрессии второго порядка на адекватность по критерию Фишера. Точность прогнозирования по уравнению регрессии второго порядка. Принятие решения на основе уравнения регрессии второго порядка. Нахождение оптимального значения фактора. Оптимизация исследуемого объекта. Алгоритм решения прикладных задач и оптимизация с помощью однофакторного уравнения регрессии второго порядка. Решение базовой задачи.

2.2. Термины и условные обозначения

МНК – метод наименьших квадратов.
РСП – равномерный симметричный план.
N – число опытов в эксперименте.
j – текущий номер опыта, $j = 1, \dots, N$.
n – количество дублей в каждом опыте.
i – текущий номер дубля, $i = 1, \dots, n$.
m – порядок уравнения регрессии полиномиального вида.
q – текущий номер порядка уравнения, $q = 0, \dots, m$.
$\beta_0, \beta_1, \beta_{11}$ – истинные коэффициенты однофакторного уравнения регрессии первого (второго) порядка в обычной форме $Y = \beta_0 + \beta_1 X_1$ ($Y = \beta_0 + \beta_1 X_1 + \beta_{11} X_1^2$).

b_0, b_1, b_{11} – оценочные коэффициенты однофакторного уравнения регрессии первого (второго) порядка в обычной форме $Y = b_0 + b_1 X_1$ ($Y = b_0 + b_1 X_1 + b_{11} X_1^2$).

b_0^*, b_1^*, b_{11}^* – оценочные коэффициенты однофакторного уравнения регрессии второго порядка в ортогонализированной форме $Y = b_0^* + b_1^* X_1 + b_{11}^* (X_1^2 - \lambda_1)$.

λ_1 – ортогонализирующий коэффициент квадратичного члена X_1^2 в однофакторном уравнении регрессии второго порядка.

x_{1j} – натуральное значение фактора x_1 в j-ом опыте.

$x_{1\max}$ ($x_{1\min}$) – максимальное (минимальное) натуральное значение фактора x_1.

$x_{1о}$ – основной уровень натурального значения фактора x_1.

Δx_1 – интервал варьирования натурального значения фактора x_1.

X_{1j} – нормированное значение фактора в j-ом опыте.

X_0 – фиктивный фактор, все значения которого $X_0 = +1$.

Y_{ji} – значение параметра Y в j-ом опыте и i-ом дубле.

$\overline{Y_j}$ – выборочное среднее значение параметра Y в j-ом опыте.

S_j^2 – выборочная дисперсия в j-ом опыте.

S_j – выборочное стандартное отклонение в j-ом опыте.

f – число степеней свободы.

$\tau_{эj}$ – экспериментальное значение критерия Смирнова–Граббса в j-ом опыте.

$\tau_{f,\,p}$ – критическое значение критерия Смирнова–Граббса с числом степеней свободы f при доверительной вероятности p (приложение 1).

$\max S_j^2$ – максимальное значение выборочной дисперсии в j-ом опыте.

$G_э$ – экспериментальное значение критерия Кохрена.

$G_{f_1, f_2, p}$ – критическое значение критерия Кохрена для сравнения N дисперсий; в числителе максимальная дисперсия с числом степеней свободы f_1, в знаменателе сумма всех дисперсий с числом степеней свободы f_2 при доверительной вероятности p (приложение 5).

φ – остаточная сумма квадратов.

Y_j^p – расчетное значение параметра Y в j-ом опыте.

$\frac{\partial \varphi}{\partial b_0}$ – частная производная φ по b_0.

$\frac{\partial \varphi}{\partial b_1}$ – частная производная φ по b_1.

$\frac{\partial \varphi}{\partial b_{11}}$ – частная производная φ по b_{11}.

$S_{воспр}^2$ – дисперсия воспроизводимости.

$S^2(b_0), S^2(b_0^*), S^2(b_1), S^2(b_{11})$ – дисперсии значимости коэффициентов однофакторного уравнения регрессии.

$t_{f,\,p}$ – критическое значение критерия Стьюдента с числом степеней свободы f при доверительной вероятности p (приложение 2).

Δb_0, Δb_0^*, Δb_1, Δb_{11} – доверительные интервалы коэффициентов уравнения регрессии первого порядка.

$S_{ад}^2$ – дисперсия адекватности, $f = N - B$.

B – число значимых коэффициентов уравнения регрессии.

$\max(S_{ад}^2, S_{воспр}^2)$ – максимальная из дисперсий $S_{ад}^2$ и $S_{воспр}^2$.

$\min(S_{ад}^2, S_{воспр}^2)$ – минимальная из дисперсий $S_{ад}^2$ и $S_{воспр}^2$.

$F_э$ – экспериментальное значение критерия Фишера.

$F_{f_1, f_2, p}$ – критическое значение критерия Фишера для сравнения двух дисперсий S_1^2 и S_2^2; в числителе бóльшая дисперсия с числом степеней свободы f_1, в знаменателе меньшая с числом степеней свободы f_2 при доверительной вероятности p (приложение 4).

$\overline{X_1}$ – среднее значение фактора X_1 по всем опытам.

$\overline{X_1}^2$ – квадрат среднего значения фактора X_1 по всем опытам.

$\overline{X_1^2}$ – среднее квадрата значения фактора X_1 по всем опытам.

Y' – первая производная параметра Y по X_1.

Y'' – вторая производная параметра Y по X_1.

2.3. Уравнение регрессии первого порядка (тема 5)

Решение задач агропромышленного комплекса в области животноводства, растениеводства, переработки сельхозпродукции, проектирования и производства машин и механизмов связано с построением математических моделей, позволяющих делать прогнозы, управлять процессами, осуществлять оптимизацию.

Математическая модель представляет собой функциональную зависимость некоторого параметра от одного или нескольких факторов. Исторически первой задачей моделирования являлся поиск функциональной зависимости параметра от одного фактора $Y = f(x_1)$, например, производительности аппарата для сушки зерна (Y – параметр) в зависимости от температуры сушильного агента (x_1–фактор). В том случае, когда аналитическая зависимость не может быть построена на основе фундаментального знания о природе объекта, математическая модель строится статистическими методами, например, в виде степенного ряда – алгебраического полинома:

$$\eta = \beta_0 + \beta_1 x_1 + \beta_{11} x_1^2 + \beta_{111} x_1^3 + ... + \beta_{11...1} x_1^m . \quad (56)$$

Однако по результатам конечного числа опытов можно получить лишь некоторую приближенную оценку истинных коэффициентов математической модели, называемой уравнением регрессии:

$$Y = b_0 + b_1 x_1 + b_{11} x_1^2 + b_{111} x_1^3 + ... + b_{11...1} x_1^m . \quad (57)$$

Следует отметить, что использование натуральных значений факторов для построения математических моделей усложняет процедуру расчетов. Поясним это на примере построения математической модели, описывающей производительность аппарата для сушки зерна в зависимости от температуры сушильного агента. При повышении степени уравнения регрессии (57) задача нахождения его коэффициентов становится достаточно громоздкой. Пусть в некотором опыте температура сушильного агента $x_{1j} = 95$ °С. Тогда $x_{1j}^4 = 81450625$ (°С)^4, и построение уравнения регрессии (57), в котором используют натуральные значения фактора x_1, потребует громоздких вычислений. Кроме того,

задача усложняется, когда исследуемый объект зависит от нескольких факторов, имеющих различную размерность.

2.3.1. Теория

Для упрощения процедуры расчетов коэффициентов уравнения регрессии (57) вводится операция нормирования факторов, которая позволяет избежать отрицательных моментов, связанных с различной размерностью и масштабом факторов. Переход от натуральных значений фактора x_{1j} к нормированным X_{1j} осуществляется по формулам:

$$X_{1j} = \frac{x_{1j} - x_{1o}}{\Delta x_1}, \quad (58)$$

$$x_{1o} = \frac{x_{1\max} + x_{1\min}}{2}, \quad (59)$$

$$\Delta x_1 = \frac{x_{1\max} - x_{1\min}}{2}. \quad (60)$$

Уравнения (58)–(60) нормируют значения фактора X_{1j} на единичный отрезок $X_{1j} \in [-1, +1]$, в то время как диапазон изменения физических значений фактора $x_{1j} \in [x_{1\min}, x_{1\max}]$. Натуральные значения фактора x_1 имеют некоторую размерность, а нормированные значения фактора X_1 безразмерны. Переход от натуральных значений фактора x_{1j} к нормированным X_{1j} и обратно проиллюстрирован на примере сушильного агрегата (таблица 7).

Практика научных исследований показала, что при отсутствии априорной информации об изучаемом объекте исследование целесообразно начинать с построения уравнения регрессии первого порядка:

$$Y = b_0 X_0 + b_1 X_1 . \quad (61)$$

В уравнение (61) для единообразия обработки и записи экспериментальных данных введен фиктивный фактор X_0, связанный с коэффициентом b_0. Все значения фиктивного фактора $X_{0j} = +1$.

Таблица 7 – Переход от натуральных значений фактора x_1 к нормированным X_1

Характеристики фактора	Натуральные значения фактора x_{1j}	Нормированные значения фактора X_{1j}
Основной уровень	65 °C	0
Интервал варьирования	25 °C	1
Верхний уровень	40 °C	+1
Нижний уровень	90 °C	–1
Формулы перехода	$X_{1j} = \dfrac{x_{1j} - 65}{25}$; $x_{1j} = 65 + 25X_{1j}$	

Прямую, которая является геометрическим образом уравнения (61), можно было бы провести через «облако» экспериментальных данных «на глаз» (рисунок 2).

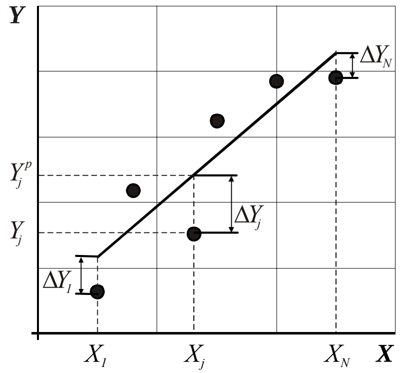

Рисунок 2 – Геометрический образ метода наименьших квадратов

Очевидно, что построение прямой таким способом неоднозначно, весьма приближенно и не имеет математического обоснования. Для строгого и однозначного построения искомого уравнения прямой (61) по экспериментальным данным, однозначно задаваемого коэффициентами b_0 и b_1, используют метод Лагранжа – метод наименьших квадратов (МНК), основанный на поиске минимума остаточной суммы квадратов φ между значениями параметров, рассчитанными по уравнению регрессии Y_j^p, и экспериментальными данными $\overline{Y}_j$ (рисунок 2):

$$\varphi = \varphi(b_0, b_1) = \sum_{j=1}^{N} \Delta Y_j^2 = \sum_{j=1}^{N} \left(Y_j^p - \overline{Y}_j \right)^2 = \sum_{j=1}^{N} \left(b_0 X_{0j} + b_1 X_{1j} - \overline{Y}_j \right)^2 = \min. \quad (62)$$

Из математического анализа известно, что необходимым условием минимума функции $\varphi = \varphi(b_0, b_1)$ является равенство нулю ее частных производных:

$$\begin{cases} \dfrac{\partial \varphi}{\partial b_0} = \sum\limits_{j=1}^{N} 2\left(b_0 X_{0j} + b_1 X_{1j} - \overline{Y}_j \right) X_{0j} = 0; \\ \dfrac{\partial \varphi}{\partial b_1} = \sum\limits_{j=1}^{N} 2\left(b_0 X_{0j} + b_1 X_{1j} - \overline{Y}_j \right) X_{1j} = 0. \end{cases} \quad (63)$$

После преобразования система (63) приобретает следующий вид:

$$\begin{cases} b_0 \sum\limits_{j=1}^{N} X_{oj}^2 + b_1 \cdot \sum\limits_{j=1}^{N} X_{0j} X_{1j} = \sum\limits_{j=1}^{N} X_{0j} \overline{Y}_j \,; \\ b_0 \sum\limits_{j=1}^{N} X_{0j} X_{1j} + b_1 \cdot \sum\limits_{j=1}^{N} X_{1j}^2 = \sum\limits_{j=1}^{N} X_{1j} \overline{Y}_j \,. \end{cases} \quad (64)$$

Системы уравнений (64) достаточно для нахождения неизвестных коэффициентов уравнения регрессии первого порядка b_0 и b_1 (два уравнения с двумя неизвестными).

Аналогичным способом можно получить систему уравнений для нахождения коэффициентов уравнения регрессии любого порядка. Например, для нахождения коэффициентов уравнения регрессии четвертого порядка (5 коэффициентов) МНК составляется система из 5-ти уравнений с 5-ю неизвестными. Решение таких систем, которые, как правило, близки к вырожденным, представляет некоторые трудности (необходимо производить достаточно громоздкие расчеты с высокой точностью). Кроме того, если по какой-то причине потребуется исключить тот или иной коэффициент, то нахождение оставшихся коэффициентов потребует повторного решения системы уравнений, хотя и на одно уравнение меньше.

Решение системы уравнений (64), полученной МНК, значительно упростилось бы, если бы система имела диагональный вид. Следует отметить, что главное достоинство системы уравнений диагонального вида заключается не столько в упрощении процедуры расчета коэффициентов уравнения регрессии, сколько в возможности независимого их определения. Система уравнений (64), приведенная к диагональному виду, выглядела бы следующим образом:

$$\begin{cases} b_0 \sum_{j=1}^{N} X_{0j}^2 + 0 = \sum_{j=1}^{N} X_{0j} \overline{Y}_j ; \\ 0 + b_1 \sum_{j=1}^{N} X_{1j}^2 = \sum_{j=1}^{N} X_{1j} \overline{Y}_j . \end{cases} \qquad (65)$$

Однако, для того чтобы система уравнений (64) имела бы диагональный вид, необходимо выполнить условие:

$$\sum_{j=1}^{N} X_{0j} X_{1j} = 0 . \qquad (66)$$

Добиться выполнения условия (66) можно только путем планирования эксперимента, при котором значения фактора подбираются специальным образом. Планирование эксперимента – один из важнейших разделов математической теории эксперимента, под которым подразумевается выбор значений факторов и их комбинаций для достижения поставленной цели.

В дальнейшем для построения однофакторных уравнений регрессии будем рассматривать только такие планы эксперимента, для которых выполняется условие (66) и которые приводят систему уравнений (64) к диагональному виду (65).

В общем случае свойство $\sum_{j=1}^{N} X_{0j} X_{1j} = 0$ называется ортогональностью факторов.

Для выполнения уравнения (66) можно построить большое количество планов. В таблице 8 приведено 4 плана, построеные для изучения производительности сушильного агрегата в зависимости от температуры сушильного агента (воздуха).

В 1-ом плане значения нормированного фактора X_{1j} неравномерны и несимметричны относительно центра плана $X_{1j} = 0$. Произведение $\sum_{j=1}^{N} X_{0j} X_{1j} = 1.20 \neq 0$. Факторы X_0 и X_{1j} неортогональны, и поэтому система уравнений (64) для такого плана не диагональна

Таблица 8 – Примеры планов однофакторного эксперимента

N	X_{0j}	1-й план			2-й план			3-й план			4-й план		
		x_{1j}	X_{1j}	$X_{0j}X_{1j}$	x_{1j}	X_{1j}	$X_{0j}X_{1j}$	x_{1j}	X_{1j}	$X_{0j}X_{1j}$	x_{1j}	X_{1j}	$X_{0j}X_{1j}$
1	1	40	–1	–1	40	–1	–1	40	–1	–1	40	–1	–1
2	1	60	–0.2	–0.2	42.5	–0.9	–0.8	45	–0.8	–0.8	50	–0.6	–0.6
3	1	65	0	0	65	0	0	50	–0.6	–0.6	60	–0.2	–0.2
4	1	80	0.6	0.6	72.5	0.3	0.2	80	0.6	0.6	70	0.2	0.2
5	1	85	0.8	0.8	80	0.6	0.6	85	0.8	0.8	80	0.6	0.6
6	1	90	1	1	90	1	1	90	1	1	90	1	1
$\sum X$	**6**	–	**1.2**	**1.2**	–	**0**	**0**	–	**0**	**0**	–	**0**	**0**
$\sum X^2$	**6**	–	**3.04**	–	–	**3.26**		–	**4.00**	–	–	**2.80**	–

Во 2-ом плане значения нормированного фактора X_{1j} неравномерны и несимметричны относительно центра плана, однако факторы X_0 и X_{1j} ортогональны, так как $\sum_{j=1}^{N} X_{0j} X_{1j} = 0$, и поэтому система уравнений (64) будет иметь диагональный вид.

В 3-ем плане значения нормированного фактора X_{1j} неравномерны, но симметричны относительно центра плана. Факторы X_0 и X_{1j} ортогональны, так как $\sum_{j=1}^{N} X_{0j} X_{1j} = 0$, и потому система уравнений (64) будет иметь диагональный вид.

В 4-ом плане приведен равномерный и симметричный план (РСП). Значения нормированного фактора X_{1j} и равномерны, и симметричны относительно центра плана. Для РСП ортогональность факторов X_0 и X_{1j} очевидна $\sum_{j=1}^{N} X_{0j} X_{1j} = 0$. Система уравнений (64) для такого плана будет иметь диагональный вид.

Из приведенных примеров можно сделать вывод, что планов, для которых выполняется условие (66), можно построить большое количество. Отметим, что планов аналогичных 2-му и 3-му можно составить бесконечное множество, а 4-й план равномерный и симметричный (РСП) – единственный. В дальнейшем для построения однофакторных уравнений регрессии любого порядка будем использовать только РСП (ниже не раз убедимся в его эффективности), натуральные значения которого рассчитываются по следующей формуле:

$$x_{1j} = x_{1\min} + \frac{j-1}{N-1} \cdot (x_{1\max} - x_{1\min}), \; j = 1, \ldots, N. \qquad (67)$$

Если натуральные значения фактора x_{1j}, задаваемые для РСП уравнением (67), подставить в уравнение (58), то для расчета нормированных значений фактора X_{1j} получим следующую формулу (вывод уравнения (68) приведен в приложении 6):

$$X_{1j} = 2 \cdot \frac{j-1}{N-1} - 1. \qquad (68)$$

Обратите внимание, что X_{1j} в РСП не зависит ни от $x_{1\max}$ ни от $x_{1\max}$, а только от числа опытов N.

При изучении исследуемого объекта натуральные и нормированные значения фактора, а также все дубли экспериментальных значений параметра Y_{ji} удобно записывать в виде таблицы 9.

Таблица 9 – План однофакторного эксперимента

№ опыта	Значения факторов		Значения параметров Y_{ji}					Предварительная обработка	
	x_{1j}	X_{1j}	Y_{j1}	…	Y_{ji}	…	Y_{jn}	$\bar{Y}_j$	S_j^2
1	x_{11}	X_{11}	Y_{11}	…	Y_{1i}	…	Y_{1n}	$\bar{Y}_1$	S_1^2
…		…	…	…	…	…	…	…	…
j	x_{1j}	X_{1j}	Y_{j1}	…	Y_{ji}	…	Y_{jn}	$\bar{Y}_j$	S_j^2
…		…	…	…	…	…	…	…	…
N	x_{1N}	X_{1N}	Y_{N1}	…	Y_{Ni}	…	Y_{Nn}	$\bar{Y}_N$	S_N^2
$G_э$			$G_{n-1,\,N,\,p}$					$\sum_{j=1}^{N} S_j^2$	

Легко показать, что РСП, задаваемый уравнением (68), в силу симметричности элементов столбца фактора X_{1j} обеспечивает выполнение условия ортогональности (66) (доказательство приведено в приложении 6):

$$\sum_{j=1}^{N} X_{0j} X_{1j} = \sum_{j=1}^{N} X_j = \sum_{j=1}^{N} \left(\frac{2(j-1)}{N-1} - 1 \right) = 0. \qquad (69)$$

Диагональный вид системы (65) для РСП позволяет независимо рассчитывать коэффициенты уравнения регрессии первого порядка:

$$b_0 = \frac{\sum_{j=1}^{N} X_{0j} \bar{Y}_j}{\sum_{j=1}^{N} X_{0j}^2}, \quad b_1 = \frac{\sum_{j=1}^{N} X_{1j} \bar{Y}_j}{\sum_{j=1}^{N} X_{1j}^2}. \qquad (70)$$

С учетом того, что $X_{0j} = 1$ для всех $j = 1, \ldots, N$, а X_{1j} для РСП рассчитываются по уравнению (68), имеем:

$$\sum_{j=1}^{N} X_{0j}^2 = N, \qquad (71)$$

$$\sum_{j=1}^{N} X_{1j}^2 = \frac{N(N+1)}{3(N-1)}. \qquad (72)$$

Справедливость уравнения (71) очевидна (см., например, таблицу 8), а доказательство уравнения (72) приводится в приложении 6.

Уравнения (70) для расчета коэффициентов регрессии первого порядка с учетом уравнений (71), (72) для РСП принимают следующий вид:

$$b_0 = \frac{\sum_{j=1}^{N} X_{0j} \bar{Y}_j}{N}, \quad b_1 = \frac{3(N-1)}{N+1} \cdot \frac{\sum_{j=1}^{N} X_{1j} \bar{Y}_j}{N}. \qquad (73)$$

Еще раз напоминаем, что прежде чем рассчитывать коэффициенты регрессии, необходимо проверить данные каждого опыта на промахи, на принадлежность их нормальному закону распределения, а выборочные дисперсии всех опытов – на однородность по критерию Кохрена (таблица 9).

Поскольку для оценки коэффициентов регрессии используется конечное число опытов N и дублей n, необходимо произвести статистическую оценку полученных коэффициентов уравнения регрессии на значимость, а найденное уравнение регрессии – на адекватность.

Проверка коэффициентов уравнения регрессии на значимость по критерию Стьюдента. После подтверждения однородности дисперсии всех опытов рассчитывается средняя дисперсия, называемая дисперсией воспроизводимости:

$$S^2_{\text{воспр}} = \frac{\sum_{j=1}^{N} S_j^2}{N}, \ f = N(n-1). \tag{74}$$

Дисперсии значимости коэффициентов уравнения регрессии первого порядка рассчитывают по следующим формулам (приводятся без доказательства):

$$S^2(b_0) = \frac{S^2_{\text{воспр}}}{n\sum_{j=1}^{N} X_{0j}^2}, \quad S^2(b_1) = \frac{S^2_{\text{воспр}}}{n\sum_{j=1}^{N} X_{1j}^2}. \tag{75}$$

Уравнения (75) с учетом уравнений (71), (72) принимают следующий вид:

$$S^2(b_0) = \frac{S^2_{\text{воспр}}}{Nn}, \qquad S^2(b_1) = \frac{3(N-1)}{N+1} \cdot \frac{S^2_{\text{воспр}}}{Nn}. \tag{76}$$

Если учесть, что практически всегда $N >> 1$, то уравнения (76) имеют достаточно простой и наглядный вид:

$$S^2(b_0) = \frac{S^2_{\text{воспр}}}{Nn}, \qquad S^2(b_1) \approx 3 \cdot \frac{S^2_{\text{воспр}}}{Nn}. \tag{77}$$

Доверительные интервалы для каждого коэффициента регрессии рассчитывают по следующей формуле (сравните с уравнением (43)):

$$\Delta b_q = t_{N(n-1),\,p} \cdot S(b_q), \ q = 0, 1, \tag{78}$$

где $t_{N(n-1),\,p}$ – критическое значение критерия Стьюдента (приложение 2).

Если выполняется неравенство $\Delta b_q < |b_q|$, то коэффициенты регрессии b_q статистически значимы. Если $\Delta b_q > |b_q|$, то коэффициенты регрессии b_q статистически незначимы. Незначимые коэффициенты следует исключить из уравнения регрессии.

Проверка уравнения регрессии на адекватность по критерию Фишера. Идея проверки уравнения регрессии на адекватность заключается в сравнении двух дисперсий: дисперсии адекватности $S^2_{\text{ад}}$ (мера рассеяния выборочных средних в каждом опыте $\overline{Y}_j$ относительно

расчетных значений параметра Y_j^p) и дисперсии воспроизводимости $S^2_{\text{воспр}}$ (мера рассеяния случайных значений параметра в каждом опыте относительно выборочных средних). Уравнение адекватно, если генеральные дисперсии равны $\sigma^2_{\text{ад}} = \sigma^2_{\text{воспр}}$. Однородность двух выборочных дисперсий $S^2_{\text{ад}}$ и $S^2_{\text{воспр}}$ определяется менее жестким критерием в силу относительно небольшого размера выборок, по которым они рассчитываются, и определяется по критерию Фишера.

Прежде чем пояснить процедуру проверки уравнений регрессии на адекватность, приведем алгоритм расчета выборочной дисперсии адекватности $S^2_{\text{ад}}$.

1. По уравнению регрессии первого порядка рассчитывают значения параметра Y_j^p в каждом опыте $Y_j^p = b_0 + b_1 X_{1j}$, $j = 1, \ldots, N$.

2. В каждом опыте рассчитывают квадрат разности между средним значением и расчетным $\left(\overline{Y}_j - Y_j^p\right)^2$, $j = 1, \ldots, N$.

3. Рассчитывают остаточную сумму квадратов:

$$\varphi = \sum_{j=1}^{N} \left(\overline{Y}_j - Y_j^p\right)^2. \tag{79}$$

4. Выборочная дисперсия адекватности $S^2_{\text{ад}}$ определяется по формуле: (???)

$$S^2_{\text{ад}} = \frac{n\varphi}{N-B}, \qquad f = N - B, \tag{80}$$

где B – число значимых коэффициентов уравнения регрессии.

Дисперсия адекватности $S^2_{\text{ад}}$, рассчитанная по формуле (80), является мерой рассеяния выборочных средних параметра в каждом опыте $\overline{Y}_j$ относительно расчетных значений Y_j^p.

Анализ уравнения (80) показывает, что минимальное количество опытов N, необходимое для проверки однофакторного уравнения регрессии первого порядка (61) на адекватность, должно быть не менее трех, то есть $N \geq 3$. Для $N = 2$ уравнение регрессии первого порядка строится однозначно, так как через две точки всегда можно провести единственную прямую. Если при этом оба коэффициента значимы, то есть $N = B,$ то согласно уравнению (80) $S^2_{ад} = \infty$. Из практического опыта с целью повышения надежности проверки полученного уравнения регрессии на адекватность рекомендуем придерживаться следующего правила: ??

$$N \geq 2(m+1)\ . \tag{81}$$

Поэтому при моделировании исследуемого объекта однофакторным уравнением регрессии первого порядка количество опытов целесообразно брать не менее четырех (показатель степени уравнения $m = 1$), а при моделировании однофакторным уравнением второго порядка – не менее шести ($m = 2$). Однако, если стоимость опытов относительно невысока, они просты и не требуют большого времени для их реализации, то количество опытов и их дублей можно увеличить, так как с ростом Nn возрастает точность и информативность полученной модели.

Сравнение двух выборочных дисперсий $S^2_{ад}$ и $S^2_{воспр}$, как было показано в разделе 1.5.1, осуществляется по критерию Фишера $F_э$:

$$F_э = \frac{\max\left(S^2_{ад}, S^2_{воспр}\right)}{\min\left(S^2_{ад}, S^2_{воспр}\right)}\ . \tag{82}$$

Рассмотрим 4 случая.

1. Если $F_э = \dfrac{S^2_{ад}}{S^2_{воспр}} < F_{N\text{-}B,\, N(n\text{-}1),\, p}$ при $S^2_{ад} > S^2_{воспр}$, (83)

где $F_{N–B,\, N(n–1),\, p}$ – критическое значение критерия Фишера (приложение 4), то $\sigma^2_{ад} = \sigma^2_{воспр}$ и уравнение регрессии первого порядка адекватно.

2. Если $F_э = \dfrac{S^2_{ад}}{S^2_{воспр}} > F_{N\text{-}B,\, N(n\text{-}1),\, p}$ при $S^2_{ад} > S^2_{воспр}$, (84)

то $\sigma^2_{ад} > \sigma^2_{воспр}$ и уравнение регрессии неадекватно. В таких случаях следует перейти к построению уравнения регрессии более высокого порядка, так как рассеяние выборочных средних относительно рассчитанных значений существенно больше рассеяния случайных значений параметра в каждом опыте относительно выборочных средних. В данном случае следует перейти к построению уравнения регрессии второго порядка. При этом необходимо учесть, что если для построения уравнения регрессии первого порядка было выполнено 4 опыта, то для построения уравнения регрессии второго порядка их следует увеличить минимум до 6 (уравнение (81)).

3. Если $F_э = \dfrac{S^2_{воспр}}{S^2_{ад}} < F_{N(n\text{-}1),\, N\text{-}B,\, p}$ при $S^2_{ад} < S^2_{воспр}$, (85)

где $F_{N(n–1),\, N–B,\, p}$ – критическое значение критерия Фишера (приложение 4), то $\sigma^2_{ад} = \sigma^2_{воспр}$ и уравнение регрессии первого порядка адекватно.

4. Если $F_э = \dfrac{S^2_{воспр}}{S^2_{ад}} > F_{N(n\text{-}1),\, N\text{-}B,\, p}$ при $S^2_{ад} < S^2_{воспр}$, (86)

то $\sigma^2_{воспр} > \sigma^2_{ад}$ (несмотря на то, что в каждом опыте наблюдается большой разброс случайных значений параметра относительно выборочных средних, тем не менее, выборочные средние очень близко ложатся на уравнение регрессии). Вероятность такого случая на практике ничтожно мала и является следствием плохой организации эксперимента, в котором нарушена идея случайности (нарушены основные статистические предпосылки). В этом случае эксперимент следует повторить, причем желательно увеличить и число опытов N, и число дублей n.

Если полученное уравнение регрессии первого порядка адекватно, то с его помощью можно оценить предельную абсолютную погрешность рассчитанного параметра Y. Так как уравнение регрессии первого порядка имеет вид:

$$Y = b_0 + b_1 X_1\ , \tag{87}$$

где b_0, b_1 – независимые коэффициенты, имеющий каждый свой доверительный интервал значимости (см. уравнение (78)), а X_1 – детерминированный фактор, то, с учетом уравнений (36), (37), (87) получим:

$$\Delta Y(X_1) = t_{N(n-1),\, p} \cdot S(Y)\ , \tag{88}$$

где $$S^2(Y) = S^2(b_0) + X_1^2 S^2(b_1). \qquad (89)$$

Для РСП с учетом уравнений (76) уравнение (89) примет следующий вид:

$$S^2(Y) = \frac{S^2_{\text{воспр}}}{Nn}\left(1 + \frac{3(N-1)}{N+1}X_1^2\right). \qquad (90)$$

Подставляя уравнение (90) в (88) получаем

$$\Delta Y(X_1) = t_{N(n-1)} \cdot \frac{S_{\text{воспр}}}{\sqrt{Nn}}\sqrt{1 + \frac{3(N-1)}{N+1}X_1^2}. \qquad (91)$$

Анализ уравнения (91) позволяет для РСП сделать несколько выводов:

а) минимальная предельная абсолютная погрешность достигается в центре факторного пространства $X_1 = 0$:

$$\Delta Y(0) = t_{N(n-1)} \cdot \frac{S_{\text{воспр}}}{\sqrt{Nn}}; \qquad (92)$$

б) максимальная предельная абсолютная погрешность достигается на границах факторного пространства при $X_1 = \pm 1$:

$$\Delta Y(\pm 1) = 2t_{N(n-1)} \cdot \frac{S_{\text{воспр}}}{\sqrt{Nn}} \cdot \sqrt{\frac{N-0.5}{N+1}}. \qquad (93)$$

Если учесть, что практически всегда $N >> 1$, то уравнение (93) примет простой и наглядный вид:

$$\Delta Y(\pm 1) = 2t_{N(n-1)} \cdot \frac{S_{\text{воспр}}}{\sqrt{Nn}} = 2 \cdot \Delta Y(0) \qquad (94)$$

в) предельная абсолютная погрешность параметра Y, рассчитанного по адекватному уравнению регрессии первого порядка, значительно меньше предельной абсолютной погрешности, найденной по экспериментальным данным в каждом опыте.

Подтвердим количественно это утверждение. Так как в эксперименте все дисперсии однородны, а это подтверждается по критерию Кохрена, и $S_{\text{воспр}}$ является средней выборочной дисперсией для всех

опытов (см. уравнение (74)), то в каждом опыте предельную абсолютную погрешность определения выборочного среднего можно рассчитать по уравнению (43):

$$\Delta \overline{Y}_j = \frac{t_{n-1,\,p} \cdot S_{\text{воспр}}}{\sqrt{n}}. \qquad (95)$$

Отношение предельных экспериментальных и расчетных абсолютных погрешностей в центре плана ($X = 0$) согласно уравнениям (92) и (95), равно

$$\frac{\Delta \overline{Y}_j}{\Delta Y(0)} \approx \frac{t_{n-1,\,p} \cdot S_{\text{воспр}} \cdot \sqrt{Nn}}{\sqrt{n} \cdot t_{N(n-1),\,p} \cdot S_{\text{воспр}}} \approx \frac{t_{n-1,\,p}\sqrt{N}}{t_{N(n-1),\,p}}. \qquad (96)$$

Так как в экспериментах $N \approx 10$, $n \approx 4$, то $t_{30,\ 0.95} \approx 2$, $t_{3,\ 0.95} \approx 3$ (приложение 2), то и из уравнения (96) следует, что в центре плана

$$\frac{\Delta \overline{Y}_j}{\Delta Y(0)} \approx \frac{3 \cdot \sqrt{10}}{2} = 4.74. \qquad (97)$$

Таким образом, предельная абсолютная погрешность рассчитанного параметра Y в центре РСП примерно в 5 раз меньше предельной абсолютной погрешности средних значений параметров, найденных экспериментально в каждом опыте, а на границе факторного пространства при $X = \pm 1$, с учетом уравнения (95), меньше примерно в 2.5 раза (убедитесь в этом самостоятельно).

Таким образом, математическая модель объекта позволяет делать прогноз точнее, чем породивший ее эксперимент!

2.3.2. Алгоритм решения прикладных задач

1. Выбрать объект исследования и на основе научных и патентных источников собрать необходимую априорную информацию о факторах и параметрах, характеризующих исследуемый объект.

2. Выбрать наиболее существенный параметр Y и n – число дублей в каждом опыте ($i = 1, \ldots, n$). Количество дублей n выбирается из соображений экономии материальных затрат (рекомендуемое число дублей не менее 4).

3. Выбрать наиболее существенный фактор x_1 и интервал его варьирования $[x_{1\min}, x_{1\max}]$, а также N – число опытов (рекомендуемое число опытов не менее $N \geq 2(m+1) = 4$).

3.1. Для построения РСП рассчитать натуральные значения фактора x_{1j} по следующей формуле:

$$x_{1j} = x_{1\min} + \frac{j-1}{N-1} \cdot (x_{1\max} - x_{1\min}),\ j = 1, \ldots, N.$$

3.2. Для РСП рассчитать нормированные значения фактора X_j:

$$X_{1j} = 2 \cdot \frac{j-1}{N-1} - 1,\ j = 1, \ldots, N.$$

4. Обеспечить приборное и технологическое оснащение эксперимента, позволяющее измерять исследуемый параметр Y и варьируемый фактор x_1 с требуемой точностью. Провести эксперимент.

5. Выполнить предварительную обработку экспериментальных данных.

5.1. Рассчитать выборочное среднее в каждом опыте:

$$\overline{Y}_j = \frac{\sum_{i=1}^{n} Y_{ji}}{n},\ j = 1, \ldots, N.$$

5.2. Рассчитать выборочную дисперсию в каждом опыте:

$$S_j^2 = \frac{\sum_{i=1}^{n} \left(Y_{ji} - \overline{Y}_j\right)^2}{n-1},\ j = 1, \ldots, N.$$

5.3. Экспериментальные данные каждого опыта проверить на промахи по критерию Смирнова–Граббса (приложение 1):

$$\tau_{эj} = \frac{\max\left|Y_{ji} - \overline{Y}_j\right|}{S_j} > \tau_{n\text{-}2,\,p},\ j = 1, \ldots, N,$$

где $\tau_{n-2,\,p}$ – критическое значение критерия Смирнова–Граббса (приложение 1).

5.4. Экспериментальные данные каждого опыта проверить на принадлежность их нормальному закону распределения:

$$\left|\frac{\text{САО}_j}{S_j} - 0.7979\right| = \left|\frac{\sum_{i=1}^{n}\left|Y_{ji} - \overline{Y}_j\right|}{nS_j} - 0.7979\right| < \frac{0.4}{\sqrt{n}},\ j = 1, \ldots, N.$$

5.5. Дисперсии всех опытов проверить на однородность по критерию Кохрена:

$$\frac{\max S_j^2}{\sum_{j=1}^{N} S_j^2} < G_{n\text{-}1,\,N,\,p},$$

где $G_{n\text{-}1,\,N,\,p}$ – критическое значение критерия Кохрена (приложение 5).

Внимание! Построение уравнения регрессии возможно только в том случае, если в каждом опыте нет промахов, случайные значения параметра в каждом опыте подчиняются нормальному закону распределения, а выборочные дисперсии всех опытов однородны.

6. Построить однофакторное уравнение регрессии первого порядка $Y = b_0X_0 + b_1X_1$, используя экспериментальные данные РСП.

6.1. В матрице планирования создать столбцы $X_{0j}\overline{Y}_j$, $X_{1j}\overline{Y}_j$ и рассчитать их суммы $\sum_{j=1}^{N} X_{0j}\overline{Y}_j$ и $\sum_{j=1}^{N} X_{1j}\overline{Y}_j$.

6.2. Рассчитать коэффициенты регрессии b_0, b_1:

$$b_0 = \frac{\sum_{j=1}^{N} X_{0j}\overline{Y}_j}{\sum_{j=1}^{N} X_{0j}^2} = \frac{\sum_{j=1}^{N} X_{0j}\overline{Y}_j}{N},\quad b_1 = \frac{\sum_{j=1}^{N} X_{1j}\overline{Y}_j}{\sum_{j=1}^{N} X_{1j}^2} = \frac{3(N-1)}{N+1} \cdot \frac{\sum_{j=1}^{N} X_{1j}\overline{Y}_j}{N}.$$

7. Проверить коэффициенты регрессии b_0, b_1 на значимость.

7.1. Рассчитать дисперсию воспроизводимости:

$$S^2_{\text{воспр}} = \frac{\sum_{j=1}^{N} S_j^2}{N},\quad f = N(n-1).$$

7.2. Рассчитать дисперсии значимости коэффициентов регрессии однофакторного уравнения регрессии первого порядка b_0, b_1:

$$S^2(b_0)=\frac{S^2_{\text{воспр}}}{n\sum_{j=1}^{N}X^2_{0j}}=\frac{S^2_{\text{воспр}}}{Nn},\quad S^2(b_1)=\frac{S^2_{\text{воспр}}}{n\sum_{j=1}^{N}X^2_{1j}}=\frac{3(N-1)}{N+1}\cdot\frac{S^2_{\text{воспр}}}{Nn}.$$

7.3. Рассчитать доверительный интервал для коэффициентов регрессии b_0, b_1:

$$\Delta b_0=t_{N(n\text{-}1),\,p}\cdot S(b_0),\quad \Delta b_1=t_{N(n\text{-}1),\,p}\cdot S(b_1),$$

где $t_{N(n-1),\,p}$ – критическое значение критерия Стьюдента (приложение 2).

7.4. Коэффициенты, для которых $\Delta b_q > |b_q|$, считать незначимыми и исключить из уравнения регрессии.

8. Проверить полученное уравнение регрессии на адекватность по критерию Фишера.

8.1. Рассчитать в каждом опыте значения параметра Y_j^p по уравнению регрессии первого порядка $Y_j^p = b_0 + b_1X_{1j}$, $j = 1, \ldots, N$.

8.2. Рассчитать в каждом опыте квадрат разности выборочного среднего и расчетного значения $(\overline{Y}_j - Y_j^p)^2$, $j = 1, \ldots, N$, а также остаточную сумму квадратов:

$$\varphi=\sum_{j=1}^{N}(\overline{Y}_j - Y_j^p)^2 .$$

8.3. Рассчитать дисперсию адекватности:

$$S^2_{\text{ад}}=\frac{n\varphi}{N-B},\quad f=N-B .$$

8.4. Рассчитать экспериментальное значение критерия Фишера:

$$F_{\text{э}}=\frac{\max(S^2_{\text{ад}},S^2_{\text{воспр}})}{\min(S^2_{\text{ад}},S^2_{\text{воспр}})} .$$

8.5. Если $F_{\text{э}}=\dfrac{S^2_{\text{ад}}}{S^2_{\text{воспр}}}<F_{N\text{-}B,\,N(n\text{-}1),\,p}$ при $S^2_{\text{ад}}>S^2_{\text{воспр}}$,

где $F_{N-B,\,N(n-1),\,p}$ – критическое значение критерия Фишера (приложение 4), то $\sigma^2_{\text{ад}}=\sigma^2_{\text{воспр}}$ и уравнение регрессии первого порядка адекватно.

8.6. Если $F_{\text{э}}=\dfrac{S^2_{\text{ад}}}{S^2_{\text{воспр}}}>F_{N\text{-}B,\,N(n\text{-}1),\,p}$ при $S^2_{\text{ад}}>S^2_{\text{воспр}}$,

где $F_{N-B,\,N(n-1),\,p}$ – критическое значение критерия Фишера (приложение 4), то $\sigma^2_{\text{ад}}>\sigma^2_{\text{воспр}}$ и уравнение регрессии первого порядка неадекватно. Следует перейти к построению уравнения регрессии второго порядка $Y=b_0+b_1X_1+b_{11}X_1^2$ (см. раздел 2.4). При этом количество опытов должно быть не менее $2(m + 1)$.

8.7. Если $F_{\text{э}}=\dfrac{S^2_{\text{воспр}}}{S^2_{\text{ад}}}<F_{N(n\text{-}1),\,N\text{-}B,\,p}$, при $S^2_{\text{ад}}<S^2_{\text{воспр}}$,

где $F_{N(n-1),\,N-B,\,p}$ – критическое значение критерия Фишера (приложение 4), то $\sigma^2_{\text{ад}}=\sigma^2_{\text{воспр}}$ и уравнение регрессии первого порядка адекватно.

8.8. Если $F_{\text{э}}=\dfrac{S^2_{\text{воспр}}}{S^2_{\text{ад}}}>F_{N(n\text{-}1),\,N\text{-}B,\,p}$ при $S^2_{\text{ад}}<S^2_{\text{воспр}}$,

то $\sigma^2_{\text{ад}}>\sigma^2_{\text{воспр}}$. Вероятность такого случая на практике ничтожно мала, следует повторить эксперимент, причем с большим числом опытов N и числом дублей n.

9. Рассчитать, используя адекватное однофакторное уравнение регрессии первого порядка, предельную абсолютную погрешность $\Delta Y(X_1)$ параметра Y внутри исследованного факторного пространства $X_{1j}\in[-1, +1]$ по формуле:

$$\Delta Y(X_1)=t_{N(n-1)}\cdot\frac{S_{\text{воспр}}}{\sqrt{Nn}}\sqrt{1+\frac{3(N-1)}{N+1}X_1^2} .$$

10. В случае необходимости преобразовать однофакторное уравнение регрессии первого порядка, заменив нормированное значение фактора X_1 натуральным x_1 по формуле:??

$$X_1=\frac{x_1-x_{1\text{o}}}{\Delta x_1}.$$

2.3.3. Базовая задача

Формулировка условия: изучить влияние витаминной добавки на суточный привес цыплят в «желторотом» возрасте (период 2–20 дней).

Анализ изучаемого объекта: В качестве параметра выберем привес цыплят Y (г/сут.), который зависит, в том числе, и от количества витаминной добавки в корме – фактор x_1 (г/сут.).

Планирование эксперимента (ссылки на алгоритм 2.3.2).

Фактор. При выборе количества опытов будем руководствоваться п. 2: $N \geq 2(m + 1)$. Так как первый этап – построение уравнения регрессии первого порядка ($m = 1$), то $N \geq 4$. Для большей информативности искомого уравнения регрессии выбираем $N = 7$, тем более, что определение массы цыплят относительно простая операция (п. 3). Для реализации эксперимента было организовано 7 контрольных вольеров (по числу опытов), в каждом из которых находилось по 100 цыплят 2-дневного возраста.

Цыплята во всех вольерах получали одинаковый корм по количеству и качеству, за исключением витаминной добавки. Априори было известно, что обычно витаминные добавки применяют в количестве до 20 г/сут на одного цыпленка. Было принято решение варьировать количество добавки от $x_{1\min} = 0$ до $x_{1\max} = 12$ г/сут. В качестве плана эксперимента выбран РСП. Количество добавки в вольерах (опыты № 2–6) рассчитывалось по уравнению (п. 3.1):

$$x_{1j} = x_{1\min} + \frac{j-1}{N-1} \cdot (x_{1\max} - x_{1\min}),\ \ j = 1, \ldots, 7.$$

Результаты расчета РСП в натуральных значениях фактора внесем в таблицу 10. Например, для $j = 5$:

$$x_{15} = 0 + (5 - 1)\cdot(12 - 0)/(7 - 1) = 4 \cdot 12/6 = 8 \text{ г/сут.}$$

Переход от натуральных значений фактора к нормированным. Для упрощения расчетов перейдем от натуральных значений фактора x_{1j} к нормированным X_{1j} по уравнению (п. 3.2):

$$X_{1j} = 2\frac{j-1}{N-1} - 1,\ j = 1, \ldots, 7.$$

Например, для $j = 5$: $X_{15} = 2 \cdot (5 - 1)/(7 - 1) - 1 = 2 \cdot 4/6 - 1 = 0.333$.

Кроме того, для универсализации расчетов в таблице 10 создан столбец фиктивного фактора X_0, для которого все $X_{0j} = 1$.

Параметр. Учитывая несложность операции взвешивания цыплят, принимаем число дублей $n = 4$.

Экспериментальные значения параметра Y (суточный привес одного цыпленка) определялись следующим образом. Сначала в каждом вольере определялась масса 10 **произвольно** (!) выбранных цыплят. Были получены следующие результаты в граммах: 498, 507, 489, 504, 496, 509, 497. Средняя масса одного цыпленка в начале эксперимента по 7-ми вольерам составила $m = (498 + 507 + 489 + 504 + 496 + 509 + 497)/70 = 50.1$ г.

Затем в течение 10 дней в корм цыплят добавляли витаминную добавку в соответствии с РСП (столбец x_{1j} таблицы 10). Например, в корм цыплят вольера № 3 добавляли $4\cdot100 = 400$ г/сут добавки, при этом были приняты меры, чтобы корм съедался всеми цыплятами одинаково. Через 10 дней из каждого вольера произвольно выбирали по 4 цыпленка и определяли массу каждого из них. Для определения суточного привеса за период вскармливания из полученной массы вычиталась масса цыпленка, равная 50.1 г, в начале эксперимента и результат делился на 10 дней (период вскармливания). Полученные результаты внесены в столбцы Y_{ji} таблицу 10.

Предварительная обработка экспериментальных данных.

Выборочное среднее в каждом опыте (п. 5.1):

$$\overline{Y_j} = \frac{\sum_{i=1}^{n} Y_{ji}}{n} = \frac{\sum_{i=1}^{4} Y_{ji}}{4},\ \ j = 1, \ldots, 7.$$

Например, $\overline{Y}_1 = (10.3 + 11.2 + 13.0 + 12.7)/4 = 11.8$.

Выборочная дисперсия в каждом опыте (п. 5.2):

$$S_j^2 = \frac{\sum_{i=1}^{n} (Y_{ji} - \overline{Y_j})^2}{n-1} = \frac{\sum_{i=1}^{4} (Y_{ji} - \overline{Y_j})^2}{4-1},\ \ j = 1, \ldots, 7.$$

Например, $S_1^2 = [(10.3 - 11.8)^2 + (11.2 - 11.8)^2 + (13.0 - 11.8)^2 +$

$+ (12.7 - 11.8)^2]/(4 - 1) = 1.620$, $S_1 = 1.273$.

Таблица 10 – План эксперимента и результаты обработки данных по вскармливанию цыплят

План эксперимента								Результаты обработки эксперимента					
N	x_{1j}	Y_{j1}	Y_{j2}	Y_{j3}	Y_{j4}	$\overline{Y}_j$	S_j^2	X_{0j}	X_{1j}	$X_{0j}\overline{Y}_j$	$X_{1j}\overline{Y}_j$	Y_j^p	$\left(\overline{Y}_j - Y_j^p\right)^2$
1	0	10.3	11.2	13.0	12.7	11.80	1.62	1	–1.000	11.80	–11.800	12.9	1.21
2	2	14.8	14.6	16.2	13.8	14.85	1.00	1	–0.667	14.85	–9.905	14.3	0.30
3	4	16.8	14.5	16.5	16.5	16.08	1.12	1	–0.333	16.08	–5.355	15.6	0.23
4	6	18.8	17.4	16.3	18.0	17.63	1.11	1	0.000	17.63	0.000	17.0	0.40
5	8	18.4	17.4	19.0	18.7	18.38	0.48	1	0.333	18.38	6.121	18.4	0.00
6	10	18.9	19.8	19.6	18.9	19.30	0.22	1	0.667	19.30	12.873	19.7	0.16
7	12	23.4	19.7	20.0	20.7	20.95	2.84	1	1.000	20.95	20.950	21.1	0.02
$N = 7$ (число опытов) $n = 4$ (число дублей)						$\sum_{j=1}^{7} S_j^2 = 8.39$		$\sum_{j=1}^{7} X_{0j} = 7$			$\sum_{j=1}^{7} X_{1j} = 0$	$\varphi = 2.32$	
$S^2_{воспр} = 1.20$, $f_{воспр} = 21$						$x_{1min} = 0$ г/сут $x_{1max} = 12$ г/сут		$\sum_{j=1}^{7} X_{0j}^2 = 7$			$\sum_{j=1}^{7} X_{1j}^2 = 3.111$	$S^2_{ад} = 1.856$, $f = 5$	
$t_{21,\,0.95} = 2.080$						$x_{1о} = 6$, $\Delta x_1 = 6$		$b_0 = 17.00$			$b_1 = 4.1$	$F_э = 1.547$	
$G_э = 0.338$, $G_{3,\,7,\,0.95} = 0.480$						$X_{1j} = (x_{1j} - 6)/6$		$S^2(b_0) = 0.0429$			$S^2(b_1) = 0.0964$	$F_{5,\,21,\,0.95} = 2.685$	
$Y = 17.00 + 4.1X_1$								$S(b_0) = 0.207$			$S(b_1) = 0.310$	Модель адекватна	
$Y = 12.9 + 0.68x_1$								$\Delta b_0 = 0.43$			$\Delta b_1 = 0.6$		

Проверка на промахи (п. 5.3). Проверка экспериментальных данных всех опытов показала, что промахов в них нет. Например, для опыта № 1 наиболее выделяющимся значением является $Y_{11} = 10.3$. Критерий Смирнова–Граббса для этого значения $\tau_{э1} = \max\left|Y_{1i} - \overline{Y}_1\right| / S_1 = |10.3 - 11.8| / 1.273 = 1.18$. Критическое значение критерия Смирнова–Граббса равно $\tau_{2,\,0.95} = 1.69$ (приложение 1, $f = n - 2 = 2$). Так как $\tau_{э1} < \tau_{2,\,0.95}$, то промахов в опыте № 1 нет.

Проверка на принадлежность данных нормальному закону распределения (п. 5.4). Все 4 дубля каждого опыта подчиняются закону нормального распределения. Например, для случайных значений параметра Y_{1i} в опыте № 1 выполняется соотношение:

$$|САО_1/S_1 - 0.7979| < 0.4/\sqrt{n} \rightarrow |1.050/1.273 - 0.7979| < 0.4/2 \rightarrow$$

$$\rightarrow |0.8248 - 0.7979| < 0.4/2 \rightarrow 0.0269 < 0.2,$$

где $САО_1 = \sum_{i=1}^{n}\left|Y_{1i} - \overline{Y}_1\right| / n = (|10.3 - 11.80| + |11.2 - 11.80| +$

$$+ |13.0 - 11.8| + |12.7 - 11.8|)/4 = 1.050.$$

Проверка дисперсий всех опытов на однородность по критерию Кохрена (п. 5.5) показала, что дисперсии всех опытов однородны, так как $G_э < G_{n-1,\,N,\,0.95} = G_{3,\,7,\,0.95} = 0.480$ (приложение 5):

$$G_э = \frac{\max(S_j^2)}{\sum_{j=1}^{7} S_j^2} = \frac{2.84}{1.62 + 1.00 + 1.12 + 1.11 + 0.48 + 0.22 + 2.84} = \frac{2.84}{8.39} = 0.338\,.$$

Поскольку предварительная обработка экспериментальных данных в каждом опыте не выявила отклонений от основных предпосылок регрессионного анализа (промахов нет, нормальный закон распределения и однородность дисперсий соблюдаются), переходим к следующему этапу – построению уравнения регрессии первого порядка.

Построение уравнения регрессии первого порядка. Так как априори неизвестен характер изучаемой зависимости, начнем исследование с построения уравнения регрессии первого порядка: $Y = b_0 + b_1 X_1$. Для этого в таблице 10 создадим столбцы $X_{0j}\overline{Y}_j$, $X_{1j}\overline{Y}_j$ и рассчитаем их суммы (п. 6.1):

$$\sum_{j=1}^{7} X_{0j}\,\overline{Y}_j = 1\cdot 11.80 + 1\cdot 14.85 + 1\cdot 16.08 + 1\cdot 17.63 +$$

$$+ 1\cdot 18.38 + 1\cdot 19.30 + 1\cdot 20.95 = 118.99\,,$$

$$\sum_{j=1}^{7} X_{1j}\,\overline{Y}_j = (-1)\cdot 11.80 + (-0.667)\cdot 14.85 + (-0.333)\cdot 16.08 +$$

$$+ 0\cdot 17.63 + 0.333\cdot 18.38 + 0.667\cdot 19.30 + 1\cdot 20.95 = 12.88\,.$$

Расчет коэффициентов уравнения регрессии (п. 6.2):

$$b_0 = \frac{\sum_{j=1}^{N} X_{0j}\,\overline{Y}_j}{N} = \frac{\sum_{j=1}^{7} X_{0j}\,\overline{Y}_j}{7} = \frac{118.99}{7} = 17.00\,,$$

$$b_1 = \frac{\sum_{j=1}^{N} X_{1j}\overline{Y}_j}{\sum_{j-1}^{N} X_{1j}^2} = \frac{3(N-1)}{N+1}\cdot\frac{\sum_{j=1}^{N} X_{1j}\overline{Y}_j}{N} = \frac{3\cdot(7-1)\cdot 12.88}{(7+1)\cdot 7} = 4.14\,.$$

Проверка на значимость коэффициентов уравнения регрессии первого порядка по критерию Стьюдента.

Дисперсия воспроизводимости (п. 7.1):

$$S^2_{воспр} = \frac{\sum_{j=1}^{N} S_j^2}{N} = \frac{\sum_{j=1}^{7} S_j^2}{7} = \frac{8.39}{7} = 1.20\,.$$

Дисперсии значимости коэффициентов регрессии (п. 7.2):

$$S^2(b_0) = \frac{S^2_{воспр}}{nN} = \frac{1.20}{4\cdot 7} = 0.0429 \;\rightarrow\; S(b_0) = 0.207\,,$$

$$S^2(b_1) = \frac{3(N-1)}{N+1}\cdot\frac{S^2_{воспр}}{nN} = \frac{3\cdot(7-1)\cdot 1.20}{(7+1)\cdot 4\cdot 7} = 0.0964 \;\rightarrow\; S(b_1) = 0.310\,.$$

Доверительные интервалы коэффициентов регрессии (п. 7.3):

$$\Delta b_0 = t_{N(n-1),\,0.95}\cdot S(b_0) = 2.080\cdot 0.207 = 0.4306 \approx 0.43\,,$$

$$\Delta b_1 = t_{N(n-1),\,0.95} \cdot S(b_1) = 2.080 \cdot 0.310 = 0.6448 \approx 0.6\,,$$

где $t_{N(n-1),\,0.95} = t_{21,\,0.95} = 2.080$ – критическое значение критерия Стьюдента (приложение 4). См. правила округления в разделе 1.3.2.

Оба коэффициента регрессии значимы, так как $\Delta b_0 < |b_0|$ (0.43 < 17.00) и $\Delta b_1 < |b_1|$ (0.6 < 4.1) (п. 7.4). Поэтому уравнение регрессии первого порядка имеет следующий вид: $Y = 17.00 + 4.1X_1$.

Проверка уравнения регрессии первого порядка на адекватность по критерию Фишера.

Расчетные значения параметра Y_j^p определяются по уравнению $Y_j^p = 17.00 + 4.14X_{1j}$. Например, $Y_1^p = 17.00 + 4.1 \cdot (-1) = 12.9$ (п. 8.1).

Остаточная сумма квадратов φ (п. 8.2):

$$\varphi = \sum_{j=1}^{7}\left(\overline{Y}_j - Y_j^p\right)^2 = (11.80 - 12.9)^2 + (14.85 - 14.3)^2 +$$

$$+ (16.08 - 15.6)^2 + (17.63 - 17.0)^2 + (18.38 - 18.4)^2 +$$

$$+ (19.30 - 19.7)^2 + (20.95 - 21.1)^2 = 1.21 + 0.30 + 0.23 + 0.40 +$$

$$+ 0.00 + 0.16 + 0.02 = 2.32.$$

Дисперсия адекватности (п. 8.3):

$$S^2_{\text{ад}} = \frac{n\varphi}{N - B} = \frac{4 \cdot 2.32}{7 - 2} = 1.856\,,\ f = N - B = 7 - 2 = 5.$$

Экспериментальное значение критерия Фишера (п. 8.4):

$$F_{\text{э}} = \frac{S^2_{\text{ад}}}{S^2_{\text{воспр}}} = \frac{1.856}{1.20} = 1.547\,,\ \text{так как}\ S^2_{\text{ад}} > S^2_{\text{воспр}}\,.$$

Так как $F_{\text{э}} = 1.567 < F_{5,\,21,\,0.95} = 2.685$ (приложение 4), то это означает, что полученное уравнение регрессии первого порядка $Y = 17.00 + 4.1X_1$ адекватно (п. 8.5).

Точность прогноза параметра Y в исследованном факторном пространстве определяется следующим уравнением (п. 9):

$$\Delta Y(X_1) = t_{N(n-1),\,p}\,\frac{S_{\text{воспр}}}{\sqrt{Nn}}\sqrt{1 + \frac{3(N-1)}{N+1}X_1^2} =$$

$$= 2.080 \cdot \frac{\sqrt{1.20}}{\sqrt{7 \cdot 4}} \cdot \sqrt{1 + \frac{3 \cdot (7-1)}{7+1}X_1^2} = 0.43 \cdot \sqrt{1 + 2.25X_1^2}\,.$$

В центре плана ($x_1 = 6$ г/сут или $X_1 = 0$) предельная абсолютная погрешность равна $\Delta Y(0) = 0.43$ г/сут, а предельная относительная погрешность $\Delta Y(0)/Y(0) = 2.5$ % (0.43/17.0 = 2.5 %). На нижней границе ($x_1 = 0$ г/сут или $X_1 = -1$) соответствующие величины равны $\Delta Y(-1) = 0.78$ г/сут, $\Delta Y(-1)/Y(-1) = 0.78/(17.0 - 4.1) = 6.0$ %. На верхней границе ($x_1 = 12$ г/сут или $X_1 = +1$) эти величины равны $\Delta Y(+1) = 0.78$ г/сут, $\Delta Y(+1)/Y(+1) = 0.78/(17.0 + 4.1) = 3.7$ %.

Полученное уравнение регрессии первого порядка в натуральных значениях фактора x с учетом того, что $X_1 = (x_1 - x_{1o})/\Delta x_1 = (x_1 - 6)/6$ (п. 10) имеет следующий вид:

$$Y = 17.00 + 4.1X_1 = 17.00 + 4.1(x_1 - 6)/6 = 12.9 + 0.68x_1,\ \text{г/сут}.$$

Интерпретация полученного уравнения.

1. Привес цыплят при использовании витаминной добавки в количестве от 0 до 12 г/сут. в период 10-дневного вскармливания («желторотый» возраст) описывается уравнением регрессии первого порядка $Y = 12.86 + 0.69x_1$.

2. Из полученного уравнения регрессии следует, что максимальный привес цыплят при использовании витаминной добавки в рационе цыплят составляет $Y_{\max} = 12.86 + 0.69 \cdot 12 = 12.87 + 8.28 = 21.1$ г/сут. при $x_1 = 12$ г/сут. Привес за счет введения витаминной добавки составляет 21.1 – 12.9 = 8.2 г/сут. или 63.6 %, так как суточный привес цыплят без добавки составляет 12.9 г/сут. ($x_1 = 0$).

3. Уравнение $Y = 12.9 + 0.68x_1$ справедливо для фактора x_1, который изменяется в интервале от 0 до 12 г/сут. Только в этом интервале значений фактора интерполяция полученного уравнения корректна (интерполяция – применение уравнения внутри исследованного факторного пространства).

4. При экстраполяции полученного уравнения расчетные значения Y могут значительно отличаться от экспериментального (экстраполяция – применение уравнения вне исследованного факторного пространства). Например, если добавить в рацион цыплят 100 г/сут. витаминной добавки, то по уравнению получается, что суточный при-

вес должен составить $Y = 12.9 + 0.68 \cdot 100 \approx 82$ г/сут. Это значит, что на 12-й день жизни «желторотые» цыплята должны были бы весить $50.1 + 82 \cdot 10 = 870$ г ≈ 0.9 кг (!), что нереально. Экстраполяция результатов полученного уравнения на $x_1 = 100$ г/сут, существенно удаленного от исследованного факторного пространства, некорректна и привела к значительной ошибке. Очевидно, что функция, отражающая привес цыплят при вскармливании, не может быть линейной в сколь угодно большом увеличении фактора. Для того чтобы выяснить, как влияет витаминная добавка при величинах добавок более 12 г/сут, необходимо расширить интервал варьирования фактора и использовать математическую модель второго порядка (см. алгоритм решения прикладных задач в разделе 2.4.2 и продолжение решения задачи в разделе 2.4.3).

2.3.4. Материалы для самостоятельной работы, лабораторных и практических занятий

Тесты

1. Перевод натуральных значений фактора x_{1j} в нормированные X_{1j}, $j = 1, \ldots, N$, осуществляется по следующим уравнениям:

а) $X_{1j} = \dfrac{x_{1j} - \Delta x_1}{x_{1о}}$; б) $X_{1j} = \dfrac{x_{1j} - x_{1о}}{\Delta x_1}$;

в) $X_{1j} = \dfrac{2x_{1j} - (x_{1\max} + x_{1\min})}{x_{1\max} - x_{1\min}}$; г) $X_j = \dfrac{2x_{1j} - (x_{1\max} - x_{1\min})}{x_{1\max} + x_{1\min}}$.

2. Диапазон изменения натуральных значений фактора $x_1 \in [x_{1\min}, x_{1\max}]$. Равномерно симметричный план эксперимента из N опытов, выраженный в натуральных значений фактора, рассчитывается по формулам:

а) $x_{1j} = x_{1\min} + \dfrac{j-1}{N-1}(x_{1\max} - x_{1\min})$; б) $x_{1j} = x_{1\min} + \dfrac{j}{N}(x_{1\max} - x_{1\min})$;

в) $x_{1j} = x_{1\max} - \dfrac{j-1}{N-1}(x_{1\max} - x_{1\min})$; г) $x_{1j} = x_{1\min} + \dfrac{j-1}{N}(x_{1\max} - x_{1\min})$.

3. Диапазон изменения нормированного фактора $X_1 \in [-1, +1]$. Равномерно симметричный план эксперимента из N опытов, выраженный в нормированных значениях фактора, рассчитывается по формулам:

а) $X_{1j} = 2 \cdot \dfrac{j-1}{N-1}$; б) $X_{1j} = 2 \cdot \dfrac{j-1}{N-1} - 1$;

в) $X_{1j} = \dfrac{j-1}{N-1} - 1$; г) $X_{1j} = 2 \cdot \dfrac{j}{N} - 1$.

4. Дисперсия воспроизводимости:

а) $S^2_{\text{воспр}} = \dfrac{\sum_{j=1}^{N} S_j^2}{N}$; б) $S^2_{\text{воспр}} = \dfrac{\sum_{j=1}^{N}\sum_{i=1}^{n}(Y_{ji} - \overline{Y}_j)^2}{N(n-1)}$;

в) $S^2_{\text{воспр}} = \dfrac{\sum_{j=1}^{N}\sum_{i=1}^{n}(Y_{ji} - \overline{Y}_j)^2}{n(N-1)}$; г) $S^2_{\text{воспр}} = \dfrac{\sum_{j=1}^{N} S_j^2}{n}$.

5. Какой план является ортогональным?

а)

N	x_{1j}, °C	X_{0j}	X_{1j}	$X_{0j}X_{1j}$
1	45	1	–1.00	–1.00
2	50	1	–0.75	–0.75
3	55	1	–0.50	–0.50
4	58	1	–0.35	–0.35
5	65	1	0.00	0.00
6	68	1	0.06	0.06
7	75	1	0.50	0.50
8	80	1	0.75	0.75
9	85	1	1.00	1.00

б)

N	x_{1j}, °C	X_{0j}	X_{1j}	$X_{0j}X_{1j}$
1	45	1	–1.00	–1.00
2	50	1	–0.75	–0.75
3	55	1	–0.50	–0.50
4	60	1	–0.25	–0.25
5	65	1	0.00	0.00
6	70	1	0.25	0.25
7	75	1	0.50	0.50
8	80	1	0.75	0.75
9	85	1	1.00	1.00

в)

N	x_{1j}, °C	X_{0j}	X_{1j}	$X_{0j}X_{1j}$
1	45	1	–1.00	–1.00
2	50	1	–0.75	–0.75
3	55	1	–0.50	–0.50
4	65	1	0.00	0.00
5	65	1	0.00	0.00
6	65	1	0.00	0.00
7	75	1	0.50	0.50
8	80	1	0.75	0.75
9	85	1	1.00	1.00

6. Коэффициенты однофакторного уравнения регрессии первого порядка рассчитываются по формулам:

$$\text{а) } b_0=\frac{\sum\limits_{j=1}^{N}X_{0j}\overline{Y}_j}{N},\ b_1=\frac{\sum\limits_{j=1}^{N}X_{1j}\overline{Y}_j}{N};\quad \text{б) } b_0=\frac{\sum\limits_{j=1}^{N}X_{0j}\overline{Y}_j}{\sum\limits_{j=1}^{N}X_{0j}^2},\ b_1=\frac{\sum\limits_{j=1}^{N}X_{1j}\overline{Y}_j}{\sum\limits_{j=1}^{N}X_{1j}^2};$$

$$\text{в) } b_0=\frac{\sum\limits_{j=1}^{N}X_{0j}\overline{Y}_j}{\sum\limits_{j=1}^{N}X_{0j}},\ b_1=\frac{\sum\limits_{j=1}^{N}X_{1j}\overline{Y}_j}{\sum\limits_{j=1}^{N}X_{1j}};\quad \text{г) } b_0=\frac{\sum\limits_{j=1}^{N}\overline{Y}_j}{\sum\limits_{j=1}^{N}X_{0j}^2},\ b_1=\frac{\sum\limits_{j=1}^{N}X_{1j}\overline{Y}_j}{\sum\limits_{j=1}^{N}X_{1j}^2}.$$

7. Если в однофакторном эксперименте $X_{1j}=2(j-1)/(N-1)-1$, то:

$$\text{а) } \sum_{j=1}^{N}X_{0j}^2=N^2,\qquad \sum_{j=1}^{N}X_{1j}^2=\frac{(N+1)}{3(N-1)};$$

$$\text{б) } \sum_{j=1}^{N}X_{0j}^2=0,\qquad \sum_{j=1}^{N}X_{1j}^2=\frac{N}{3(N-1)},$$

$$\text{в) } \sum_{j=1}^{N}X_{0j}^2=N,\qquad \sum_{j=1}^{N}X_{1j}^2=\frac{N(N+1)}{3(N-1)},$$

$$\text{г) } \sum_{j=1}^{N}X_{0j}^2=\frac{N}{2},\qquad \sum_{j=1}^{N}X_{1j}^2=\frac{N(N+1)}{3}.$$

8. Дисперсии значимости коэффициентов регрессии первого порядка в однофакторном эксперименте:

$$\text{а) } S^2(b_0)=\frac{S^2_{\text{воспр}}}{n\sum\limits_{j=1}^{N}X_{0j}^2},\qquad S^2(b_1)=\frac{S^2_{\text{воспр}}}{n\sum\limits_{j=1}^{N}X_{1j}^2};$$

$$\text{б) } S^2(b_0)=\frac{S^2_{\text{воспр}}}{N\sum\limits_{j=1}^{N}X_{0j}^2},\qquad S^2(b_1)=\frac{S^2_{\text{воспр}}}{N\sum\limits_{j=1}^{N}X_{1j}^2};$$

$$\text{в) } S^2(b_0)=\frac{S^2_{\text{воспр}}}{(n\text{-}1)\sum\limits_{j=1}^{N}X_{0j}^2},\qquad S^2(b_1)=\frac{S^2_{\text{воспр}}}{(n\text{-}1)\sum\limits_{j=1}^{N}X_{1j}^2};$$

$$\text{г) } S^2(b_0)=\frac{S^2_{\text{воспр}}}{nN\sum\limits_{j=1}^{N}X_{0j}^2},\qquad S^2(b_1)=\frac{S^2_{\text{воспр}}}{nN\sum\limits_{j=1}^{N}X_{1j}^2}.$$

9. Доверительный интервал коэффициентов регрессии первого порядка в однофакторном эксперименте (m = 0, 1):

$$\text{а) } \Delta b_q=t_{N(n-1)}S^2(b_q);\qquad \text{б) } \Delta b_q=t_{N(n-1)}S(b_q);$$

$$\text{в) } \Delta b_m=t_{(N-1)n}S(b_q);\qquad \text{г) } \Delta b_q=t_{N(n-1)}\frac{S(b_q)}{\sqrt{n}}.$$

10. Остаточная сумма квадратов:

$$\text{а) } \varphi=\sum_{j=1}^{N}(\overline{Y}_j-Y_j^p);\qquad \text{б) } \varphi=\sum_{j=1}^{N}(\overline{Y}_j-Y_j^p)^2;$$

$$\text{в) } \varphi=\sum_{j=1}^{N}(Y_j^p-\overline{Y}_j)^2;\qquad \text{г) } \varphi=\sum_{j=1}^{N}\left(\overline{Y}_j^2-Y_j^{p2}\right).$$

11. Дисперсия адекватности:

$$\text{а) } S^2_{\text{ад}}=\frac{n\varphi}{N-1};\qquad \text{б) } S^2_{\text{ад}}=\frac{n\varphi}{N-(B-1)};$$

$$\text{в) } S^2_{\text{ад}}=\frac{n\varphi}{N-B};\qquad \text{г) } S^2_{\text{ад}}=\frac{(n-1)\varphi}{N-B}.$$

12. Уравнение регрессии первого порядка адекватно, если

а) $F_{э} = \dfrac{S^2_{ад}}{S^2_{воспр}} < F_{N-B,\,N(n-1),\,p}$ при $S^2_{ад} > S^2_{воспр}$;

б) $F_{э} = \dfrac{S^2_{ад}}{S^2_{воспр}} > F_{N-B,\,N(n-1),\,p}$ при $S^2_{ад} > S^2_{воспр}$;

в) $F_{э} = \dfrac{S^2_{воспр}}{S^2_{ад}} < F_{N(n-1),\,N-B,\,p}$ при $S^2_{ад} < S^2_{воспр}$;

г) $F_{э} = \dfrac{S^2_{воспр}}{S^2_{ад}} < F_{N-B,\,N(n-1),\,p}$ при $S^2_{ад} < S^2_{воспр}$.

Вопросы для самопроверки

1. Какие требования предъявляются к параметру?
2. Какие требования предъявляются к фактору?
3. Натуральные и нормированные значения факторов. Формулы перехода.
4. Какими свойствами должен обладать план однофакторного эксперимента для того, чтобы можно было независимо определять коэффициенты уравнения регрессии первого порядка?
5. Какими свойствами обладает РСП?
6. Напишите математическое уравнение, выражающее ортогональность факторов в однофакторном эксперименте.
7. Что общего и в чем различие в математических формулах

$$\eta = \sum_{q=o}^{m} \beta_q X_q \quad \text{и} \quad Y = \sum_{q=o}^{m} b_q X_q \ ?$$

8. Напишите формулы расчета коэффициентов уравнения регрессии первого порядка для однофакторного эксперимента.
9. Напишите формулу расчета дисперсии воспроизводимости (N опытов, n дублей). Чему равно число степеней свободы для этого параметра?
10. Напишите формулы расчета дисперсий значимости коэффициентов уравнения регрессии первого порядка для однофакторного эксперимента.

11. Напишите формулы расчета доверительных интервалов коэффициентов уравнения регрессии первого порядка для однофакторного эксперимента.
12. Сформулируйте условие значимости коэффициентов регрессии?
13. Напишите формулу для остаточной суммы квадратов.
14. Напишите формулу расчета дисперсии адекватности (N опытов, n дублей). Чему равно число степеней свободы для этого параметра?
15. Сформулируйте критерий адекватности уравнения регрессии любого порядка.
16. Напишите уравнение для определения предельной абсолютной погрешности параметра Y, рассчитанного по однофакторному уравнению регрессии первого порядка

Задачи для самостоятельного решения

Уровень 1

1. Докажите, что в однофакторном эксперименте (N опытов)

$$\sum_{j=1}^{N} X_{0j}^2 = N .$$

Уровень 2

2. Докажите, что в однофакторном эксперименте для РСП

$$\sum_{j=1}^{N} X_{1j}^2 = \frac{N(N+1)}{3(N-1)} .$$

3. Составьте РСП для однофакторного эксперимента, состоящего из 6 опытов, если фактором является температура в печи, варьируемая в диапазоне от 20 до 1000 °С.

Ответ:

N	x_{1j}, °С	X_{1j}
1	20	–1.0
2	216	–0.6
3	412	–0.2
4	608	0.2
5	804	0.6
6	1000	1.0

4. Температура сушильного агента в аппарате для сушки зерна варьируется в диапазоне $x \in [45, 95$ ºС]. Для исследования процесса сушки постройте РСП, состоящий из 11 опытов.

Ответ:

N	x_{1j}, °С	X_{1j}
1	45	–1.0
2	50	–0.8
3	55	–0.6
4	60	–0.4
5	65	–0.2
6	70	0.0
7	75	0.2
8	80	0.4
9	85	0.6
10	90	0.8
11	95	1.0

Уровень 3

5. Измерения ТЭДС термопары алюмель–хромель в температурном диапазоне $x \in (0, 520$ ºС) дали следующие результаты:

N	x_{1j}, °С	Y_{j1}, мВ	Y_{j2}, мВ	Y_{j3}, мВ
1	20	0.2	0.4	2.1
2	120	19.7	18.5	17.3
3	220	36.0	34.8	35.9
4	320	56.0	55.5	53.6
5	420	71.8	73.4	72.3
6	520	91.6	90.0	91.9

Постройте уравнение регрессии первого порядка для термопары алюмель-хромель в нормированных значениях фактора X_1. Проверьте коэффициенты уравнения на значимость, а само уравнение на адекватность.

Ответ: Нормированные значения фактора X_1 определяются следующим образом: $X_{1j} = (x_{1j} – 270)/250$. Уравнение регрессии первого порядка $Y = 45.61 + 45.2X_1$. Оба коэффициента уравнения значимы, так как $\Delta b_0 = 0.53$, $\Delta b_1 = 0.8$. Уравнение регрессии первого порядка адекватно, так как $F_э = 1.13$, а $F_{4,\ 12,\ 0.95} = 3.26$.

6. Измерения ТЭДС термопары алюмель-хромель в температурном диапазоне $x_1 \in (0, 620$ ºС) дали следующие результаты:

N	x_{1j}, °С	Y_{j1}, мВ	Y_{j2}, мВ	Y_{j3}, мВ
1	20	0.1	0.0	0.2
2	140	19.1	17.3	16.2
3	260	36.2	36.0	36.4
4	380	57.7	56.1	54.3
5	500	76.0	75.6	76.7
6	620	96.4	97.9	96.7

Постройте уравнение регрессии первого порядка для термопары алюмель-хромель в нормированных значениях фактора X. Проверьте коэффициенты уравнения на значимость, а само уравнение на адекватность.

Ответ: Нормированные значения фактора X_1 определяются следующим образом: $X_{1j} = (x_{1j} – 320)/300$. Уравнение регрессии первого порядка $Y = 47.16 + 48.6 \cdot X_1$. Оба коэффициента уравнения значимы, так как $\Delta b_0 = 0.51$, $\Delta b_1 = 0.8$. Уравнение регрессии первого порядка неадекватно, так как $F_э = 4.80$, а $F_{4,\ 12,\ 0.95} = 3.26$.

2.4. Уравнение регрессии второго порядка (тема 6)

2.4.1. Теория

Если уравнение регрессии первого порядка неадекватно, то исследуемый объект целесообразно моделировать уравнением регрессии второго порядка в обычной форме ($m = 2$):

$$Y = b_0 + b_1 X + b_{11} X_1^2. \qquad (98)$$

Для определения коэффициентов уравнения регрессии второго порядка b_0, b_1, b_{11} с помощью МНК составляется система уравнений (в данном случае - три уравнения с тремя неизвестными (по аналогии с уравнениями (63), (64)). Однако, система уравнений для определения b_0, b_1, b_{11}, полученная по уравнению регрессии второго порядка в форме (98), не будет иметь диагональный вид, так как добавленный к уравнению регрессии первого порядка $Y = b_0X_0 + b_1X_1$ фактор X_1^2 не ортогонален фактору X_0:

$$\sum_{j=1}^{N} X_{0j} X_{1j}^2 = \sum_{j=1}^{N} X_{1j}^2 > 0. \qquad (99)$$

Поскольку для РСП факторы X_0 и X_1 ортогональны (см. уравнение (66)), то для ортогонализации квадратичного фактора X_1^2 с X_0 и X_1 представим его в виде квадратного многочлена:

$$X_1^2 \to X_1^2 - \gamma_1 X_1 - \lambda_1, \qquad (100)$$

где γ_1, λ_1 – константы.

Значения констант γ_1 и λ_1 подбираются из условия взаимной ортогональности всех трех факторов: X_0, X_1, $X_1^2 - \gamma_1 X_1 - \lambda_1$:

$$\begin{cases} \sum_{j=1}^{N} X_{0j}(X_{1j}^2 - \gamma_1 X_{1j} - \lambda_1) = 0; \\ \sum_{j=1}^{N} X_{1j}(X_{1j}^2 - \gamma_1 X_{1j} - \lambda_1) = 0. \end{cases} \qquad (101)$$

В результате решения уравнений (101) относительно γ_1 и λ_1 с учетом того, что используется РСП, задаваемый уравнением (68), получим $\gamma_1 = 0$ и $\lambda_1 = \overline{X_1^2}$ (вывод представлен в приложении 6).

Таким образом, преобразованный фактор $X_1^2 - \lambda_1$ ортогонален факторам X_0 и X_1, а ортогонализирующий коэффициент λ_1 с учетом РСП эксперимента, задаваемого уравнением (72), имеет следующий вид:

$$\lambda_1 = \overline{X_1^2} = \frac{\sum_{j=1}^{N} X_{1j}^2}{N} = \frac{N+1}{3(N-1)}. \qquad (102)$$

Так как факторы X_0, X_1, $X_1^2 - \lambda_1$ взаимно ортогональны, то для соответствующего им ортогонализированного уравнения регрессии второго порядка

$$Y = b_0^* X_0 + b_1^* X_1 + b_{11}^* (X_1^2 - \lambda_1) \qquad (103)$$

система уравнений для нахождения регрессионных коэффициентов b_0^*, b_1^*, b_{11}^*, полученная МНК, имеет диагональный вид. Коэффициенты ортогонализированного уравнения регрессии второго порядка (уравнение (103)) помечены звездочкой, чтобы отличать их от регрессионных коэффициентов b_0, b_1, b_{11}, входящих в уравнение регрессии второго порядка в обычной форме (98).

Стандартная процедура нахождения коэффициентов уравнения регрессии (103) в этом случае будет аналогична процедуре, описанной выше (см. уравнения (62) – (65)):

$$\varphi = \sum_{j=1}^{N} \left[b_0^* X_{0j} + b_1^* X_{1j} + b_2^* \left(X_{1j}^2 - \lambda_1 \right) - \overline{Y}_j \right]^2 = \min. \qquad (104)$$

Минимум функции $\varphi = \varphi(b_0^*, b_1^*, b_2^*)$ достигается при следующем условии:

$$\begin{cases} \frac{\partial \varphi}{\partial b_0} = \sum_{j=1}^{N} 2\left[b_0^* X_{0j} + b_1^* X_{1j} + b_{11}^* \left(X_{1j}^2 - \lambda_1 \right) - \overline{Y}_j \right] X_{0j} = 0; \\ \frac{\partial \varphi}{\partial b_1} = \sum_{j=1}^{N} 2\left[b_0^* X_{0j} + b_1^* X_{1j} + b_{11}^* \left(X_{1j}^2 - \lambda_1 \right) - \overline{Y}_j \right] X_{1j} = 0; \\ \frac{\partial \varphi}{\partial b_{11}} = \sum_{j=1}^{N} 2\left[b_0^* X_{0j} + b_1^* X_{1j} + b_{11}^* \left(X_{1j}^2 - \lambda_1 \right) - \overline{Y}_j \right] \left(X_{1j}^2 - \lambda_1 \right) = 0. \end{cases} \qquad (105)$$

После преобразования система уравнений (105) с учетом ортогональности факторов (уравнения (66), (101)) примет следующий вид:

$$\begin{cases} b_0^* \sum_{j=1}^{N} X_{0j}^2 + 0 + 0 = \sum_{j=1}^{N} X_{0j} \overline{Y}_j \ ; \\ 0 + b_1^* \sum_{j=1}^{N} X_{1j}^2 + 0 = \sum_{j=1}^{N} X_{1j} \overline{Y}_j \ ; \\ 0 + 0 + b_{11}^* \sum_{j=1}^{N} \left(X_{1j}^2 - \lambda_1 \right)^2 = \sum_{j=1}^{N} \left(X_{1j}^2 - \lambda_1 \right) \overline{Y}_j . \end{cases} \qquad (106)$$

Из системы уравнений (106) следует, что

$$b_0^* = \frac{\sum_{j=1}^{N} X_{0j} \overline{Y}_j}{\sum_{j=1}^{N} X_{0j}^2}, \quad b_1^* = \frac{\sum_{j=1}^{N} X_{1j} \overline{Y}_j}{\sum_{j=1}^{N} X_{1j}^2}, \quad b_2^* = \frac{\sum_{j=1}^{N} \left(X_{1j}^2 - \lambda_1 \right) \overline{Y}_j}{\sum_{j=1}^{N} \left(X_{1j}^2 - \lambda_1 \right)^2}. \qquad (107)$$

Ортогонализированное уравнение второго порядка регрессии (103) можно преобразовать в уравнение в обычной форме (уравнение (98)):

$$Y = b_0^* + b_1^* X_1 + b_{11}^* \left(X_1^2 - \lambda_1 \right) = \left(b_0^* - \lambda_1 b_{11}^* \right) + b_1^* X_1 + b_{11}^* X_1^2 . \qquad (108)$$

Сравнивая уравнения (103) и (98), нетрудно заметить, что:

$$b_0 = b_0^* - \lambda_1 b_{11}^* ; \quad b_1 = b_1^* ; \quad b_{11} = b_{11}^* . \qquad (109)$$

Для однофакторного РСП эксперимента в дополнение к уже приведенным ранее формулам (см. уравнения (71), (72)), приведем еще одну полезную формулу (доказательство дано в приложении 6):

$$\sum_{j=1}^{N} \left(X_{1j}^2 - \lambda_1 \right)^2 = \frac{4N(N+1)(N^2-4)}{45(N-1)^3} . \qquad (110)$$

Уравнения (109) с учетом уравнений (107), (102), (71), (72), (110) примут следующий вид:

$$\begin{cases} b_0 = \frac{\sum_{j=1}^{N} X_{0j} \overline{Y}_j}{\sum_{j=1}^{N} X_{0j}^2} - \frac{\lambda_1 \sum_{j=1}^{N} \left(X_{1j}^2 - \lambda_1 \right) \overline{Y}_j}{\sum_{j=1}^{N} \left(X_{1j}^2 - \lambda_1 \right)^2} = \frac{\sum_{j=1}^{N} X_{0j} \overline{Y}_j}{N} - \frac{15(N-1)^2}{4(N^2-4)} \cdot \frac{\sum_{j=1}^{N} \left(X_{1j}^2 - \lambda_1 \right) \overline{Y}_j}{N} ; \\ b_1 = \frac{\sum_{j=1}^{N} X_{1j} \overline{Y}_j}{\sum_{j=1}^{N} X_{1j}^2} = \frac{3(N-1)}{N+1} \cdot \frac{\sum_{j=1}^{N} X_{1j} \overline{Y}_j}{N} ; \\ b_{11} = \frac{\sum_{j=1}^{N} \left(X_{1j}^2 - \lambda_1 \right) \overline{Y}_j}{\sum_{j=1}^{N} \left(X_{1j}^2 - \lambda_1 \right)^2} = \frac{45(N-1)^3}{4(N+1)(N^2-4)} \cdot \frac{\sum_{j=1}^{N} \left(X_{1j}^2 - \lambda_1 \right) \overline{Y}_j}{N} . \end{cases} \qquad (111)$$

Отметим, что обходной маневр для нахождения коэффициентов b_0, b_1, b_{11} уравнения регрессии второго порядка в обычной форме (98), связанный с ортогонализацией квадратичного фактора, был предпринят для существенного упрощения процедуры расчета и, что самое главное, независимого определения этих коэффициентов.

Проверка полученного уравнения регрессии второго порядка на статистическую значимость (проверка коэффициентов регрессии на значимость, а само уравнение на адекватность) во многом повторяет аналогичную процедуру с уравнением регрессии первого порядка.

Проверка коэффициентов регрессии на значимость по критерию Стьюдента**.** Дисперсия воспроизводимости $S^2_{\text{воспр}}$ рассчитывается так же, как и для уравнения регрессии первого порядка (см. уравнение (74)).

Формулы расчета дисперсий значимости коэффициентов уравнения регрессии второго порядка b_0, b_1, b_{11} имеют некоторое отличие от формул расчета дисперсий значимости коэффициентов уравнения регрессии первого порядка (см. уравнения (75)). Сначала рассчитывают дисперсии значимости коэффициентов ортогонализированного уравнения регрессии (103) b_0^*, b_1^*, b_{11}^* по формулам аналогичным уравнениям (75):

$$S^2(b_0^*) = \frac{S^2_{\text{воспр}}}{n \sum_{j=1}^{N} X_{0j}^2}, \ S^2(b_1^*) = \frac{S^2_{\text{воспр}}}{n \sum_{j=1}^{N} X_{1j}^2}, \ S^2(b_{11}^*) = \frac{S^2_{\text{воспр}}}{n \sum_{j=1}^{N} \left(X_{1j}^2 - \lambda_1 \right)^2}, \qquad (112)$$

или для РСП с учетом уравнений (71), (72), (110):

$$\begin{cases} S^2(b_0^*) = \dfrac{S^2_{\text{воспр}}}{Nn}; \\ S^2(b_1^*) = \dfrac{3(N-1)}{N+1} \cdot \dfrac{S^2_{\text{воспр}}}{Nn}; \\ S^2(b_{11}^*) = \dfrac{45(N-1)^3}{4(N+1)(N^2-4)} \cdot \dfrac{S^2_{\text{воспр}}}{Nn}. \end{cases} \tag{113}$$

Так как между коэффициентами регрессии b_0, b_1, b_{11} и b_0^*, b_1^*, b_{11}^* существует взаимосвязь, приведенная в уравнении (109), то, используя уравнения (36) и (37), с учетом того, что величины b_0^*, b_1^*, b_{11}^* независимы, получим:

$$\begin{cases} S^2(b_0) = S^2(b_0^*) + \lambda_1^2 S^2(b_{11}^*); \\ S^2(b_1) = S^2(b_1^*); \\ S^2(b_{11}) = S^2(b_{11}^*). \end{cases} \tag{114}$$

Уравнения (114) с учетом уравнений (102), (113) примут следующий вид:

$$\begin{cases} S^2(b_0) = \dfrac{S^2_{\text{воспр}}}{n\sum\limits_{j=1}^{N} X_{0j}^2} + \dfrac{\lambda_1^2 S^2_{\text{воспр}}}{n\sum\limits_{j=1}^{N}\left(X_{1j}^2 - \lambda_1\right)^2} = \left[1 + \dfrac{5(N^2-1)}{4(N^2-4)}\right] \cdot \dfrac{S^2_{\text{воспр}}}{Nn}; \\ S^2(b_1) = \dfrac{S^2_{\text{воспр}}}{n\sum\limits_{j=1}^{N} X_{1j}^2} = \dfrac{3(N-1)}{N+1} \cdot \dfrac{S^2_{\text{воспр}}}{Nn}; \\ S^2(b_{11}) = \dfrac{S^2_{\text{воспр}}}{n\sum\limits_{j=1}^{N}\left(X_{1j}^2 - \lambda_1\right)^2} = \dfrac{45(N-1)^3}{4(N+1)(N^2-4)} \cdot \dfrac{S^2_{\text{воспр}}}{Nn}. \end{cases} \tag{115}$$

Так как число опытов $N >> 1$, то уравнения (115) имеют наглядный вид:

$$\begin{cases} S^2(b_0) = 2.25 \cdot \dfrac{S^2_{\text{воспр}}}{Nn}; \\ S^2(b_1) = 3 \cdot \dfrac{S^2_{\text{воспр}}}{Nn}; \\ S^2(b_{11}) = 11.25 \cdot \dfrac{S^2_{\text{воспр}}}{Nn}. \end{cases} \tag{116}$$

Процедура построения доверительных интервалов для каждого коэффициента уравнения регрессии второго порядка полностью совпадает с процедурой построения доверительных интервалов для коэффициентов уравнения регрессии первого порядка (см. уравнение (78)).

Проверка уравнения регрессии второго порядка на адекватность по критерию Фишера полностью совпадает с аналогичной процедурой для однофакторного уравнения регрессии первого порядка (см. уравнения (79)–(86)). Сделаем только одно уточнение: в случае неадекватности однофакторного уравнении регрессии второго порядка целесообразно перейти к построению уравнения регрессии третьего порядка, однако в данном учебно-методическом комплексе построение уравнений регрессии третьего порядка рассматривается.

Уравнение регрессии второго порядка в ортогонализированном виде (103) позволяет оценить предельную абсолютную погрешность рассчитанного параметра $Y(X_1)$ в исследованном факторном пространстве. Факторы X_0, X_1, X_1^2 уравнения регрессии второго порядка в обычной форме $Y = b_0X_0 + b_1X_1 + b_{11}X_1^2$ неортогональны (см. уравнение (99)) и поэтому регрессионные коэффициенты b_0, b_1, b_{11} зависимы (см. уравнение (109)). Поэтому применять уравнение регрессии второго порядка в обычной форме (37) для расчета дисперсии $S^2(Y)$ нельзя. А вот факторы X_0, X_1, $X_1^2 - \lambda_1$ ортогонализированного уравнения регрессии $Y = b_0^* + b_1^* X + b_{11}^*(X_1^2 - \lambda_1)$ ортогональны, и поэтому коэффициенты b_0^*, b_1^*, b_{11}^* независимы. Применяя к ортогонализированному уравнению регрессии второго порядка уравнения (36) и (37) получим:

$$\Delta Y = t_{N(n-1),\,p} \cdot S(Y)\,, \qquad (117)$$

где $S^2(Y) = S^2(b_0^*) + X_1^2 S^2(b_1^*) + (X_1^2 - \lambda_1)^2 S^2(b_{11}^*)\,.$ (118)

С учетом уравнений ((113), (102) перепишем уравнение (118):

$$S^2(Y) = \frac{S^2_{\text{воспр}}}{Nn}\left[1 + \frac{3(N-1)}{N+1}X_1^2 + \frac{45(N-1)^3}{4(N+1)(N^2-4)}\left(X_1^2 - \frac{N+1}{3(N-1)}\right)^2\right]. \quad (119)$$

Подставляя уравнение (119) в (117), получим:

$$\Delta Y(X_1) = t_{N(n-1),\,p}\frac{S_{\text{воспр}}}{\sqrt{Nn}}\sqrt{1 + \frac{3(N-1)}{N+1}X_1^2 + \frac{45(N-1)^3}{4(N+1)(N^2-4)}\left(X_1^2 - \frac{N+1}{3(N-1)}\right)^2}\,. \quad (120)$$

Анализ уравнения (120) позволяет сделать несколько выводов:

а) минимальная предельная абсолютная погрешность достигается не в центре плана $X_1 = 0$, как для уравнения регрессии первого порядка (см. уравнение (92)), а в точках $X_1 = \pm\sqrt{(N^2+1)/5(N-1)^2}$, а максимальная - на границах факторного пространства $X_1 = \pm 1$ (вывод приведен в приложении 6):

$$\Delta Y\left(\pm\sqrt{\frac{N^2+1}{5(N-1)^2}}\right) \approx 1.34 \cdot t_{N(n-1),\,p}\frac{S_{\text{воспр}}}{\sqrt{Nn}}\sqrt{\frac{N^2-2/3}{N^2-1}} \approx 1.34 \cdot t_{N(n-1),\,p} \cdot \frac{S_{\text{воспр}}}{\sqrt{Nn}}\,, \quad (121)$$

$$\Delta Y(0) = 1.5 \cdot t_{N(n-1),\,p}\frac{S_{\text{воспр}}}{\sqrt{Nn}}\sqrt{\frac{N^2-7/3}{N^2-4}} \approx 1.5 \cdot t_{N(n-1),\,p}\frac{S_{\text{воспр}}}{\sqrt{Nn}}\,, \quad (122)$$

$$\Delta Y(\pm 1) = 3 \cdot t_{N(n-1),\,p}\frac{S_{\text{воспр}}}{\sqrt{Nn}}\sqrt{\frac{N^2-N+2/3}{N^2+3N+2)}} \approx 3 \cdot t_{N(n-1),\,p}\frac{S_{\text{воспр}}}{\sqrt{Nn}}\left(1 - \frac{2}{N}\right). \quad (123)$$

б) предельная абсолютная погрешность параметра $Y(X_1)$, рассчитанная по адекватному уравнению регрессии второго порядка в ортогонализированной форме, также как и для уравнения регрессии первого порядка, значительно меньше предельной абсолютной погрешности, найденной по экспериментальным данным в каждом опыте.

Подтвердим количественно это утверждение. Предельная абсолютная погрешность параметра $\overline{Y}_j$ в каждом опыте рассчитывается по уравнению (95). Соотношение предельных экспериментальных и

расчетных ошибок в центре плана $X_1 = 0$, используя уравнения (122) и (95), равно

$$\frac{\Delta \overline{Y}_j}{\Delta Y(0)} \approx \frac{t_{n-1,\,p} \cdot S_{\text{воспр}} \cdot \sqrt{Nn}}{\sqrt{n} \cdot t_{N(n-1),\,p} \cdot 1.5 \cdot S_{\text{воспр}}} = \frac{t_{n-1,\,p}\sqrt{N}}{1.5 \cdot t_{N(n-1),\,p}}\,. \quad (124)$$

Так как в экспериментах $N \approx 10$, $n \approx 4$, то $t_{30,\,0.95} \approx 2$, $t_{3,\,0.95} \approx 3$ (см. приложение 2) и из уравнения (124) следует, что в центре плана

$$\frac{\Delta \overline{Y}_j}{\Delta Y(0)} = \frac{3 \cdot \sqrt{10}}{1.5 \cdot 2} \approx 3.2\,. \quad (125)$$

Из уравнения (125) следует, что предельная абсолютная погрешность параметра Y в центе плана, рассчитанная по однофакторному уравнению регрессии второго порядка, построенному по РСП, в 3.2 раза меньше предельной абсолютной погрешности выборочных средних, найденных экспериментально в каждом опыте.

Таким образом, и в этом случае математическая модель объекта позволяет делать прогноз точнее, чем породивший ее эксперимент.

Если уравнение регрессии второго порядка адекватно, то следует найти максимум (минимум) исследуемого параметра. Необходимым условием максимума (минимума) является равенство нулю первой производной уравнения регрессии второго порядка в обычной форме $Y = b_0 + b_1 X_1 + b_{11} X_1^2$:

$$Y' = b_1 + 2b_2 X_1 = 0\,. \quad (126)$$

Достаточное условие максимума - $Y'' < 0$ (для уравнения регрессии второго порядка это условие эквивалентно условию $b_{11} < 0$). Достаточное условие минимума - $Y'' > 0$ (для уравнения регрессии второго порядка это условие эквивалентно условию $b_{11} > 0$). Из уравнения (126) следует, что значение фактора, при котором уравнение регрессии достигает максимум (минимум), равно $X_{\text{опт}} = -\, b_1/2b_{11}$, а само максимальное (минимальное) значение параметра определяется следующим выражением:

$$Y_{\max}\ (Y_{\min}) = b_0 + b_1\left(-\frac{b_1}{2b_2}\right) + b_2\left(-\frac{b_1}{2b_2}\right)^2 = b_0 - \frac{b_1^2}{4b_2}\,. \quad (127)$$

Если значение фактора $X_{\text{опт}}$, при котором исследуемый параметр достигает максимум (минимум), лежит вне исследованного факторного пространства, то погрешность параметра может оказаться больше рассчитанной величины. В этом случае для проверки необходимо провести дополнительные исследования.

2.4.2. Алгоритм решения прикладных задач

1. Выбрать объект исследования и на основе литературных и патентных источников собрать необходимую априорную информацию о факторах и параметрах, характеризующих исследуемый объект.

2. Выбрать наиболее существенный параметр Y и n – число дублей в каждом опыте. Рекомендуемое число дублей не менее 4.

3. Выбрать наиболее существенный фактор x_1 и интервал его варьирования $[x_{1\min}, x_{1\max}]$, а также N – число опытов. Рекомендуемое число опытов для $m = 2$ не менее $N \geq 2(m + 1) = 6$.

3.1. Для РСП рассчитать натуральные значения фактора x_{1j}:

$$x_{1j} = x_{1\min} + \frac{j-1}{N-1} \cdot (x_{1\max} - x_{1\min}),\ j = 1, \ldots, N.$$

3.2. Для РСП рассчитать нормированные значения фактора X_{1j}:

$$X_{1j} = 2 \cdot \frac{j-1}{N-1} - 1,\ j = 1, \ldots, N.$$

4. Обеспечить приборное и технологическое оснащение эксперимента, позволяющее измерять исследуемый параметр Y и варьируемый фактор x_1 с требуемой точностью.

5. Провести предварительную обработку экспериментальных данных.

5.1. Рассчитать выборочное среднее в каждом опыте:

$$\overline{Y}_j = \frac{\sum_{i=1}^{n} Y_{ji}}{n},\ j = 1, \ldots, N.$$

5.2. Рассчитать выборочную дисперсию в каждом опыте:

$$S_j^2 = \frac{\sum_{i=1}^{n} \left(Y_{ji} - \overline{Y}_j\right)^2}{n-1},\ j = 1, \ldots, N..$$

5.3. Экспериментальные данные каждого опыта проверить на промахи по критерию Смирнова–Граббса (приложение 1). Опыты, в которых выявлены промахи, переделать:

$$\tau_{э} = \frac{\max\left|Y_{ji} - \overline{Y}_j\right|}{S} > \tau_{n\text{-}2,\,p},\ j = 1, \ldots, N,$$

где $\tau_{n-2,\,p}$ – критическое значение критерия Смирнова–Граббса (приложение 1).

5.4. Экспериментальные данные каждого опыта проверить на принадлежность их к нормальному закону распределения:

$$\left|\frac{\text{САО}_j}{S_j} - 0.7979\right| = \left|\frac{\sum_{i=1}^{n}\left|Y_{ji} - \overline{Y}_j\right|}{nS_j} - 0.7979\right| < \frac{0.4}{\sqrt{n}},\ j = 1, \ldots, N.$$

5.5. Дисперсии всех опытов проверить на однородность по критерию Кохрена:

$$\frac{\max S_j^2}{\sum_{j=1}^{N} S_j^2} < G_{n\text{-}1,\,N,\,p},$$

где $G_{n\text{-}1,\,N,\,p}$ – критическое значение критерия Кохрена (приложение 5).

Внимание! Дальнейшая обработка экспериментальных данных возможна только в том случае, если в каждом опыте нет промахов, случайные значения параметра в каждом опыте подчиняются нормальному закону распределения, а дисперсии всех опытов однородны.

6. Рассчитать коэффициенты b_0, b_1, b_{11} уравнения регрессии второго порядка $Y = b_0 X_0 + b_1 X_1 + b_{11} X_1^2$.

6.1. Образовать столбцы $X_{0j}\overline{Y}_j$, $X_{1j}\overline{Y}_j$, $(X_{1j}^2 - \lambda_1)\overline{Y}_j$ и рассчитать их суммы, где $\lambda_1 = \dfrac{\sum_{j=1}^{N} X_{1j}^2}{N} = \dfrac{N+1}{3(N-1)}$.

6.2. Рассчитать коэффициенты регрессии b_0, b_1, b_{11}:

$$b_0 = \frac{\sum_{j=1}^{N} X_{0j}\overline{Y}_j}{\sum_{j=1}^{N} X_{oj}^2} - \lambda_1 \frac{\sum_{j=1}^{N}\left(X_{1j}^2 - \lambda_1\right)\overline{Y}_j}{\sum_{j=1}^{N}\left(X_{1j}^2 - \lambda_1\right)^2} = \frac{\sum_{j=1}^{N} X_{0j}\overline{Y}_j}{N} - \frac{15(N-1)^2}{4(N^2-4)} \frac{\sum_{j=1}^{N}\left(X_{1j}^2 - \lambda_1\right)\overline{Y}_j}{N},$$

$$b_1 = \frac{\sum_{j=1}^{N} X_{1j}\overline{Y}_j}{\sum_{j=1}^{N} X_{1j}^2} = \frac{3(N-1)}{N+1} \cdot \frac{\sum_{j=1}^{N} X_{1j}\overline{Y}_j}{N},$$

$$b_{11} = \frac{\sum_{j=1}^{N}\left(X_{1j}^2 - \lambda_1\right)\overline{Y}_j}{\sum_{j=1}^{N}\left(X_{1j}^2 - \lambda_1\right)^2} = \frac{45(N-1)^3}{4(N+1)(N^2-4)} \cdot \frac{\sum_{j=1}^{N}\left(X_{1j}^2 - \lambda_1\right)\overline{Y}_j}{N}.$$

7. Проверить коэффициенты b_0, b_1, b_{11} уравнения регрессии второго порядка вида $Y = b_0 X_0 + b_1 X_1 + b_{11} X_1^2$ на значимость.

7.1. Рассчитать дисперсию воспроизводимости:

$$S_{\text{воспр}}^2 = \frac{\sum_{j=1}^{N} S_j^2}{N}, f = N(n-1).$$

7.2. Рассчитать дисперсии значимости коэффициентов b_0, b_1, b_{11}:

$$\begin{cases} S^2(b_0) = \dfrac{S_{\text{воспр}}^2}{n\sum_{j=1}^{N} X_{0j}^2} + \lambda_1^2 \dfrac{S_{\text{воспр}}^2}{n\sum_{j=1}^{N}\left(X_{1j}^2 - \lambda_1\right)^2} = \left[1 + \dfrac{5(N^2-1)}{4(N^2-4)}\right] \cdot \dfrac{S_{\text{воспр}}^2}{Nn}; \\ S^2(b_1) = \dfrac{S_{\text{воспр}}^2}{n\sum_{j=1}^{N} X_{1j}^2} = \dfrac{3(N-1)}{N+1} \cdot \dfrac{S_{\text{воспр}}^2}{Nn}; \\ S^2(b_{11}) = \dfrac{S_{\text{воспр}}^2}{n\sum_{j=1}^{N}\left(X_{1j}^2 - \lambda_1\right)^2} = \dfrac{45(N-1)^3}{4(N+1)(N^2-4)} \cdot \dfrac{S_{\text{воспр}}^2}{Nn}. \end{cases}$$

7.3. Рассчитать доверительный интервал для коэффициентов b_0, b_1, b_{11}:

$$\Delta b_0 = t_{N(n-1),\, p} S(b_0), \ \Delta b_1 = t_{N(n-1),\, p} S(b_1), \ \Delta b_{11} = t_{N(n-1),\, p} S(b_{11}),$$

где $t_{N(n\text{-}1),\, p}$ – критическое значение критерия Стьюдента (приложение 2).

7.4. Коэффициенты, для которых $\Delta b > |b|$, незначимы и исключаются из уравнения регрессии.

8. Проверить полученное уравнение регрессии второго порядка $Y = b_0 X_0 + b_1 X_1 + b_{11} X_1^2$ на адекватность по критерию Фишера.

8.1. В каждом опыте по уравнению регрессии второго порядка рассчитать $Y_j^p = b_0 + b_1 X_{1j} + b_{11} X_{1j}^2$, $j = 1, \ldots, N$.

8.2. В каждом опыте рассчитать квадрат разности экспериментальных и расчетных значений $(\overline{Y}_j - Y_j^p)^2$, $j = 1, \ldots, N$, а также остаточную сумму квадратов $\varphi = \sum_{j=1}^{N}\left(\overline{Y}_j - Y_j^p\right)^2$.

8.3. Рассчитать дисперсию адекватности:

$$S_{\text{ад}}^2 = \frac{n\varphi}{N-B}, \ f = N - B.$$

8.4. Рассчитать экспериментальное значение критерия Фишера:

$$F_{\text{э}} = \frac{\max\,(S_{\text{ад}}^2, S_{\text{воспр}}^2)}{\min\,(S_{\text{ад}}^2, S_{\text{воспр}}^2)}.$$

8.5. Если $F_{\text{э}} = \dfrac{S_{\text{ад}}^2}{S_{\text{воспр}}^2} < F_{N\text{-}B,\, N(n\text{-}1),\, p}$ при $S_{\text{ад}}^2 > S_{\text{воспр}}^2$,

где $F_{N\text{–}B,\, N(n\text{–}1),\, p}$ – критическое значение критерия Фишера (приложение 4), то $\sigma_{\text{ад}}^2 = \sigma_{\text{воспр}}^2$ и уравнение регрессии второго порядка адекватно.

8.6. Если $F_{\text{э}} = \dfrac{S_{\text{ад}}^2}{S_{\text{воспр}}^2} > F_{N\text{-}B,\, N(n\text{-}1),\, p}$ при $S_{\text{ад}}^2 > S_{\text{воспр}}^2$,

то $\sigma_{\text{ад}}^2 > \sigma_{\text{воспр}}^2$ и уравнение регрессии неадекватно. Следует перейти к построению уравнения регрессии третьего порядка (рассеяние средних величин относительно рассчитанных значений существенно больше рассеяния случайных значений параметра относительно выборочных средних). В данном учебно-методическом комплексе построение уравнений регрессии третьего порядка не рассматривается.

8.7. Если $F_э = \dfrac{S^2_{воспр}}{S^2_{ад}} < F_{N(n-1),\, N-B,\, p}$ при $S^2_{ад} < S^2_{воспр}$,

где $F_{N(n-1),\, N-B,\, p}$ – критическое значение критерия Фишера (приложение 4), то $\sigma^2_{ад} = \sigma^2_{воспр}$ и уравнение регрессии первого порядка адекватно.

8.8. Если $F_э = \dfrac{S^2_{воспр}}{S^2_{ад}} > F_{N(n-1),\, N-B,\, p}$ при $S^2_{ад} < S^2_{воспр}$,

то $\sigma^2_{ад} < \sigma^2_{воспр}$. Вероятность такого случая на практике ничтожно мала. Его возникновение может означать плохую организацию эксперимента, в котором нарушена идея случайности (нарушены основные статистические предпосылки). В этом случае эксперимент следует повторить, причем желательно увеличить и число опытов N, и число дублей n.

9. Рассчитать абсолютную погрешность рассчитанного параметра $Y(X_1)$ внутри исследованного факторного пространства $X_j \in [-1, +1]$:

$$\Delta Y(X_1) = t_{N(n-1),\, p}\frac{S_{воспр}}{\sqrt{Nn}}\sqrt{1+\frac{3(N-1)}{N+1}X_1^2+\frac{45(N-1)^3}{4(N+1)(N^2-4)}\left(X_1^2-\frac{N+1}{3(N-1)}\right)^2}\,.$$

10. Если уравнение регрессии второго порядка $Y = b_0X_0 + b_1X_1 + b_{11}X_1^2$ адекватно, рассчитать оптимальное значение фактора $X_{1опт}$, при котором исследуемый параметр достигает максимума (минимума):

$$Y_{max} = b_0 - \frac{b_1^2}{4b_{11}},\ \text{при}\ X_{1опт} = -\frac{b_1}{2b_{11}}\ \text{и}\ b_{11} < 0,$$

$$Y_{min} = b_0 - \frac{b_1^2}{4b_{11}},\ \text{при}\ X_{1опт} = -\frac{b_1}{2b_{11}}\ \text{и}\ b_{11} > 0\,.$$

Экспериментально проверить рассчитанное значение Y_{max} (Y_{min}) при $X_1 = X_{1опт}$. Если $|X_{1опт}| > 1$ (экспериментальная точка лежит вне исследованного факторного пространства), то возможно, что при экспериментальной проверке погрешность окажется больше рассчитанной.

11. В заключение можно получить (если в этом есть необходимость) уравнение регрессии, в котором фактор выражен в натуральных значениях. Для этого фактор X_1, выраженный в нормированных значениях, в обычном уравнении регрессии второго порядка необходимо заменить фактором x_1, выраженным в натуральных значениях, по формуле:

$$X_1 = \frac{x_1 - x_{1о}}{\Delta x_1}\,.$$

2.4.3. Базовая задача

Формулировка условия: изучить влияние витаминной добавки на суточный привес цыплят в «желторотом» периоде вскармливания в расширенном, по сравнению с базовой задачей (2.3.3), диапазоне изменения фактора.

Анализ изучаемого объекта: Изучаемым параметром, как и в базовой задаче (2.3.3), является привес цыплят Y (г/сут), который зависит, в том числе, от количества витаминной добавки в корме – фактор x_1 (г/сут). В интервале варьирования фактора от 0 до 12 г/сут привес цыплят описывается уравнением регрессии первого порядка $Y = 12.9 + 0.69x_1$. Поскольку линейный рост веса цыплят с увеличением добавки противоречит здравому смыслу, представляет интерес выявить истинный характер зависимости путем расширения диапазона варьирования фактора.

Планирование эксперимента.

Фактор. В базовой задаче (2.3.3) содержание витаминной добавки x_1 варьировалось от 0 до 12 г/сут. Для новой серии экспериментов принято решение варьировать содержание витаминной добавки в расширенном диапазоне от 0 до 20 г/сут. Для экономии затрат используем результаты опытов из базовой задачи (2.3.3) по РСП: x_1 = 0, 2, 4, 6, 8, 10, 12 г/сут (7 опытов). Логично дополнить этот РСП следующими равномерно симметричными значениями: x_1 = 14, 16, 18, 20 г/сут. В результате общее количество опытов возрастет до $N = 7 + 4 = 11$, но это тоже РСП.

Для реализации такого РСП дополнительно было организовано еще 4 контрольных вольера, в каждом из которых находилось по 100 цыплят «желторотого» возраста. Цыплята, так же как и в базовой задаче (2.3.3), в новых вольерах получали одинаковый корм по количеству и составу, за исключением количества витаминной добавки.

Параметр. Так же как и в базовой задаче (2.3.3) принимаем число дублей $n = 4$. Экспериментальные значения параметра Y (суточный привес одного цыпленка) в дополнительных 4-х вольерах определялись аналогичным, как в базовой задаче (2.3.3), способом.

Методика эксперимента. В начале исследования в каждом дополнительном вольере определялась масса 10 произвольно выбранных цыплят. Были получены следующие результаты в граммах: 497, 507, 495, 506. Таким образом, средняя масса одного цыпленка в начале эксперимента по 4-м дополнительным вольерам составила $m_0 = (497 + 507 + 495 + 506)/40 = 50.1$ г (достаточно хорошее совпадение со средним выборочным начального веса цыплят по 7-ми вольерам (см. задачу 2.3.3)). Затем в течение 10 дней в корм цыплят в каждом вольере добавлялась витаминная добавка в соответствии с планом эксперимента (см. столбец x_{1j} таблицы 11). Например, в корм цыплят в вольере № 11 ежедневно добавлялось $20 \cdot 100 = 2000$ г/сут витаминной добавки.

Предварительная обработка экспериментальных данных дополнительных опытов. Из-за отсутствия априорной информации моделирование исследуемого объекта начнем с построения уравнения регрессии первого порядка: $Y = b_0 + b_1 x_1$. Поэтому расчеты для его построения выполним по алгоритму 2.3.2.

Построение уравнения регрессии первого порядка.

Выборочное среднее в каждом дополнительном опыте (п. 5.1):

$$\overline{Y}_j = \frac{\sum_{i=1}^{4} Y_{ji}}{4}.$$

Например, $\overline{Y}_{10} = (19.4 + 23.0 + 22.8 + 22.1)/4 = 21.83$.

Выборочная дисперсия в каждом дополнительном опыте (п. 5.2):

$$S_j^2 = \sum_{i=1}^{4} (Y_{ji} - \overline{Y}_j)^2 / (4-1).$$

Например, $$S_{10}^2 = \frac{(19.4 - 21.83)^2 + (23.0 - 21.83)^2}{4-1} +$$

$$+ \frac{(22.8 - 21.83)^2 + (22.1 - 21.83)^2)}{(4-1)} = 2.76, \; S_{10} = 1.66.$$

Анализ дополнительных опытов на промахи (п. 5.3). В 4-х дополнительных опытах промахов нет. Например, для опыта № 10 наиболее выделяющимся значением является $Y_{10,1} = 19.4$. Экспериментальное значение критерия Смирнова–Граббса $\tau_{э10} = \max\left|Y_{10,i} - \overline{Y}_{10}\right| / S_{10} = \left|19.4 - 21.83\right| / 1.66 = 1.46$. Критическое значение критерия Смирнова–Граббса равно $\tau_{2,\,0.95} = 1.69$ (приложение 1, $f = 4 - 2 = 2$). Следовательно, в опыте № 10 промахов нет, так как $\tau_{э10} < \tau_{2,\,0.95}$.

Анализ результатов дополнительных опытов на закон распределения (п. 5.4). Все случайные значения 4-х дополнительных опытов подчиняются нормальному закону распределения. Например, для опыта № 10:

$$\left|\text{САО}_{10} / S_{10} - 0.7979\right| < 0.4/\sqrt{n} \to \left|1.21/1.66 - 0.7979\right| < 0.4/\sqrt{4} \to$$

$$\to \left|\, 0.7289 - 0.7979 \,\right| = 0.0690 < 0.2,$$

где $$\text{САО}_{10} = \frac{\sum_{i=1}^{n}\left|Y_{10,i} - \overline{Y}_{10}\right|}{n} = \frac{\left|19.4 - 21.83\right| + \left|23.0 - 21.83\right|}{4} +$$

$$+ \frac{\left|22.8 - 21.83\right| + \left|22.1 - 21.83\right|}{4} = 1.21.$$

Проверка дисперсий всех опытов на однородность по критерию Кохрена (п. 5.5) показала, что все дисперсии однородны, так как $G_э < G_{3,\,11\,0.95} = 0.35$ (приложение 5).

$$G_э = \frac{\max(S_j^2)}{\sum_{j=1}^{11} S_j^2} = \frac{2.84}{1.62 + 1.00 + 1.12 + 1.11 + 0.48 + 0.22} +$$

$$+ \frac{2.84}{2.84 + 0.78 + 1.23 + 2.76 + 0.46} = = \frac{2.84}{13.62} = 0.21.$$

Поскольку предварительная обработка экспериментальных данных в дополнительных опытах не выявила отклонений от основных предпосылок регрессионного анализа (нормальный закон распределения и однородность дисперсий), переходим к следующему этапу – построению уравнения регрессии первого порядка.

Таблица 11 – План эксперимента и результаты обработки данных по вскармливанию цыплят

План эксперимента								Уравнение регрессии 1-го порядка				Уравнение регрессии 2-го порядка			
N	x_{1j}	Y_{j1}	Y_{j2}	Y_{j3}	x_{1j}	$\overline{Y}_j$	S_j^2	X_{0j}	X_{1j}	Y_j^p	$\left(\overline{Y}_j - Y_j^p\right)^2$	X_{1j}^2	$X_{1j}^2 - 0.4$	Y_j^p	$\left(\overline{Y}_j - Y_j^p\right)^2$
1	0	10.3	11.2	13.0	12.7	11.80	1.62	1	-1.0	13.96	4.67	1.00	0.60	12.21	0.168
2	2	14.8	14.6	16.2	13.8	14.85	1.00	1	-0.8	14.91	0.00	0.64	0.24	14.21	0.412
3	4	16.8	14.5	16.5	16.5	16.08	1.12	1	-0.6	15.87	0.04	0.36	-0.04	15.97	0.011
4	6	18.8	17.4	16.3	18.0	17.63	1.11	1	-0.4	16.82	0.66	0.16	-0.24	17.51	0.015
5	8	18.4	17.4	19.0	18.7	18.38	0.48	1	-0.2	17.78	0.36	0.04	-0.36	18.81	0.185
6	10	18.9	19.8	19.6	18.9	19.30	0.22	1	0.0	18.73	0.32	0.00	-0.40	19.88	0.336
7	12	23.4	19.7	20.0	20.7	20.95	2.84	1	0.2	19.68	1.61	0.04	-0.36	20.72	0.054
8	14	22.0	23.4	21.3	22.5	22.30	0.78	1	0.4	20.64	2.76	0.16	-0.24	21.32	0.953
9	16	19.8	21.9	20.5	22.1	21.08	1.23	1	0.6	21.59	0.26	0.36	-0.04	21.70	0.382
10	18	19.4	23.0	22.8	22.1	21.83	2.76	1	0.8	22.55	0.52	0.64	0.24	21.84	0.000
11	20	21.2	21.4	22.7	22.0	21.83	0.46	1	1.0	23.50	2.79	1.00	0.60	21.75	0.006
N = 11 (опыты) n = 4 (дубли)						$\sum_{j=1}^{11} S_j^2 = 13.62$			$\sum_{j=1}^{11} X_{0j}^2 = 11$		φ=13.99	$\sum_{j=1}^{11} (X_j^2 - 0.4)^2 = 1.373$			φ = 2.522
$x_{1о}$ = 10 Δx_1 = 10						$S^2_{воспр}$ =1.238, $f_{воспр}$ = 33			$\sum_{j=1}^{11} X_j^2$ =4.40		$S^2_{ад}$ =6.22, $f_{ад}$=9	b_0 =19.88	b_{11} = – 2.9		$S^2_{ад}$ =1.261 $f_{ад}$ = 8
$t_{33,\ 0.95}$ = 2.035						b_0 = 18.73			b_1 = 4.77		$F_э$=5.02	$S^2(b_0)$ = = 0.0643	$S^2(b_{11})$ = = 0.226		$F_э$ =1.018
$G_э$ = 0.21, $G_{3,11,0.95}$ = 0.35						$S^2(b_0)$ = 0.0281			$S^2(b_1)$ = 0.0705		$F_{9,\ 33,\ 0.95}$= = 2.18				$F_{8,\ 33,\ 0.95}$= = 2.235
Y=19.88+4.77X_1 -2.9 X_1^2						$S(b_0)$ = 0.168			$S(b_1)$ = 0.265		Модель неадекватна	$S(b_0)$ = 0.253	$S(b_{11})$ = 0.475		Модель адекватна
Y=12.21+1.057x_1–0.029 x_1^2						Δb_0 = 0.34			Δb_1 = 0.54			Δb_0 =0.52	Δb_{11} =1.0		

Расчет нормированных значений фактора (п. 3.2). Для составленного РСП из 11 опытов рассчитаем для каждого опыта значения нормированного фактора:

$$X_j = 2 \cdot \frac{j-1}{N-1} - 1, \; j = 1, ..., 11.$$

Например, $X_{1,8} = \frac{2 \cdot (8-1)}{11-1} - 1 = \frac{14}{10} - 1 = 0.4$.

Для универсализации расчетов создадим фиктивный фактор X_0.

Расчет коэффициентов уравнения регрессии первого порядка b_0, b_1 (пп. 6.1, 6.2):

Создадим столбцы $X_{0j}\overline{Y}_j$, $X_{1j}\overline{Y}_j$ и рассчитаем их суммы:

$$\sum_{j=1}^{11} X_{0j}\overline{Y}_j = 1 \cdot 11.80 + 1 \cdot 14.65 + 1 \cdot 16.08 + 1 \cdot 17.63 + 1 \cdot 18.38 +$$

$$+ 1 \cdot 19.30 + 1 \cdot 20.95 + 1 \cdot 22.30 + 1 \cdot 21.08 + 1 \cdot 21.83 + 1 \cdot 21.83 = 206.0,$$

$$\sum_{j=1}^{11} X_{1j}\overline{Y}_j = (-1) \cdot 11.80 + (-0.8) \cdot 14.85 + (-0.6) \cdot 16.08 +$$

$$+(-0.4) \cdot 17.63 + (-0.2) \cdot 18.38 + 0 \cdot 19.30 + 0.2 \cdot 20.95 + 0.4 \cdot 22.30 +$$

$$+0.6 \cdot 21.08 + 0.8 \cdot 21.83 + 1 \cdot 21.83 = 21.00.$$

Рассчитаем коэффициенты регрессии b_0, b_1:

$$b_0 = \frac{\sum_{j=1}^{11} X_{0j}\overline{Y}_j}{11} = \frac{206.0}{11} = 18.73,$$

$$b_1 = \frac{3(N-1)}{N+1} \cdot \frac{\sum_{j=1}^{11} X_{1j}\overline{Y}_j}{N} = \frac{3 \cdot (11-1) \cdot 21.0}{(11+1) \cdot 11} = 4.77.$$

Проверка коэффициентов на значимость.

Дисперсия воспроизводимости (п. 7.1):

$$S^2_{\text{воспр}} = \frac{\sum_{j=1}^{11} S_j^2}{11} = \frac{13.62}{11} = 1.24.$$

Дисперсии значимости коэффициентов регрессии (п. 7.2):

$$S^2(b_0) = \frac{S^2_{\text{воспр}}}{Nn} = \frac{1.24}{4 \cdot 11} = 0.0282 \;\rightarrow\; S(b_0) = 0.168,$$

$$S^2(b_1) = \frac{3(N-1)}{N+1} \cdot \frac{S^2_{\text{воспр}}}{Nn} = \frac{3 \cdot (11-1) \cdot 1.24}{(11+1) \cdot 11 \cdot 4} = 0.0705 \;\rightarrow\; S(b_1) = 0.265.$$

Доверительные интервалы коэффициентов регрессии (п. 7.3):

$$\Delta b_0 = t_{N(n-1),\,0.95} \cdot S(b_0) = 2.035 \cdot 0.168 = 0.3419 \approx 0.34,$$

$$\Delta b_1 = t_{N(n-1),\,0.95} \cdot S(b_1) = 2.035 \cdot 0.265 = 0.5393 \approx 0.54,$$

где $t_{N(n-1),\,p} = t_{33,\,0.95} = 2.035$ критическое значение критерия Стьюдента (приложение 2).

Оба коэффициента регрессии значимы, так как $\Delta b_0 < |b_0|$ ($0.34 < 18.73$) и $\Delta b_1 < |b_1|$ ($0.56 < 4.77$) (п. 7.4).

Таким образом, в уравнении регрессии первого порядка $Y = 18.73 + 4.77X_1$ все коэффициенты регрессии значимы.

Анализ уравнения на адекватность по критерию Фишера.

Расчетные значения параметра Y_j^p в каждом опыте $j = 1, \ldots, N$ определяются по уравнению $Y_j^p = 18.73 + 4.77X_{1j}$ (п. 8.1).

Например, $Y_{10}^p = 18.73 + 4.77 \cdot 0.8 = 22.55$, $Y_{11}^p = 18.73 + 4.77 \cdot 1 = 23.50$.

Остаточная сумма квадратов φ (п. 8.2):

$$\varphi = \sum_{j=1}^{11} (\overline{Y}_j - Y_j^p)^2 = (11.80 - 13.96)^2 + (14.85 - 14.91)^2 +$$

$$+ (16.08 - 15.87)^2 + (17.63 - 16.82)^2 + (18.38 - 17.78)^2 +$$

$$+ (19.30 - 18.73)^2 + (20.95 - 19.68)^2 + (22.30 - 20.64)^2 +$$

$$+ (21.08 - 21.59)^2 + (21.83 - 22.55)^2 + (21.83 - 23.50)^2 =$$

$$= 4.67 + 0.00 + 0.04 + 0.66 + 0.36 + 0.32 +$$

$$+ 1.61 + 2.76 + 0.26 + 0.52 + 2.79 = 13.99.$$

Дисперсия адекватности (п. 8.3):

$$S^2_{\text{ад}} = \frac{n\varphi}{N-B} = \frac{4\cdot 13.99}{11-2} = 6.22\,,\;\; f = N - B = 11 - 2 = 9\,.$$

Экспериментальное значение критерия Фишера (п. 8.4):

$$F_{\text{э}} = \frac{S^2_{\text{ад}}}{S^2_{\text{воспр}}} = \frac{6.22}{1.24} = 5.02\,, \text{ так как } S^2_{\text{ад}} > S^2_{\text{воспр}}\,,$$

где $F_{(N-B),\,N(n-1),\,p} = F_{9,\,33,\,0.95} = 2.18$ – критическое значение критерия Фишера (приложение 4). Выполнение неравенства $F_{\text{э}} > F_{9,\,33,\,0.95}$ ($5.02 > 2.18$) означает, что полученное уравнение регрессии первого порядка $Y = 18.73 + 4.77X$ неадекватно (п. 8.6).

Повысим степень уравнения регрессии и перейдем к построению уравнения регрессии второго порядка. Дальнейшие операции будут проводится **по алгоритму 2.4.2.** При этом необходимо учесть, что для построения уравнения регрессии второго порядка рекомендуемое условие $N \geq 6$ (п. 3)) выполнено, так как после проведения дополнительных опытов общее количество $N = 11$, что вполне достаточно, и в проведении дополнительных опытов нет необходимости.

Построение уравнения регрессии второго порядка: $Y = b_0 + b_1X_1 + b_{11}X_1^2$.

Рассчитаем ортогонализирующий коэффициент λ_1 (п. 6.1):

$$\lambda_1 = \overline{X_1^2} = \frac{\sum_{j=1}^{11} X_{1j}^2}{N} = \frac{N+1}{3(N-1)} = \frac{11+1}{3(11-1)} = \frac{12}{30} = 0.40\,.$$

В таблице 11 построим столбец фактора $X_{1j}^2 - \lambda_1 = X_{1j}^2 - 0.4$. Легко убедиться в том, что для РСП эксперимента, состоящего из 11 опытов, факторы X_0, X_1, $X_1^2 - 0.4$ ортогональны (таблица 11):

$$\sum_{j=1}^{11} X_{0j}X_{1j} = \sum_{j=1}^{11} X_{1j} = 0\,,\;\; \sum_{j=1}^{11} X_{0j}(X_{1j}^2 - 0.4) = \sum_{j=1}^{11}(X_{1j}^2 - 0.4) = 0\,.$$

Рассчитаем сумму произведений $(X_{1j}^2 - 0.4)\overline{Y}_j$ (п. 6.1), поскольку суммы произведений $X_{0j}\overline{Y}_j$, $X_{1j}\overline{Y}_j$ уже рассчитаны:

$$\sum_{j=1}^{11}(X_{1j}^2 - 0.4)\overline{Y}_j = 0.60\cdot 11.80 + 0.24\cdot 14.85 + (-0.04)\cdot 16.08 +$$
$$+(-0.24)\cdot 17.63 + (-0.36)\cdot 18.38 + (-0.40)\cdot 19.30 + (-0.36)\cdot 20.95 +$$
$$+(-0.24)\cdot 22.30 + (-0.04)\cdot 21.08 + 0.24\cdot 21.83 + 0.60\cdot 21.83 = -3.97\,.$$

Расчет коэффициентов регрессии b_0, b_1, b_{11} (п. 6.2):

$$b_0 = \frac{\sum_{j=1}^{11} X_{0j}\overline{Y}_j}{N} - \frac{15(N-1)^2}{4(N^2-4)}\cdot\frac{\sum_{j=1}^{11}\left(X_{1j}^2 - \lambda_1\right)\overline{Y}_j}{N} =$$

$$= \frac{206.0}{11} - \frac{15\cdot(11-1)^2\cdot(-3.97)}{4\cdot(11^2-4)\cdot 11} = 18.73 + 1.15 = 19.88\,;$$

$$b_1 = \frac{3(N-1)}{N+1}\cdot\frac{\sum_{j=1}^{N} X_{1j}\overline{Y}_j}{N} = \frac{3\cdot(11-1)\cdot 21.00}{(11+1)\cdot 11} = 4.77\,;$$

$$b_{11} = \frac{45(N-1)^3}{4(N+1)(N^2-4)}\frac{\sum_{j=1}^{N}\left(X_j^2 - \lambda\right)\overline{Y}_j}{N} = \frac{45\cdot(11-1)^3\cdot(-3.97)}{4\cdot(11+1)\cdot(11^2-4)\cdot 11} = -2.89\,.$$

Проверка коэффициентов регрессии второго порядка на значимость.

Дисперсия воспроизводимости уже рассчитана при построении уравнения регрессии первого порядка $S^2_{\text{воспр}} = 1.24$ (п. 7.1).

Дисперсии значимости коэффициентов b_0, b_1, b_{11} уравнения регрессии второго порядка (п. 7.2):

$$S^2(b_0) = \left[1 + \frac{5(N^2-1)}{4(N^2-4)}\right]\cdot\frac{S^2_{\text{воспр}}}{Nn} = \left[1 + \frac{5\cdot(11^2-1)}{4\cdot(11^2-4)}\right]\cdot\frac{1.24}{11\cdot 4} = 0.0642\,,$$

$$S^2(b_{11}) = \frac{45(N-1)^3}{4(N+1)(N^2-4)} \cdot \frac{S^2_{\text{воспр}}}{Nn} = \frac{45 \cdot (11-1)^2 \cdot 1.24}{4 \cdot (11+1) \cdot (11^2-4) \cdot 11 \cdot 4} = 0.226\,.$$

Соответствующие среднеквадратичные отклонения равны:

$$S(b_0) = \sqrt{0.0642} = 0.253, \qquad S(b_{11}) = \sqrt{0.226} = 0.475\,.$$

Расчет доверительных интервалов коэффициентов регрессии b_0, b_1, b_{11} (п. 7.3) (использованы правила округления, см. раздел 1.3.2):

$$\Delta b_0 = t_{N(n-1),\,p} S(b_0) = 2.035 \cdot 0.253 = 0.5149 \approx 0.52\,;$$

$$\Delta b_1 = t_{N(n-1),\,p} S(b_1) = 2.035 \cdot 0.265 = 0.5393 \approx 0.54\,;$$

$$\Delta b_{11} = t_{N(n-1),\,p} \cdot S(b_{11}) = 2.035 \cdot 0.475 = 0.9666 \approx 1.0\,.$$

где $t_{N(n-1),\,\text{p}} = t_{33,\ 0.95} = 2.035$ – критическое значение критерия Стьюдента (приложение 2). Все коэффициенты уравнения регрессии второго порядка $Y = 19.88 + 4.77X_1 - 2.9X_1^2$ значимы, так как $\Delta b_0 < |b_0|$ (0.52 < 19.88), $\Delta b_1 < |b_1|$ (0.54 < 4.77), $\Delta b_{11} < |b_{11}|$ (1.0 < 2.9) (п. 7.4).

Анализ уравнения регрессии второго порядка на адекватность.

Расчетные значения параметра Y_j^p (п. 8.1) определяются по уравнению регрессии $Y = 19.88 + 4.77X_1 - 2.9X_1^2$.

Например, $Y_{10}^p = 19.88 + 4.77 \cdot 0.8 - 2.9 \cdot 0.8^2 = 21.84$.

Эти и другие расчетные значения Y_j^p внесены в таблицу 11.

Остаточная сумма квадратов φ (п. 8.2):

$$\varphi = \sum_{j=1}^{11} (\overline{Y}_j - Y_j^p)^2 = (11.80 - 12.21)^2 + (14.85 - 14.21)^2 +$$

$$+ (16.08 - 15.97)^2 + (17.63 - 17.51)^2 + (18.38 - 18.81)^2 +$$

$$+ (19.30 - 19.88)^2 + (20.95 - 20.72)^2 + (21.30 - 21.32)^2 +$$

$$+ (21.08 - 21.70)^2 + (21.83 - 21.84)^2 + (21.83 - 21.75)^2 =$$

$$= 0.168 + 0.412 + 0.011 + 0.015 + 0.185 + 0.336 + 0.054 + 0.953 +$$

$$+ 0.382 + 0.000 + 0.006 = 2.522\,.$$

Дисперсия адекватности (п. 8.3):

$$S^2_{\text{ад}} = \frac{n\varphi}{N-B} = \frac{4 \cdot 2.522}{11-3} = 1.261\,,\ \ f = N - B = 11 - 3 = 8\,.$$

Экспериментальное значение критерия Фишера (п. 8.4):

$$F_э = \frac{S^2_{\text{ад}}}{S^2_{\text{воспр}}} = \frac{1.261}{1.24} = 1.017\text{, так как } S^2_{\text{ад}} > S^2_{\text{воспр}}\,.$$

Критическое значение критерия Фишера $F_{N\text{-}B,\,N(n\text{-}1),\,0.95} = F_{8,\,33,\,0.95} = 2.235$ (приложение 4).

Так как $F_э < F_{8,\,33,\,0.95}$ (1.017 < 2.235), то уравнение регрессии второго порядка $Y = 19.88 + 4.77X_1 - 2.9X_1^2$ **адекватно** (п. 8.5).

Нахождение максимума уравнения регрессии второго порядка.

Необходимое и достаточное условия максимума $Y' = 0$ и $Y'' < 0$: $Y'' = -2 \cdot 2.9 = -5.8 < 0$ (функция имеет максимум) (п. 10);

$$Y' = 4.77 - 2{\cdot}2.9X_1 = 0 \rightarrow X_{1\text{опт}} = 4.77/(2{\cdot}2.9) = 0.822,$$

$$Y_{\max} = 19.88 + 4.77{\cdot}0.822 - 2.9 \cdot 0.822^2 = 21.84 \text{ г/сут.}$$

Натуральное значение фактора $x_{1\text{опт}}$ равно:

$$x_{1\text{опт}} = x_{1\text{о}} + X_{1\text{опт}}{\cdot}\Delta x_1 = 10 + 0.822\ 10 = 18.2 \text{ г/сут (п. 11).}$$

Определение предельной абсолютной погрешности параметра $Y_{\max}$ выполняется с помощью уравнения регрессии второго порядка ортогонализированного вида (п. 9):

$$\Delta Y(X_{1\text{опт}}) = t_{N(n-1),\,p} \frac{S_{\text{воспр}}}{\sqrt{Nn}} \sqrt{1 + \frac{3(N-1)}{N+1} X_1^2 + \frac{45(N-1)^3}{4(N+1)(N^2-4)} \left(X_1^2 - \frac{N+1}{3(N-1)} \right)^2}\,.$$

Численное значение предельной абсолютной погрешности $Y_{\max}$ равно:

$$\Delta Y(X_{1\text{опт}}) = 2.035 \cdot \frac{\sqrt{1.24}}{\sqrt{11 \cdot 4}} \times$$

$$\times \sqrt{1 + \frac{3 \cdot (11-1)}{12} \cdot 0.822 + \frac{45 \cdot (11-1)^3}{4 \cdot 12 \cdot (11^2-4)} \left(0.822^2 - 0.4 \right)^2} = 0.65 \text{ г/сут,}$$

$$\frac{\Delta Y(X_{1\text{опт}})}{Y_{\max}} = \frac{0.65}{21.84} = 0.030 \quad (3.0\,\%)\,.$$

Полученное уравнение регрессии второго порядка в натуральных значениях фактора с учетом того, что $X_1 = (x_1 - x_{1о})/\Delta x_1$, а $x_{1о} = 10$, $\Delta x_1 = 10$, будет иметь следующий вид (п. 11):

$$Y = 19.88 + 4.77 X_1 - 2.9 X_1^2 = 19.88 + 4.77(x_1 - 10)/10 -$$

$$-2.9(x_1 - 10)^2/10^2 = 12.21 + 1.057 x_1 - 0.029 x_1^2 .$$

Резюме. 1. Привес цыплят без использования витаминной добавки составил 12.2 г/сут. За 10 дней без витаминной добавки средний вес цыпленка составил бы: 50 + 12.2·10 =172 г.

2. Максимальный привес цыплят при использовании витаминной добавки в количестве 18.2 г/сут составляет 21.8 г/сут. За 10 дней кормления с витаминной добавкой средний вес цыпленка составит: 50 + 21.8·10 = 268 г. Таким образом, использование витаминной добавки увеличивает средний привес цыплят за 10 дней на 268 – 172 = 96 г (≈ 60 %).

3. Полученное уравнение регрессии справедливо для возрастного периода 2–20 дней. Допускается экстраполяция, например, еще на 3–4 дня. Однако, если нужно получить более полную и достоверную количественную информацию о привесе цыплят за относительно большой срок вскармливания, например, за несколько месяцев, да еще с учетом влияния не только витаминной добавки, но и других факторов (возраст, ультрафиолетовое облучение и т. п.), то исследование процесса вскармливания должно базироваться на многофакторном моделировании.

2.4.4. Материалы для самостоятельной работы, лабораторных и практических занятий

Тесты

1. Однофакторное уравнение регрессии второго порядка строится в ортогональной форме $Y = b_0^* X_0 + b_1^* X_1 + b_{11}^* (X_1^2 - \lambda_1)$, где:

а) $\lambda_1 = \overline{X_1^2}$; б) $\lambda_1 = \dfrac{\sum_{j=1}^{N} X_{1j}^2}{N}$; в) $\lambda = \sum_{j=1}^{N} X_{1j}^2$; г) $\lambda_1 = \dfrac{\sum_{j=1}^{N} X_j}{N}$.

2. Коэффициенты регрессии второго порядка однофакторного эксперимента рассчитываются по формулам:

а) $b_0^* = \dfrac{\sum_{j=1}^{N} X_{0j} \overline{Y}_j}{\sum_{j=1}^{N} X_{0j}}$, $b_1^* = \dfrac{\sum_{j=1}^{N} X_{1j} \overline{Y}_j}{\sum_{j=1}^{N} X_{1j}}$, $b_{11}^* = \dfrac{\sum_{j=1}^{N} \left(X_{1j}^2 - \overline{X^2}\right) \overline{Y}_j}{\sum_{j=1}^{N} \left(X_{1j}^2 - \overline{X^2}\right)}$;

б) $b_0^* = \dfrac{\sum_{j=1}^{N} X_{0j} \overline{Y}_j}{\sum_{j=1}^{N} X_j^2}$, $b_1^* = \dfrac{\sum_{j=1}^{N} \overline{Y}_j}{\sum_{j=1}^{N} X_{1j}^2}$, $b_{11}^* = \dfrac{\sum_{j=1}^{N} (X_{1j}^2 - \overline{X_1^2}) \overline{Y}_j}{\sum_{j=1}^{N} X_{1j}^2}$;

в) $b_0^* = \dfrac{\sum_{j=1}^{N} X_{0j} \overline{Y}_j}{N \sum_{j=1}^{N} X_{0j}^2}$, $b_1^* = \dfrac{\sum_{j=1}^{N} X_{1j} \overline{Y}_j}{N \sum_{j=1}^{N} X_{1j}^2}$, $b_{11}^* = \dfrac{\sum_{j=1}^{N} \left(X_{1j}^2 - \overline{X_1^2}\right) \overline{Y}_j}{N \sum_{j=1}^{N} \left(X_{1j}^2 - \overline{X_1^2}\right)^2}$;

г) $b_0^* = \dfrac{\sum_{j=1}^{N} X_{0j} \overline{Y}_j}{\sum_{j=1}^{N} X_{0j}^2}$, $b_1^* = \dfrac{\sum_{j=1}^{N} X_{1j}^2 \overline{Y}_j}{\sum_{j=1}^{N} X_{1j}^2}$, $b_{11}^* = \dfrac{\sum_{j=1}^{N} \left(X_{1j}^2 - \overline{X_1^2}\right) \overline{Y}_j}{\sum_{j=1}^{N} \left(X_{1j}^2 - \overline{X_1^2}\right)^2}$.

3. Соотношение коэффициентов уравнения регрессии второго порядка в общей форме и в ортогональной форме:

а) $b_0 = b_0^* - b_{11}^*$; $b_1 = b_1^* - \overline{X_1^2}$; $b_{11} = b_{11}^*$;

б) $b_о = b_о^* - \overline{X^2} \cdot b_2^*$; $b_1 = b_1^*$; $b_{11} = b_{11}^*$;

в) $b_0 = b_0^* - \dfrac{\sum_{j=1}^{N} X_{1j}^2}{N} \cdot b_{11}^*$; $b_1 = b_1^*$; $b_{11} = b_{11}^*$;

г) $b_0 = b_0^*$; $b_1 = b_1^* - \lambda_1 b_{11}^*$; $b_{11} = b_{11}^*$.

4. В однофакторном эксперименте для РСП имеем:

$$\text{а) } \sum_{j=1}^{N}\left(X_{1j}^2-\overline{X_1^2}\right)^2=\frac{4(N+1)(N^2-4)}{45(N-1)^3};$$

$$\text{б) } \sum_{j=1}^{N}\left(X_{1j}^2-\lambda_1\right)^2=\frac{4N(N^2-4)}{45(N-1)^3};$$

$$\text{в) } \sum_{j=1}^{N}\left(X_{1j}^2-\overline{X_1^2}\right)^2=\frac{4N(N+1)(N^2-4)}{45(N-1)^3}.$$

5. Дисперсии значимости коэффициентов однофакторного уравнения регрессии второго порядка в обычной форме:

$$\text{а) } S^2(b_1)=\frac{S^2_{\text{воспр}}}{n\sum\limits_{j=1}^{N}X_{1j}^2}, \quad S^2(b_{11})=\frac{S^2_{\text{воспр}}}{n\sum\limits_{j=1}^{N}\left(X_{1j}^2-\overline{X_1^2}\right)^2},$$

$$S^2(b_0)=\frac{S^2_{\text{воспр}}}{n\sum\limits_{j=1}^{N}X_{0j}^2}+\left(\overline{X_1^2}\right)^2S^2(b_{11});$$

$$\text{б) } S^2(b_1)=\frac{S^2_{\text{воспр}}}{N\sum\limits_{j=1}^{N}X_{1j}^2}, \quad S^2(b_{11})=\frac{S^2_{\text{воспр}}}{N\sum\limits_{j=1}^{N}\left(X_{1j}^2-\lambda_1\right)^2},$$

$$S^2(b_0)=\frac{S^2_{\text{воспр}}}{N\sum\limits_{j=1}^{N}X_{0j}^2}+\lambda_1^2S^2(b_{11});$$

$$\text{в) } S^2(b_1)=\frac{S^2_{\text{воспр}}}{n\sum\limits_{j=1}^{N}X_{1j}^2}, \quad S^2(b_{11})=\frac{S^2_{\text{воспр}}}{n\sum\limits_{j=1}^{N}\left(X_{1j}^2-\lambda_1\right)^2},$$

$$S^2(b_0)=\frac{S^2_{\text{воспр}}}{n\sum\limits_{j=1}^{N}X_{0j}^2}+\lambda_1^2S^2(b_{11}).$$

Вопросы для самопроверки

1. Напишите уравнение регрессии второго порядка в общем и ортогонализированном видах.
2. Напишите формулу для расчета ортогонализирующего коэффициента квадратичного члена уравнения регрессии второго порядка.
3. Напишите формулы для расчета коэффициентов уравнения регрессии второго порядка в общем и ортогонализированном видах (*N* опытов, *n* дублей).
5. Напишите формулу для расчета дисперсии воспроизводимости (*N* опытов, *n* дублей). Чему равно число степеней свободы для этого параметра?
6. Напишите формулы для расчета дисперсий значимости коэффициентов уравнения регрессии второго порядка в общем и ортогонализированном видах (*N* опытов, *n* дублей).
7. Напишите формулы для расчета доверительных интервалов коэффициентов уравнения регрессии второго порядка.
8. Сформулируйте критерий значимости коэффициентов регрессии в общем виде.
9. Напишите формулу для остаточной суммы квадратов.
10. Напишите формулу для расчета дисперсии адекватности (*N* опытов, *n* дублей). Чему равно число степеней свободы для этого параметра?
11. Сформулируйте критерий адекватности уравнения регрессии любого порядка.
12. Сформулируйте необходимое и достаточное условие экстремума для однофакторного уравнения регрессии второго порядка. Напишите выражение для экстремума однофакторного уравнения регрессии второго порядка? Всегда ли оптимальное значение фактора находится внутри исследованного факторного пространства?
13. Напишите формулу для предельной абсолютной погрешности параметра $Y(X)$, рассчитанного по однофакторному уравнению регрессии второго порядка.

Задачи для самостоятельного решения

Уровень 1

1. Пусть температура сушильного агента в аппарате для сушки зерна является фактором x_1. Представляет интерес следующий диапазон изменения фактора x_1: $x_{1\min}$ = 45 °C, $x_{1\max}$ = 93 °C. Построить РСП для 9 опытов.

Ответ:

РСП для 9 опытов

N	$x_{1j} = T_j$, °C	X_{1j}
1	45	-1.00
2	51	-0.75
3	57	-0.50
4	63	-0.25
5	69	0.00
6	75	0.25
7	81	0.50
8	87	0.75
9	93	1.00

Уровень 2

2. Рассчитать величины $\sum_{j=1}^{21} X_{0j}^2$, $\sum_{j=1}^{21} X_{1j}^2$, $\sum_{j=1}^{21} (X_{1j}^2 - \lambda_1)^2$ и построить РСП для 21 опыта.

Ответ: РСП: $X_{1j} = -1, -0.9, -0.8, -0.7, -0.6, -0.5, -0.4, -0.3, -0.2, -0.1, 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1$; $\lambda_1 = 0.3667$, $\sum_{j=1}^{16} X_{0j}^2 = 21$, $\sum_{j=1}^{21} X_{1j}^2 = 7.70$, $\sum_{j=1}^{21} (X_{1j}^2 - 0.3667)^2 = 2.243$.

Уровень 3

1. Измерения ТЭДС термопары алюмель-хромель в температурном диапазоне $x_1 \in (0, 620$ ºC) дали следующие результаты:

N	x_{1j}, ºC	Y_{j1}, мВ	Y_{j2}, мВ	Y_{j3}, мВ
1	20	0.1	0.0	0.2
2	140	19.1	17.3	16.2
3	260	36.2	36.0	36.4
4	380	57.7	56.1	54.3
5	500	76.0	75.6	76.7
6	620	96.4	97.9	96.7

Исследовать температурную зависимость термопары алюмель-хромель уравнением второго порядка. Уравнение регрессии построить в нормированных значениях фактора X_1. Проверить коэффициенты уравнения на значимость, а само уравнение – на адекватность.

Ответ: Уравнение первого порядка $Y = 47.16 + 48.6X_1$, где $X_1 = (x_1 - 320)/300$, неадекватно, так как $F_э = 4.80$, а $F_{4,\,12,\,0.95} = 3.26$. Оба коэффициента уравнения значимы, так как $\Delta b_0 = 0.51$, $\Delta b_1 = 0.8$. Уравнение регрессии второго порядка $Y = 46.0 + 48.6X + 2.6X_1^2$, где $X_1 = (x_1 - 320)/300$, адекватно, так как $F_э = 6.02$, а $F_{12,\,3,\,0.95} = 8.74$. Все три коэффициента уравнения значимы, так как $\Delta b_0 = 0.8$, $\Delta b_1 = 0.8$, $\Delta b_{11} = 1.3$.

Модуль 3
МНОГОФАКТОРНЫЙ ЭКСПЕРИМЕНТ

В результате изучения модуля студент должен:

– **уметь** составлять планы эксперимента для построения многофакторных уравнений регрессии первого и второго порядка при решении прикладных задач;

– **уметь** осуществлять статистическую оценку полученных моделей (значимость коэффициентов регрессии, адекватность уравнения регрессии);

– **уметь** решать оптимизационные задачи (крутое восхождение, поиск оптимума по моделям второго порядка) и оценивать точность рассчитываемого параметра;

– **обладать** навыками альтернативного мышления (предлагать альтернативные планы и модели исследуемого объекта).

3.1. Перечень и содержание лекций

Лекция 1. Методология построения многофакторных уравнений регрессии m-го порядка. Виды многофакторных уравнений регрессии: первого порядка, неполного второго порядка, второго порядка. Выбор существенных факторов. Требования к факторам. Переход от натуральных значений факторов к нормированным. Планирование многофакторных экспериментов. Свойства планов: симметричность, ортогональность, нормирование. Стратегия поиска оптимальных решений: центральный полный (дробный) факторный эксперимент $\rightarrow$ крутое восхождение $\rightarrow$ ортогональное центральное композиционное планирование $\rightarrow$ оптимизация. Уравнение регрессии первого порядка – начальный этап исследования изучаемого объекта. Планы центрального полного факторного эксперимента (ЦПФЭ). Планы центрального дробного факторного эксперимента (ЦДФЭ). Генерирующие соотношения, определяющий контраст. Смешанные эффекты. Расчет коэффициентов уравнения регрессии первого порядка для планов ЦПФЭ, ЦДФЭ методом наименьших квадратов. Дисперсия воспроизводимости. Дисперсии значимости коэффициентов уравнения регрессии первого порядка. Проверка коэффициентов уравнения регрессии первого порядка для планов ЦПФЭ, ЦДФЭ на значимость по критерию Стьюдента. Остаточная

сумма квадратов многофакторных моделей. Дисперсия адекватности. Проверка уравнения регрессии первого порядка многофакторного эксперимента, построенных по планам ЦПФЭ, ЦДФЭ, на адекватность по критерию Фишера. Оценка точности рассчитываемого параметра. Алгоритм решения прикладных задач ЦПФЭ, ЦДФЭ. Решение базовой задачи.

Лекция 2. Анализ уравнения регрессии первого порядка и принятие решения. Метод крутого восхождения – эффективный способ оптимизации исследуемого объекта. Построение плана эксперимента для крутого восхождения. Принятие решения после проведения экспериментов по крутому восхождению. Стационарная область. Алгоритм решения прикладных задач по крутому восхождению. Решение базовой задачи.

Лекция 3. Центральное ортогональное композиционное планирование (ОЦКП) эксперимента для построения многофакторного уравнения регрессии второго порядка (k факторов). Расчет коэффициентов уравнения регрессии второго порядка для ОЦКП методом наименьших квадратов. Расчет звездного плеча и ортогонализирующего коэффициента в общем случае (k факторов). Расчет натуральных и нормированных значений факторов в звездных точках плана в общем случае (k факторов). Расчет суммы квадратов для линейных и квадратичных факторов, а также для эффектов взаимодействия. Расчет коэффициентов уравнения регрессии второго порядка в многофакторном эксперименте в общем случае (k – факторов). Дисперсия воспроизводимости. Дисперсии значимости коэффициентов уравнения регрессии второго порядка для ОЦКП. Доверительные интервалы коэффициентов уравнения регрессии второго порядка. Проверка коэффициентов уравнения регрессии второго порядка на значимость по критерию Стьюдента.

Лекция 4. Остаточная сумма квадратов уравнения регрессии второго порядка. Дисперсия адекватности. Проверка уравнения регрессии второго порядка многофакторного эксперимента, построенного по ОЦКП, на адекватность по критерию Фишера. Принятие решения по уравнению регрессии второго порядка. Оптимизация исследуемого объекта. Оценка точности рассчитываемого параметра. Алгоритм решения прикладных задач и оптимизация с помощью многофакторного уравнения регрессии второго порядка. Решение базовой задачи.

3.2. Термины и условные обозначения

Полный факторный эксперимент (ПФЭ) – эксперимент, реализующий все возможные неповторяющиеся комбинации уровней независимых факторов, которые варьируются на двух уровнях.

Центральный полный факторный эксперимент (ЦПФЭ) – эксперимент, реализованный по композиционному плану, состоящему из плана ПФЭ и одного опыта в центре плана.

Дробный факторный эксперимент (ДФЭ) – эксперимент, реализующий некоторое подмножество всех возможных неповторяющихся комбинаций уровней независимых факторов, которые варьируются на двух уровнях.

Центральный дробный факторный эксперимент (ЦДФЭ) – эксперимент, реализованный по композиционному плану, состоящему из плана ДФЭ и одного опыта в центре плана.

Ортогональный центральный композиционный план (ОЦКП) – композиционный план эксперимента, состоящий из плана ЦПФЭ или ЦДФЭ и опытов в звездных точках (два опыта на каждый фактор).

МНК – метод наименьших квадратов.

Априори – до проведения эксперимента.

Апостериори – после проведения эксперимента.

N_k – число опытов: $N_k = N_{k0} = 2^k$ в ПФЭ; $N_k = N_{k0} + 1$ в ЦПФЭ; $N_k = N_{kg} = 2^{k-g}$ в ДФЭ; $N_k = N_{kg} + 1$ в ЦДФЭ; $N_k = N_{k0} + 2k + 1$ в ОЦКП на базе ЦПФЭ; $N_k = N_{kg} + 2k + 1$ в ОЦКП на базе ЦДФЭ.

j – текущий номер опыта: $j = 1, \ldots, N_k$.

n – количество дублей в каждом опыте.

i – текущий номер дубля $i = 1, \ldots, n$.

k – число факторов.

r, s, v, w – текущие номера факторов, $r, s, v, w = 1, \ldots, k$.

g – число генерирующих соотношений в ДФЭ.

m – степень уравнения регрессии полиноминального типа.

q – текущий индекс степени уравнения регрессии полиноминального типа $q = 0, \ldots, m$.

x_r – натуральное значение r-го фактора.

$x_{r\text{o}}$ – основной уровень фактора x_r.

$\Delta x_{r\text{o}}$ – интервал варьирования фактора x_r.

$x_{r\,\min}$ – минимальное значение фактора x_r.

$x_{r\,\max}$ – максимальное значение фактора x_r.

x_{rj} – натуральное значение r-го фактора в j-ом опыте.

X_r – нормированное значение r-го фактора.

X_{rj} – нормированное значение r-го фактора в j-ом опыте.

X_0 – фиктивный фактор, все значения которого $X_{0j} = 1$.

Y – параметр, имеющий определенную размерность.

Y_{ji} – значение параметра Y в j-ом опыте и в i-ом дубле.

$\overline{Y}_j$ – выборочное среднее параметра Y в j-ом опыте.

f – число степеней свободы.

$\tau_{\text{э}j}$ – экспериментальное значение критерия Смирнова–Граббса в j-ом опыте.

$\tau_{f,\,p}$ – критическое значение критерия Смирнова–Граббса с числом степеней свободы f при доверительной вероятности p (приложение 1).

S_j^2 – выборочная дисперсия параметра Y в j-ом опыте.

S_j – выборочное стандартное отклонение параметра Y в j-ом опыте.

$G_\text{э}$ – экспериментальное значение критерия Кохрена.

$G_{f_1,\,f_2,\,p}$ – критическое значение критерия Кохрена для сравнения N дисперсий, в числителе дисперсия с числом степеней свободы f_1, в знаменателе сумма дисперсий с числом степеней свободы f_2 при доверительной вероятности p (приложение 5).

β_0, β_r, β_{rs}, β_{rr} – истинные коэффициенты уравнения регрессии при свободном, линейных, бинарных и квадратичных членах.

b_0, b_r, b_{rs}, b_{rr} – оценочные коэффициенты обычного уравнения регрессии при свободном, линейных, бинарных и квадратичных членах.

b_0^*, b_r^*, b_{rs}^*, b_{rr}^* – оценочные коэффициенты ортогонализированного уравнения регрессии при свободном, линейных, бинарных и квадратичных членах.

$S^2_{\text{воспр}}$ – дисперсия воспроизводимости.

$S^2(b_0)$, $S^2(b_r)$, $S^2(b_{rs})$, $S^2(b_{rr})$ – дисперсии значимости коэффициентов уравнения регрессии b_0, b_r, b_{rs}, b_{rr}.

$t_{f,\,p}$ – критическое значение критерия Стьюдента с числом степеней свободы f и доверительной вероятности p (приложение 2).

Δb_0, Δb_r, Δb_{rs}, Δb_{rr} – доверительные интервалы коэффициентов уравнения регрессии b_0, b_r, b_{rs}, b_{rr}.

Y_j^p – значение параметра Y в j-ом опыте, рассчитанное по уравнению регрессии.

φ – остаточная сумма квадратов.

B_k – число коэффициентов в уравнении регрессии.

B – число значимых коэффициентов в уравнении регрессии.

$S^2_{\text{ад}}$ – выборочная дисперсия адекватности.

$F_{\text{э}}$ – экспериментальное значение критерия Фишера.

$F_{f_1, f_2, p}$ – критическое значение критерия Фишера, в числителе бóльшая дисперсия с числом степеней свободы f_1, а в знаменателе меньшая дисперсия с числом степеней свободы f_2 при доверительной вероятности p (приложение 4).

$C_k^2 = \dfrac{k(k-1)}{1 \cdot 2}$ – число сочетаний из k по 2.

$C_k^3 = \dfrac{k(k-1)(k-2)}{1 \cdot 2 \cdot 3}$ – число сочетаний из k по 3.

3.3. Центральный полный факторный эксперимент (тема 7)

Большинство явлений в природе происходит под влиянием не одного, а многих факторов. Изучение и построение многофакторных зависимостей требует достаточно развитой математической теории эксперимента. На ранних этапах научных исследований изучение многофакторных зависимостей упрощали, сводя их к однофакторному эксперименту, поочередно варьируя одним фактором, а все остальные фиксировали на постоянных уровнях. Однако такой метод малоэффективен, связан с большим объемом экспериментов и не позволяет выявить взаимодействия между различными факторами. В начале XX века был разработан новый методологический подход к проведению многофакторных экспериментов, отличительной особенностью которого является одновременное варьирование всеми факторами по определенному плану. Использование оптимальных многофакторных планов позволяет при минимальном количестве опытов (минимальные материальные затраты) получить максимально возможную информацию с требуемой точностью.

Математическая теория многофакторного эксперимента включает в себя как формализуемые, так и неформализуемые этапы. Первый этап математической теории эксперимента заключается в выборе значимых k факторов x_r ($r = 1, \ldots, k$), их основных уровней и интервалов варьирования $[x_{r\,\min}, x_{r\,\max}]$. Этот этап неформализуем. Он базируется на априорной информации об изучаемом объекте и во многом определяется интуицией и опытом экспериментатора.

Построение уравнений регрессии, как показано в модуле 2, целесообразно производить в нормированных значениях факторов X_r. По тем же соображениям в многофакторном эксперименте также целесообразно перейти от натуральных значений факторов x_r к нормированным X_r (таблица 12). Такой переход осуществляется по следующим формулам:

$$X_r = \frac{x_r - x_{r0}}{\Delta x_r}, \quad r = 1, \ldots, k; \tag{128}$$

$$x_{r0} = \frac{x_{r\,\max} + x_{r\,\min}}{2}, \quad r = 1, \ldots, k; \tag{129}$$

$$\Delta x_r = \frac{x_{r\,\max} - x_{r\,\min}}{2}, \quad r = 1, \ldots, k. \tag{130}$$

Обратный переход от нормированных значений к натуральным производится, как следует из уравнения (128), по формуле

$$x_r = x_{r0} + \Delta x_r \cdot X_r, \quad r = 1, \ldots, k. \tag{131}$$

Таблица 12 – Уровни и интервалы варьирования факторов

Факторы	1-й фактор		r-й фактор		k-й фактор	
	x_1	X_1	x_r	X_r	x_k	X_k
Основной уровень	x_{10}	0	x_{r0}	0	x_{k0}	0
Интервал варьирования	Δx_1	1	Δx_r	1	Δx_k	1
Верхний уровень	$x_{1\max}$	+1	$x_{r\max}$	+1	$x_{k\max}$	+1
Нижний уровень	$x_{1\min}$	–1	$x_{r\min}$	–1	$x_{k\min}$	–1
Формулы перевода натуральных значений факторов x_j в нормированные X_j и обратно	$X_r = \dfrac{x_r - x_{r0}}{\Delta x_r}$; $x_r = x_{r0} + \Delta x_r X_r$, $x_{r0} = \dfrac{x_{r\max} + x_{r\min}}{2}$; $\Delta x_r = \dfrac{x_{r\max} - x_{r\min}}{2}$.					

Следующий этап исследования изучаемого объекта заключается в выборе математической модели. В данном учебно-методическом комплексе, как уже говорилось выше, мы ограничимся рассмотрением только моделей полиномиального вида:

а) математическая модель первого порядка:

$$\eta = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k = \beta_0 + \sum_{r=1}^{k} \beta_r X_r ; \tag{132}$$

б) математическая модель неполного второго порядка:

$$\eta = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k + \beta_{12} X_1 X_2 + \ldots + \beta_{k-1,k} X_{k-1} X_k =$$
$$= \beta_0 + \sum_{r=1}^{k} \beta_r X_r + \sum_{\substack{r=1 \\ r<s}}^{k} \beta_{r,s} X_r X_s ; \tag{133}$$

в) математическая модель второго порядка:

$$\eta = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k + \beta_{12} X_1 X_2 + \ldots + \beta_{k-1,k} X_{k-1} X_k +$$
$$+ \beta_{11} X_1^2 + \ldots + \beta_{kk} X_k^2 = \beta_0 + \sum_{r=1}^{k} \beta_r X_r + \sum_{\substack{r=1 \\ r<s}}^{k} \beta_{r,s} X_r X_s + \sum_{r=1}^{k} \beta_{rr} X_r^2 . \tag{134}$$

Точная оценка коэффициентов математических моделей (132)–(134) потребовала бы очень большого количества опытов. На практике коэффициенты математических моделей, рассчитанные на основе относительно небольшого количества опытов (выборочные данные), являются оценкой истинных значений β и обозначаются как *b*. Поэтому уравнения регрессии (132)–(134), полученные на основе выборочных данных, имеют следующий вид:

а) уравнение регрессии первого порядка:

$$Y = b_0 + b_1 X_1 + b_2 X_2 + \ldots + b_k X_k = b_0 + \sum_{r=1}^{k} b_r X_r , \tag{135}$$

с $B_k = (1+k)$ коэффициентами уравнения регрессии;

б) уравнение регрессии неполного второго порядка:

$$Y = b_0 + b_1 X_1 + b_2 X_2 + \ldots + b_k X_k + b_{12} X_1 X_2 + \ldots + b_{k-1,k} X_{k-1} X_k =$$
$$= b_0 + \sum_{r=1}^{k} b_r X_r + \sum_{\substack{r=1 \\ r<s}}^{k} b_{r,s} X_r X_s , \tag{136}$$

с $B_k = (1 + k + C_k^2) = \left(1 + k + \dfrac{k(k-1)}{2}\right) = \left(\dfrac{k^2 + k + 2}{2}\right)$ коэффициентами уравнения регрессии;

в) уравнение регрессии второго порядка:

$$Y = b_0 + b_1 X_1 + b_2 X_2 + \ldots + b_k X_k + b_{12} X_1 X_2 + \ldots + b_{k-1,k} X_{k-1} X_k +$$
$$+ b_{11} X_1^2 + \ldots + b_{kk} X_k^2 = b_0 + \sum_{r=1}^{k} b_r X_r + \sum_{\substack{r=1 \\ r<s}}^{k} b_{r,s} X_r X_s + \sum_{r=1}^{k} b_{rr} X_r^2 , \tag{137}$$

с $B_k = (1 + k + C_k^2 + k) = \left(1 + k + \dfrac{k(k-1)}{2} + k\right) = \dfrac{(k+1)(k+2)}{2}$ коэффициентами уравнения регрессии.

3.3.1. Теория

Так же, как и для однофакторных, исследование многофакторных объектов целесообразно начинать с построения уравнения регрессии первого порядка (уравнение (135)). Минимальное число уровней факторов, позволяющее построить уравнение регрессии первого порядка (уравнение (135)), равно двум. Эксперимент, реализующий полный перебор всех комбинаций факторов на 2-х уровнях, будем называть полным факторным экспериментом (ПФЭ). Примем эти уровни равными минимальному и максимальному натуральным значениям факторов. Тогда из формул (128)–(130) следует, что уровни нормированных значений факторов равны ± 1 (для краткости уровень +1 и –1 будем обозначать как «+» и «–»). Число опытов ПФЭ при варьировании каждого фактора на 2-х уровнях равно $N_{k0} = 2^k$. Так же, как и в случае однофакторных экспериментов для придания формулам для расчета коэффициентов регрессии и дисперсий их значимости универсального вида (в этом убедимся в дальнейшем), введем фиктивный фактор X_0, все значения которого $X_{0j} = +1$.

План ПФЭ с числом опытов $N_{k0} = 2^k$ позволяет оценить не только k линейных коэффициентов регрессии, но и, как будет показано ниже, коэффициенты регрессии при произведениях факторов $X_r X_s$, $X_r X_s X_v$, $X_r X_s X_v X_w$ (эффекты взаимодействия).

Однако плану ПФЭ свойственны два недостатка. Во-первых, попытка использовать план ПФЭ для оценки квадратичных членов β_{rr} не даст результата, так как столбцы X_{rj}^2 и X_{0j} идентичны, и поэтому свободный член b_0 включает в себя значения и свободного, и квадратичных членов. Математически эффект смешанной оценки коэффициентов регрессии записывается так:

$$b_0 \to \beta_0 + \sum_{r=1}^{k} \beta_{rr} \,. \tag{138}$$

Второй недостаток плана ПФЭ связан с тем, что каждый фактор варьируется только на двух уровнях. Для однофакторного эксперимента при построении уравнения регрессии первого порядка двух уровней концептуально недостаточно, так как через две точки всегда проводится единственная прямая, и проверка уравнения на адекватность невозможна (не остается степеней свободы). Для многофакторных экспериментов при построении уравнения регрессии

первого порядка в принципе можно использовать планы, в которых факторы варьируются на двух уровнях. Однако, если с помощью планов ПФЭ оценивать еще и эффекты взаимодействия, получаемые как произведения факторов $X_r X_s$, $X_r X_s X_v$, $X_r X_s X_v X_w$ и т. п., то число опытов N_{k0} равно числу определяемых коэффициентов регрессии, и, следовательно, проверка уравнения регрессии на адекватность невозможна (не остается ни одной степени свободы).

Для устранения указанных недостатков усилим план ПФЭ, добавив еще один опыт в центре плана. В этом случае, во-первых, каждый фактор будет варьироваться на трех уровнях. Во-вторых, дополнительный опыт в центре плана позволит оценить коэффициент β_0 независимо от суммы квадратичных факторов (уравнение (138)), а также повышает надежность оценки адекватности полученного уравнения за счет дополнительной степени свободы. Композиционный план ПФЭ с дополнительным опытом в центре будем называть ЦПФЭ. Приведем несколько примеров таких планов (таблицы 13 и 14):

Таблица 13 – План ЦПФЭ для $k = 2$

$N_{20} + 1$	X_{0j}	X_{1j}	X_{2j}
1	+	–	–
2	+	+	–
3	+	–	+
4	+	+	+
5	+	0	0

Таблица 14 – План ЦПФЭ для $k = 3$

$N_{30} + 1$	X_{0j}	X_{1j}	X_{2j}	X_{3j}
1	+	–	–	–
2	+	+	–	–
3	+	–	+	–
4	+	+	+	–
5	+	–	–	+
6	+	+	–	+
7	+	–	+	+
8	+	+	+	+
9	+	0	0	0

Перечислим основные свойства планов ЦПФЭ.

1. Число опытов $N_k = N_{k0} + 1 = 2^k + 1$, где k – число факторов (убедитесь в этом самостоятельно).

2. При построении плана полный перебор всех возможных комбинаций уровней достигается следующим образом: для первого фактора уровни варьирования чередуются; для второго фактора чередуются пары одинаковых уровней; для третьего чередуются четыре одинаковых уровня и т. д. Кроме того, добавляется еще один опыт в центре, в котором значения факторов равны 0.

3. Выполняются следующие соотношения:

$$\sum_{j=1}^{N_k} X_{0j} = N_{k0} + 1; \tag{139}$$

$$\sum_{j=1}^{N_k} X_{0j}^2 = N_k = N_{k0} + 1 \ \text{(нормирование)}; \tag{140}$$

$$\sum_{j=1}^{N_k} X_{rj} = 0, \ r = 1, ..., k \ \text{(симметричность)}; \tag{141}$$

$$\sum_{j=1}^{N_k} X_{rj}^2 = N_{k0}, \ r = 0, ..., k \ \text{(нормирование)}. \tag{142}$$

4. Все факторы ЦПФЭ ортогональны (проверить самостоятельно):

$$\begin{cases} \sum_{j=1}^{N_k} X_{0j} X_{rj} = 0, \ r = 1, ..., k; \\ \sum_{j=1}^{N_k} X_{rj} X_{sj} = 0, \quad r < s, \ r = 1, ..., k-1, \ s = 2, ..., k. \end{cases} \tag{143}$$

Выведем формулы для расчета коэффициентов многофакторного уравнения регрессии первого порядка (уравнение (135)) на примере двухфакторного уравнения $Y = b_0 X_0 + b_1 X_1 + b_2 X_2$. Методом наименьших квадратов составим систему уравнений, аналогичную системе (64):

$$\begin{cases} b_0 \sum_{j=1}^{N_k} X_{0j}^2 + b_1 \sum_{j=1}^{N_k} X_{0j} X_{1j} + b_2 \sum_{j=1}^{N_k} X_{0j} X_{2j} = \sum_{j=1}^{N_k} X_{0j} \overline{Y}_j; \\ b_0 \sum_{j=1}^{N_k} X_{0j} X_{1j} + b_1 \sum_{j=1}^{N_k} X_{1j}^2 + b_2 \sum_{j=1}^{N_k} X_{1j} X_{2j} = \sum_{j=1}^{N_k} X_{1j} \overline{Y}_j; \\ b_0 \sum_{j=1}^{N_k} X_{0j} X_{2j} + b_1 \sum_{j=1}^{N_k} X_{1j} X_{2j} + b_2 \sum_{j=1}^{N_k} X_{2j}^2 = \sum_{j=1}^{N_k} X_{2j} \overline{Y}_j. \end{cases} \tag{144}$$

С учетом взаимной ортогональности всех факторов (уравнение (143)) система уравнений (144) приобретает диагональный вид:

$$\begin{cases} b_0 \sum_{j=1}^{N_k} X_{0j}^2 + 0 + 0 = \sum_{j=1}^{N_k} X_{0j} \overline{Y}_j; \\ 0 + b_1 \sum_{j=1}^{N_k} X_{1j}^2 + 0 = \sum_{j=1}^{N_k} X_{1j} \overline{Y}_j; \ . \\ 0 + 0 + b_2 \sum_{j=1}^{N_k} X_{2j}^2 = \sum_{j=1}^{N_k} X_{2j} \overline{Y}_j. \end{cases} \tag{145}$$

Из системы уравнений (145) с учетом уравнений (140), (142) следует, что

$$b_0 = \frac{\sum_{j=1}^{N_k} X_{0j} \overline{Y}_j}{\sum_{j=1}^{N_k} X_{0j}^2} = \frac{\sum_{j=1}^{N_k} X_{rj} \overline{Y}_j}{N_k} = \frac{\sum_{j=1}^{N_k} X_{rj} \overline{Y}_j}{N_{k0} + 1}; \tag{146}$$

$$b_r = \frac{\sum_{j=1}^{N_k} X_{rj} \overline{Y}_j}{\sum_{j=1}^{N_k} X_{rj}^2} = \frac{\sum_{j=1}^{N_k} X_{rj} \overline{Y}_j}{N_{k0}}, \ r = 1, 2. \tag{147}$$

Можно показать, что уравнение (147) справедливо для любого числа факторов $r = 1, \ldots, k$ (убедитесь в этом самостоятельно).

Статистический анализ многофакторных уравнений регрессии первого порядка, построенных по планам ЦПФЭ, (значимость коэффициентов, адекватность уравнения) во многом совпадает с анализом однофакторных уравнений регрессии первого порядка (см. раздел 2.3.1).

Проверка коэффициентов уравнения регрессии на значимость по критерию Стьюдента.

1. Рассчитать дисперсию воспроизводимости $S^2_{\text{воспр}}$:

$$S^2_{\text{воспр}} = \frac{\sum_{j=1}^{N_k} S_j^2}{N_k}, \ f = N_k\left(n-1\right). \qquad (148)$$

2. Рассчитать дисперсии значимости коэффициентов регрессии $S^2(b_0), \ S^2(b_r)$:

$$S^2(b_0) = \frac{S^2_{\text{воспр}}}{n\sum_{j=1}^{N_k} X_{0j}^2} = \frac{S^2_{\text{воспр}}}{nN_k}; \qquad (149)$$

$$S^2(b_r) = \frac{S^2_{\text{воспр}}}{n\sum_{j=1}^{N_k} X_{rj}^2} = \frac{S^2_{\text{воспр}}}{nN_{k0}}, \qquad r = 1, ..., k. \qquad (150)$$

Доверительные интервалы коэффициентов регрессии Δb_r рассчитываются по уравнению (78) по критерию Стьюдента при числе степеней свободы $f = N_k\left(n-1\right)$.

Проверка на адекватность многофакторного уравнения регрессии первого порядка, полученного по плану ЦПФЭ, выполняется по такой же процедуре, что и для однофакторного уравнения регрессии первого порядка (см. уравнения (79)–(86)).

Если полученное многофакторное уравнение регрессии первого порядка адекватно, то оно позволяет оценить предельную абсолютную погрешность параметра Y в исследованном факторном пространстве. Рассчитаем предельную абсолютную погрешность параметра Y, например, по двухфакторному уравнению регрессии первого порядка:

$$Y = b_0 X_0 + b_1 X_1 + b_2 X_2. \qquad (151)$$

Так как факторы этого уравнения X_0, X_1, X_2 ортогональны (уравнение (143)), то b_0, b_1, b_2 – независимые коэффициенты, каждый из которых имеет свой доверительный интервал значимости. Отметим, что X_0, X_1, X_2 – детерминированные факторы. Применяя к уравнению (151) уравнения (36), (37), получим (вывод приведен в приложении 6)

$$\Delta Y(X_1, X_2) = t_{N_k(n-1),\, p} \cdot S(Y), \qquad (152)$$

где

$$S^2(Y) = S^2(b_0) + X_1^2 S^2(b_1) + X_2^2 S^2(b_2). \qquad (153)$$

С учетом уравнений (149), (150) уравнение (153) примет следующий вид:

$$S^2(Y) = \frac{S^2_{\text{воспр}}}{nN_k}\left[1 + \frac{N_k}{N_{k0}}\left(X_1^2 + X_2^2\right)\right]. \qquad (154)$$

Подставляя уравнение (154) в (152), получаем

$$\Delta Y(X_1, X_2) = t_{N_k(n-1),\, p} \cdot \frac{S_{\text{воспр}}}{\sqrt{nN_k}} \sqrt{1 + \frac{N_k}{N_{k0}}\left(X_1^2 + X_2^2\right)}. \qquad (155)$$

Можно показать, что в общем случае для произвольного числа факторов уравнение (155) будет иметь следующий вид (убедитесь в этом самостоятельно):

$$\Delta Y(X_1, ..., X_k) = t_{N_k(n-1),\, p} \cdot \frac{S_{\text{воспр}}}{\sqrt{nN_k}} \sqrt{1 + \frac{N_k}{N_{k0}} \sum_{r=1}^{k} X_r^2}. \qquad (156)$$

Анализ уравнения (156) позволяет сделать несколько выводов:

а) минимальная предельная абсолютная погрешность достигается в центре плана $X_1 = X_2 = \ldots = X_k = 0$:

$$\Delta Y(0) = t_{N_k(n-1),\, p} \cdot \frac{S_{\text{воспр}}}{\sqrt{nN_k}}; \qquad (157)$$

б) максимальная предельная абсолютная погрешность достигается на границах факторного пространства при $X_1 = X_2 = \ldots = X_k = \pm 1$:

$$\Delta Y(\pm 1, ..., \pm 1) = t_{N_k(n-1),\, p} \cdot \frac{S_{\text{воспр}}}{\sqrt{nN_k}} \sqrt{1 + \frac{kN_k}{N_{k0}}}; \qquad (158)$$

в) предельная абсолютная погрешность параметра, рассчитанного по адекватному уравнению регрессии первого порядка, значительно меньше предельной абсолютной погрешности, найденной по экспериментальным данным в каждом опыте.

Подтвердим количественно это утверждение. Предельная абсолютная погрешность в каждом опыте $\Delta\overline{Y}_j$ рассчитывается так же, как и для однофакторного уравнения регрессии первого и второго порядка (см. уравнения (95)–(97)). Соотношение предельных экспериментальных и расчетных ошибок в центре плана ($X = 0$) для двухфакторного плана, используя уравнения (157) и (95), равно

$$\frac{\Delta\overline{Y}_j}{\Delta Y(0)} \approx \frac{t_{n-1,\,p} \cdot S_{\text{воспр}} \cdot \sqrt{nN_k}}{\sqrt{n} \cdot t_{N_k(n-1),\,p} \cdot S_{\text{воспр}}} = \frac{t_{n-1,\,p}\sqrt{N_k}}{t_{N_k(n-1),\,p}}. \tag{159}$$

Так как в двухфакторном эксперименте $N_k = 5$, $n \approx 4$, то $t_{15,\,0.95} \approx 2$, $t_{3,\,0.95} \approx 3$ (приложение 2) и из уравнения (159) следует, что в центре плана

$$\frac{\Delta\overline{Y}_j}{\Delta Y(0)} = \frac{3\sqrt{5}}{2} \approx 3.4\,. \tag{160}$$

Предельная абсолютная погрешность параметра Y в центре плана, рассчитанная по многофакторному уравнению регрессии первого порядка, построенному по плану ЦПФЭ, в 3.4 раза меньше предельной абсолютной погрешности выборочных средних, найденных экспериментально в каждом опыте. Для трехфакторного эксперимента $N_k = 9$ и приведенное выше отношение будет равно 4.5, для четырехфакторного эксперимента $N_k = 17$ отношение равно 6.2 (убедитесь в этом самостоятельно)

Таким образом, можно опять повторить: математическая теория эксперимента позволяет делать прогноз точнее, чем породивший ее эксперимент.

3.3.2. Алгоритм решения прикладных задач

1. Выбрать объект исследования и на основе литературных и патентных источников собрать необходимую априорную информацию о факторах и параметрах, характеризующих исследуемый объект.

2. Выбрать параметр Y, который наиболее полно характеризует исследуемый объект, а также число дублей в каждом опыте. Количество дублей n выбирается интуитивно (рекомендуем не менее 3).

3. Выбрать варьируемые факторы x_k, которые влияют на изменение параметра, а также диапазон их варьирования $[x_{k\,\min}, x_{k\,\max}]$. С учетом того, что в планах ЦПФЭ каждый фактор варьируется только на 3-х уровнях, натуральные значения факторов будут равняться максимальному, основному и минимальному значениям, а нормированные ± 1 и 0.

4. Построить матрицу планирования для ЦПФЭ с числом опытов $N_k = N_{k0} + 1 = 2^k + 1$. Создать фиктивный фактор X_0 ($X_{0j} = +1$).

5. Обеспечить приборное и технологическое оснащение эксперимента, позволяющее измерять исследуемый параметр Y и варьируемые факторы x_r с требуемой точностью. Подобрать технологическое оборудование, позволяющее поддерживать варьируемые факторы на требуемых уровнях. Провести эксперимент.

6. Провести предварительную обработку экспериментальных данных.

6.1. Рассчитать выборочное среднее в каждом опыте:

$$\overline{Y}_j = \frac{\sum_{i=1}^{n} Y_{ji}}{n}, \quad j = 1, \ldots, N_k.$$

6.2. Рассчитать выборочную дисперсию в каждом опыте:

$$S_j^2 = \frac{\sum_{i=1}^{n}\left(Y_j - \overline{Y}_j\right)^2}{n-1}, \quad j = 1, \ldots, N_k.$$

6.3. Проверить экспериментальные данные каждого опыта на промахи по критерию Смирнова–Граббса (приложение 1):

$$\tau_{эj} = \max\left|\frac{Y_{ji} - \overline{Y}_j}{S_j}\right| > \tau_{n-2,\,p}, \quad j = 1, \ldots, N_k,$$

где $\tau_{n-2,\ p}$ – критическое значение критерия Смирнова–Граббса (приложение 1).

6.4. Проверить экспериментальные данные каждого опыта на принадлежность к нормальному закону распределения:

$$\left|\frac{CAO_j}{S_j} - 0.7979\right| = \left|\frac{\sum\limits_{i=1}^{n}\left|Y_{ji} - \overline{Y}_j\right|}{nS_j} - 0.7979\right| < \frac{0.4}{\sqrt{n}}, \quad j = 1, \ldots, N_k.$$

6.5. Проверить дисперсии всех опытов на однородность по Критерию Кохрена (приложение 5):

$$\frac{\max(S_j^2)}{\sum\limits_{j=1}^{N_k} S_j^2} < G_{n-1,\, N_k,\, p}\,,$$

где $G_{n-1,\, N_k,\, p}$ – критическое значение критерия Кохрена (приложение 5).

Внимание! Дальнейшая обработка экспериментальных данных с целью построения уравнения регрессии возможна только в том случае, если в каждом опыте нет промахов, данные каждого опыта подчиняются нормальному закону распределения, а дисперсии всех опытов однородны. Все опыты, в которых эти требования не выполнены, следует переделать.

7. Перейти к построению уравнения регрессии первого порядка $Y = b_0 X_0 + b_1 X_1 + \ldots b_k X_k$. Рассчитать коэффициенты регрессии $b_0, b_1, \ldots, b_k$ по формулам:

$$b_0 = \frac{\sum\limits_{j=1}^{N_k} X_{0j}\overline{Y}_j}{\sum\limits_{j=1}^{N_k} X_{0j}^2} = \frac{\sum\limits_{j=1}^{N_k} X_{rj}\overline{Y}_j}{N_k} = \frac{\sum\limits_{j=1}^{N_k} X_{rj}\overline{Y}_j}{N_{k0}+1};$$

$$b_r = \frac{\sum\limits_{j=1}^{N_k} X_{rj}\overline{Y}_j}{\sum\limits_{j=1}^{N_k} X_{rj}^2} = \frac{\sum\limits_{j=1}^{N_k} X_{rj}\overline{Y}_j}{N_{k0}}, \quad r = 1, \ldots, k.$$

8. Проверить коэффициенты регрессии $b_0, b_1, \ldots, b_k$ на значимость по критерию Стьюдента.

8.1. Рассчитать дисперсию воспроизводимости:

$$S^2_{\text{воспр}} = \frac{\sum\limits_{j=1}^{N_k} S_j^2}{N_k}, \quad f = N_k(n-1).$$

8.2. Рассчитать дисперсии значимости для коэффициентов регрессии $b_0, b_1, \ldots, b_k$ следующим способом:

$$S^2(b_0) = \frac{S^2_{\text{воспр}}}{n\sum\limits_{j=1}^{N_k} X_{0j}^2} = \frac{S^2_{\text{воспр}}}{nN_k};$$

$$S^2(b_r) = \frac{S^2_{\text{воспр}}}{n\sum\limits_{j=1}^{N_k} X_{rj}^2} = \frac{S^2_{\text{воспр}}}{nN_{k0}}, \quad r = 1, \ldots, k.$$

8.3. Рассчитать доверительный интервал для коэффициентов регрессии $b_0, b_1, \ldots, b_k$:

$$\Delta b_r = t_{N_k(n-1),\, p} \cdot S(b_r), \quad r = 0, 1, \ldots, k,$$

где $t_{N_k(n-1),\, p}$ – критическое значение критерия Стьюдента (приложение 2).

8.4. Коэффициенты, для которых $\Delta b_r > |b_r|$ незначимы, исключить из уравнения регрессии.

9. Проверить уравнения регрессии первого порядка на адекватность по критерию Фишера.

9.1. Рассчитать значения параметра Y_j^p по уравнению регрессии первого порядка $Y = b_0 X_0 + b_1 X_1 + \ldots b_k X_k$ в каждом опыте $j = 1, \ldots, N_k$.

9.2. Рассчитать квадрат разности $(\overline{Y}_j - Y_j^p)^2$ экспериментальных $\overline{Y}_j$ и расчетных значений Y_j^p в каждом опыте $j = 1, \ldots, N_k$, а также остаточную сумму квадратов φ для всех опытов:

$$\varphi = \sum_{j=1}^{N_k}(\overline{Y}_j - Y_j^p)^2.$$

9.3. Рассчитать дисперсию адекватности:

$$S_{\text{ад}}^2 = \frac{n\varphi}{N_k - B}, \quad f = N_k - B.$$

9.4. Рассчитать экспериментальное значение критерия Фишера:

$$F_{\text{э}} = \frac{\max(S_{\text{ад}}^2, S_{\text{воспр}}^2)}{\min(S_{\text{ад}}^2, S_{\text{воспр}}^2)}.$$

9.5. Если $F_{\text{э}} = \dfrac{S_{\text{ад}}^2}{S_{\text{воспр}}^2} < F_{(N_k - B),\, N_k(n-1),\, p}$ при $S_{\text{ад}}^2 > S_{\text{воспр}}^2$,

где $F_{(N_k - B),\, N_k(n-1),\, p}$ – критическое значение критерия Фишера (приложение 4), то $\sigma_{\text{ад}}^2 = \sigma_{\text{воспр}}^2$ и уравнение регрессии первого порядка адекватно. Перейти к экспериментам по крутому восхождению (см. раздел 3.5.1).

9.6. Если $F_{\text{э}} = \dfrac{S_{\text{ад}}^2}{S_{\text{воспр}}^2} > F_{(N_k - B),\, N_k(n-1),\, p}$ при $S_{\text{ад}}^2 > S_{\text{воспр}}^2$,

то $\sigma_{\text{ад}}^2 > \sigma_{\text{воспр}}^2$ и уравнение регрессии первого порядка неадекватно. Перейти к построению уравнения регрессии второго порядка с использованием ОЦКП (см. раздел 3.6.1).

9.7. Если $F_{\text{э}} = \dfrac{S_{\text{воспр}}^2}{S_{\text{ад}}^2} < F_{N_k(n-1),\, (N_k - B),\, p}$ при $S_{\text{ад}}^2 < S_{\text{воспр}}^2$,

где $F_{N_k(n-1),\, (N_k - B),\, p}$ – критическое значение критерия Фишера (приложение 4), то $\sigma_{\text{ад}}^2 = \sigma_{\text{воспр}}^2$ и уравнение регрессии первого порядка адекватно. Перейти к экспериментам по крутому восхождению (см. раздел 3.5.1).

9.8. Если $F_{\text{э}} = \dfrac{S_{\text{воспр}}^2}{S_{\text{ад}}^2} > F_{N_k(n-1),\, (N_k - B),\, p}$ при $S_{\text{ад}}^2 < S_{\text{воспр}}^2$,

то $\sigma_{\text{воспр}}^2 > \sigma_{\text{ад}}^2$. Вероятность такого случая на практике мала. Такой случай является следствием либо плохой организации эксперимента, в котором нарушена идея случайности (нарушены основные статистические предпосылки), либо недостаточно степеней свободы для оценки дисперсии адекватности. В этом случае необходимо повторить эксперимент, причем желательно увеличить число дублей n.

10. Расчет предельной абсолютной погрешности:

$$\Delta Y(X_1, ..., X_k) = t_{N_k(n-1),\, p} \cdot \frac{S_{\text{воспр}}}{\sqrt{nN_k}} \sqrt{1 + \frac{N_k}{N_{k0}} \sum_{r=1}^{k} X_r^2}.$$

11. В заключение, если есть в этом необходимость, построить многофакторное уравнение регрессии в натуральных значениях факторов. Для выполнения этой процедуры в уравнении регрессии первого порядка в обычной форме нормированные значения факторов X_r необходимо заменить на натуральные значения x_r по формуле

$$X_r = \frac{x_r - x_{r0}}{\Delta x_r}.$$

3.3.3. Базовая задача

Формулировка условия: изучить влияние количества посевного материала и неорганического удобрения на урожайность пшеницы.

Анализ исследуемого объекта: изучаемый параметр – урожайность пшеницы Y (ц/га), которая зависит от количества посевного материала x_1 (ц/га) и неорганического удобрения x_2 (ц/га).

Планирование эксперимента. Ссылки на алгоритм 3.3.2.

Параметр. Для повышения статистической достоверности изучаемого параметра примем число дублей в каждом опыте $n = 4$ (п. 2).

Факторы. На основе литературных данных выбраны основные уровни и интервалы варьирования факторов x_r (таблица 15) (п. 3).

План эксперимента. Так как априори характер искомой зависимости неизвестен, начнем с построения уравнения регрессии первого порядка по плану ЦПФЭ (п. 4). В случае ЦПФЭ для двух факторов ($k = 2$) количество опытов $N_{k0} = 2^k = 4$, $N_k = N_{k0} = 2^k + 1 = 2^2 + 1 = 5$. План ЦПФЭ для двух факторов представлен в таблице 16.

Методика эксперимента. Для уменьшения влияния неучтенных факторов эксперимент проводится в теплице. Так как общее количество опытов $N_k = 5$ и число дублей $n = 4$, организуем 20 делянок по 1 м^2. Количество посевного материала и неорганического удобрения вносятся в каждую делянку согласно плану ЦПФЭ. Все остальные факторы (количество воды, света, тепла, сроки посева и сбора и т. п.) поддерживаются на фиксированных, одинаковых для всех делянок уровнях.

Таблица 15 – Уровни и интервалы варьирования факторов

Факторы	1-й фактор (семена)		2-й фактор (удобрение)	
	x_1, ц/га	X_1	x_2, ц/га	X_2
Основной уровень	$x_{10} = 1$	0	$x_{20} = 0.5$	0
Интервал варьирования	$\Delta x_1 = 0.5$	–	$\Delta x_2 = 0.25$	–
Верхний уровень	$x_{1\,max} = 1.5$	+1	$X_{2\,max} = 0.75$	+1
Нижний уровень	$x_{1\,min} = 0.5$	–1	$X_{2\,min} = 0.25$	–1
Формулы перевода натуральных значений факторов x_r в нормированные X_r и обратно	$X_{1j} = \dfrac{x_{1j} - 1}{0.5}$; $x_{1j} = 1 + 0.5X_{1j}$ $X_{2j} = \dfrac{x_{1j} - 0.5}{0.25}$; $x_{2j} = 0.5 + 0.25X_{2j}$			

Предварительная обработка экспериментальных данных.
Вычисление выборочного среднего в каждом опыте (п. 6.1):

$$\overline{Y}_j = \frac{\sum_{i=1}^{n} Y_{ji}}{n}, \quad j = 1, \ldots, N_k.$$

Например, $\overline{Y}_2 = \dfrac{29.6 + 29.5 + 29.9 + 29.9}{4} = 29.73$.

Вычисление выборочной дисперсии в каждом опыте (п. 6.2):

$$S_j^2 = \frac{\sum_{i=1}^{n} (Y_{ji} - \overline{Y}_j)^2}{n-1}, \quad j = 1, \ldots, N_k.$$

Например, $S_2^2 = \dfrac{(29.6 - 29.73)^2 + (29.5 - 29.73)^2}{4-1} +$

$+\dfrac{(29.9 - 29.73)^2 + (29.9 - 29.73)^2}{4-1} = 0.0425$, $S_2 = \sqrt{0.0425} = 0.206$.

Таблица 16 – Матрица планирования и результаты анализа экспериментальных данных

$N_{20}+1$	X_{0j}	X_{1j}	X_{2j}	Y_{j1}	Y_{j2}	Y_{j3}	Y_{j4}	$\overline{Y}_j$	S_j^2	Y_j^p	$(\overline{Y}_j - Y_j^p)^2$
1	1	–	–	16.2	15.8	15.8	16.2	16.00	0.0533	16.06	0.0036
2	1	+	–	29.6	29.5	29.9	29.9	29.73	0.0425	29.60	0.0169
3	1	–	+	27.2	27.5	27.7	27.3	27.43	0.0492	27.35	0.0064
4	1	+	+	40.6	40.7	41.0	40.8	40.78	0.0292	40.84	0.0036
5	1	0	0	28.5	28.2	28.5	28.1	28.33	0.0425	28.45	0.0144
$\sum_{j=1}^{5} X_{rj}$	5	0	0					$\sum_{j=1}^{5} S_j^2 = 0.217$		$\varphi = 0.0449$	
$\sum_{j=1}^{5} X_{rj}^2$	5	4	4					$S_{\text{воспр}}^2 = 0.0434$		$S_{\text{ад}}^2 = 0.0898$	
b_r	28.45	6.77	5.62							$F_{э} = 2.07$	
$S_{b_r}^2$	0.0022	0.0027	0.0027							$F_{2,\,15,\,0.95} = 3.68$	
S_{b_r}	0.047	0.052	0.052	$t_{15,\,095} = 2.13$						Модель адекватна	
Δb_r	0.10	0.11	0.11								

Анализ результатов на промахи (п. 6.3), показал, что во всех опытах промахов нет. Например, для опыта № 2 наиболее выделяющимся значением является $Y_{2,2} = 29.5$. Критерий Смирнова–Граббса, рассчитанный по уравнению (37), равен:

$$\tau_{э2} = \max\left|Y_{2,i} - \overline{Y}_2\right| / S_2 = \left|29.5 - 29.73\right| / 0.206 = 1.17 .$$

Промахов в опыте № 2 нет, так как критическое значение критерия Смирнова–Граббса, найденное из приложения 1 при числе степеней свободы $f = n - 2 = 4 - 2 = 2$, равно $\tau_{2,\ 0.95} = 1.69$ ($\tau_{\min} = 1.17 < \tau_{3,\ 0.95} = 1.69$).

Анализ на принадлежность случайных значений параметра $\boldsymbol{Y_{ji}}$ в каждом опыте к нормальному закону распределения (п. 6.4) показал, что экспериментальные данные подчиняются нормальному закону распределения. Например, случайные значения параметра Y_{2i} подчиняются нормальному закону распределения, так как

$$\left|САО_2 / S_2 - 0.7979\right| < 0.4/\sqrt{n} \rightarrow \left|0.175/0.206 - 0.7979\right| < 0.4/\sqrt{4} \rightarrow$$

$$\rightarrow 0.0516 < 0.2 \text{, где } САО_2 = \frac{\sum_{i=1}^{n}\left|Y_{2,i} - \overline{Y}_2\right|}{n} = \frac{\left|29.6 - 29.73\right|}{4} +$$

$$+ \frac{\left|29.5 - 29.73\right| + \left|29.9 - 29.73\right| + \left|29.9 - 29.73\right|}{4} = 0.175 .$$

Проверка дисперсий всех опытов на однородность по критерию Кохрена (п. 6.5) показала, что все дисперсии однородны, так как $G_э < G_{n-1,\ N_k,\ p} = G_{3,\ 5,\ 0.94} = 0.60$ (приложение 5).

$$G_э = \frac{\max(S_j^2)}{\sum_{j=1}^{5} S_j^2} = \frac{0.0533}{0.0533 + 0.0425 + 0.0492 + 0.0292 + 0.0425} =$$

$$= \frac{0.0533}{0.217} = 0.25 .$$

Поскольку предварительная обработка экспериментальных данных в каждом опыте не выявила отклонений от основных предпосылок регрессионного анализа (отсутствие промахов, выполняется

нормальный закон распределения, все дисперсии однородны), переходим к следующему этапу – построению многофакторного уравнения регрессии первого порядка.

Построение уравнения регрессии первого порядка (п. 7) $Y = b_0 X_0 + b_1 X_1 + b_2 X_2$.

Образуем столбцы $X_{0j}\overline{Y}_j$, $X_{1j}\overline{Y}_j$, $X_{2j}\overline{Y}_j$ и рассчитаем их суммы:

$$\sum_{j=1}^{5} X_{0j}\overline{Y}_j = 1 \cdot 16.00 + 1 \cdot 29.73 + 1 \cdot 27.43 + 1 \cdot 40.78 + 1 \cdot 28.33 = 142.3 ,$$

$$\sum_{j=1}^{5} X_{0j}\overline{Y}_j = (-1) \cdot 16.00 + 1 \cdot 29.73 + (-1) \cdot 27.43 +$$

$$1 \cdot 40.78 + 0 \cdot 28.33 = 27.08 ,$$

$$\sum_{j=1}^{5} X_{2j}\overline{Y}_j = (-1) \cdot 16.00 + (-1) \cdot 29.73 + 1 \cdot 27.43 +$$

$$1 \cdot 40.78 + 0 \cdot 28.33 = 22.48 .$$

Рассчитаем коэффициенты регрессии b_0, b_1, b_2 (п. 7):

$$b_0 = \frac{\sum_{j=1}^{N_k} X_{rj}\overline{Y}_j}{N_k} = \frac{142.27}{5} = 28.45 ,$$

$$b_1 = \frac{\sum_{j=1}^{N_k} X_{1j}\overline{Y}_j}{N_{k0}} = \frac{27.08}{4} = 6.77 ,$$

$$b_2 = \frac{\sum_{j=1}^{N_k} X_{2j}\overline{Y}_j}{N_{k0}} = \frac{22.48}{4} = 5.62 .$$

Проверка коэффициентов уравнения регрессии первого порядка на значимость по критерию Стьюдента.

Расчет дисперсии воспроизводимости (п. 8.1):

$$S^2_{\text{воспр}} = \frac{\sum_{j=1}^{N_k} S_j^2}{N_k} = \frac{0.217}{5} = 0.0434 .$$

Расчет дисперсии значимости коэффициентов регрессии b_0, b_1, b_2 (п. 8.2):

$$S^2(b_0) = \frac{S^2_{\text{воспр}}}{nN_k} = \frac{0.0434}{5 \cdot 4} = 0.0022 \rightarrow S(b_0) = 0.0466 ,$$

$$S^2(b_1) = S^2(b_2) = \frac{S^2_{\text{воспр}}}{nN_{k0}} = \frac{0.0434}{4 \cdot 4} = 0.0027 \rightarrow S(b_1) = S(b_2) = 0.0521 .$$

Расчет доверительных интервалов коэффициентов (п. 8.3):

$$\Delta b_0 = t_{N_k(n-1),\, p} \cdot S(b_0) = 2.131 \cdot 0.0466 = 0.0993 \approx 0.10 ,$$

$$\Delta b_1 = \Delta b_2 = t_{N_k(n-1),\, p} \cdot S(b_1) = 2.131 \cdot 0.0521 = 0.1110 \approx 0.11 ,$$

где $t_{N_k(n-1),\, p} = t_{15,\, 0.95} = 2.131$ – критическое значение критерия Стьюдента (приложение 2).

Все коэффициенты регрессии b_0, b_1, b_2 значимы, так как $\Delta b_0 < |b_0|$ (0.10 < 28.45); $\Delta b_1 < |b_1|$ (0.11 < 6.77); $\Delta b_2 < |b_2|$ (0.11 < 5.62) (п. 8.4). Таким образом, в уравнении регрессии первого порядка $Y = 28.48 + 6.77X_1 + 5.62X_2$ все коэффициенты значимы.

Анализ уравнения регрессии первого порядка на адекватность по критерию Фишера.

Значения параметра Y_j^p рассчитываются по уравнению $Y = 28.45 + 6.77X_1 + 5.62X_2$ (п. 9.1).

Например, $Y_1^p = 28.45 + 6.77 \cdot (-1) + 5.62 \cdot (-1) = 16.06$.

Остаточная сумма квадратов φ (п. 9.2):

$$\varphi = \sum_{j=1}^{5} (\overline{Y}_j - Y_j^p)^2 == (16.00 - 16.09)^2 + (29.73 - 29.60)^2 +$$

$$(27.43 - 27.35)^2 + (40.78 - 40.84)^2 + (28.33 - 28.45)^2 =$$

$$= 0.0036 + 0.0169 + 0.0064 + 0.0036 + 0.0144 = 0.0449 .$$

Дисперсия адекватности (п. 9.3):

$$S^2_{\text{ад}} = \frac{n\varphi}{N_k - B} = \frac{4 \cdot 0.0499}{5 - 3} = 0.0898 ,\ f = N_k - B = 4 + 1 - 3 = 2 .$$

Экспериментальное значение критерия Фишера (п. 8.4):

$$F_{\text{э}} = \frac{S^2_{\text{ад}}}{S^2_{\text{воспр}}} = \frac{0.0898}{0.0434} = 2.07 \text{, так как } S^2_{\text{ад}} > S^2_{\text{воспр}} .$$

Так как $F_{\text{э}} = 2.07 < F_{(N_k - B),\, N_k(n-1),\, p} = F_{2,\, 15,\, 0.95} = 3.68$ (приложение 4), то $\sigma^2_{\text{ад}} = \sigma^2_{\text{воспр}}$ и полученное уравнение регрессии первого порядка $Y = 28.48 + 6.77X_1 + 5.62X_2$ адекватно (п. 8.5).

Интерпретация уравнения $Y = 28.45 + 6.77X_1 + 5.62X_2$.

1. Из полученного уравнения следует, что для повышения урожайности пшеницы необходимо увеличивать и количество посевного материала, и количество неорганического удобрения.

2. Так как линейное уравнение адекватно, то для нахождения условий, соблюдение которых существенно повысит урожайность пшеницы, целесообразно применить метод крутого восхождения (см. 3.5.3).

3.3.4. Материалы для самостоятельной работы, лабораторных и практических занятий

Тесты

1. Уравнения пересчета натуральных и нормированных значений факторов в многофакторном эксперименте:

а) $X_r = \dfrac{x_r - x_{r\max} - x_{r\min}}{\Delta x_r}$, $x_r = \Delta x_r X_r$;

б) $X_r = \dfrac{x_r - x_{r0}}{x_{r\max} + x_{r\min}}$, $x_r = x_{r0} + X_r$;

в) $X_r = \dfrac{x_r - x_{r0}}{\Delta x_r}$, $x_r = x_{ro} + \Delta x_r X_r$.

где $x_{r0} = \dfrac{x_{r\max} + x_{r\min}}{2}$, $\Delta x_r = \dfrac{x_{r\max} - x_{r\min}}{2}$, $r = 1, ..., k$.

2. Двухуровневый план ЦПФЭ для $k = 2$:

а)

$N_{20} + 1$	X_{0j}	X_{1j}	X_{2j}
1	+	–	–
2	+	–	–
3	+	–	+
4	+	+	+
5	+	0	0

б)

$N_{20} + 1$	X_{0j}	X_{1j}	X_{2j}
1	+	–	+
2	+	+	–
3	+	–	+
4	+	+	+
5	+	0	0

в)

$N_{20} + 1$	X_{0j}	X_{1j}	X_{2j}
1	+	–	–
2	+	+	–
3	+	–	+
4	+	+	+
5	+	0	0

г)

$N_{20} + 1$	X_{0j}	X_{1j}	X_{2j}
1	+	–	+
2	+	+	+
3	+	–	–
4	+	+	–
5	+	0	0

3. Трехуровневый план ЦПФЭ для $k = 3$:

а)

$N_{30} + 1$	X_{0j}	X_{1j}	X_{2j}	X_{3j}
1	+	–	–	–
2	+	+	–	+
3	+	–	+	–
4	+	+	+	+
5	+	–	–	–
6	+	+	–	+
7	+	–	+	–
8	+	+	+	+
9	+	0	0	0

б)

$N_{30} + 1$	X_{0j}	X_{1j}	X_{2j}	X_{3j}
1	–	+	–	–
2	+	+	–	–
3	–	+	+	–
4	+	+	+	–
5	–	+	–	+
6	+	+	–	+
7	–	+	+	+
8	+	+	+	+
9	+	0	0	0

в)

$N_{30} + 1$	X_{0j}	X_{1j}	X_{2j}	X_{3j}
1	+	–	–	–
2	+	+	–	–
3	+	–	+	–
4	+	+	+	–
5	+	–	–	+
6	+	+	–	+
7	+	–	+	+
8	+	+	+	+
9	+	0	0	0

4. Число опытов N_k в ЦПФЭ при $k = 2$:

а) $N_k = 2^n + 1$; б) $N_k = k^2$; в) $N_k = 2^k + 1$; г) $N_k = 3^k + 1$.

5. Коэффициенты уравнения регрессии первого порядка многофакторного ЦПФЭ ($r = 0, 1, \ldots, k$):

$$\text{а)}\ b_r = \frac{\sum_{j=1}^{N_k} X_{rj}\overline{Y}_j}{\sum_{j=1}^{N_k} X_{rj}^2};\quad \text{б)}\ b_r = \frac{\sum_{j=1}^{N_k} X_{rj}\overline{Y}_j}{N_k};$$

$$\text{в)}\ b_r = \frac{\sum_{j=1}^{N_{k0}} X_{rj}\overline{Y}_j}{N_{k0}};\quad \text{г)}\ b_r = \frac{\sum_{j=1}^{N_{k0}+1} X_{rj}\overline{Y}_j}{n(N_{k0}+1)}.$$

6. Дисперсия значимости коэффициента b_0 уравнения регрессии первого порядка для ЦПФЭ:

а) $S^2(b_0) = \dfrac{S^2_{\text{воспр}}}{nN_{k0}}$;   б) $S^2(b_0) = \dfrac{S^2_{\text{воспр}}}{n\sum\limits_{j=1}^{N_k} X_{0j}}$;

в) $S^2(b_0) = \dfrac{S^2_{\text{воспр}}}{n\sum\limits_{j=1}^{N_k} X^2_{0j}}$;   г) $S^2(b_0) = \dfrac{S^2_{\text{воспр}}}{n(N_{k0}+1)}$.

7. Дисперсии значимости коэффициентов b_r уравнения регрессии первого порядка для ЦПФЭ:

а) $S^2(b_r) = \dfrac{S^2_{\text{воспр}}}{N_{k0}}, \quad r = 1, ..., k$;

б) $S^2(b_r) = \dfrac{S^2_{\text{воспр}}}{nN_{k0}}, \quad r = 1, ..., k$;

в) $S^2(b_r) = \dfrac{S^2_{\text{воспр}}}{n\sum\limits_{j=1}^{N_k} X_{rj}}, \quad r = 1, ..., k$.

г) $S^2(b_r) = \dfrac{S^2_{\text{воспр}}}{n\sum\limits_{j=1}^{N_{k0}} X^2_{rj}}, \quad r = 1, ..., k$.

Вопросы для самопроверки

1. Напишите уравнения регрессии для четырех факторов ($k = 4$) следующих видов: первого порядка, неполного первого порядка, второго квадрата.
2. Какие требования предъявляются к факторам многофакторного эксперимента?
3. Рассчитать основной уровень и интервал варьирования факторов, если выбран диапазон исследований $x_r \in [x_{r\,\min}, x_{r\,\max}]$.

4. Напишите формулы перехода от натуральных значений факторов к нормированным.
5. Каким свойством должен обладать план многофакторного эксперимента, позволяющий определять коэффициенты регрессии независимо друг от друга? Проиллюстрируйте свое утверждение на примере плана ЦПФЭ для $k = 4$.
6. Сформулируйте алгоритм построения плана ЦПФЭ для любого количества факторов. Какое количество опытов для k факторов содержит план ЦПФЭ?
7. Составьте план для ЦПФЭ при $k = 2$.
8. Составьте план для ЦПФЭ при $k = 3$.
9. Составьте план для ЦПФЭ при $k = 4$.
10. Напишите формулы для расчета коэффициентов уравнения регрессии в ЦПФЭ для b_0, b_r.
11. Можно ли, выполнив эксперимент по плану ЦПФЭ, рассчитать коэффициенты регрессии b_{rr} при квадратичных членах?
12. Сформулируйте алгоритм проверки значимости коэффициентов уравнения регрессии, полученного по плану ЦПФЭ.
12. Сформулируйте алгоритм проверки на адекватность уравнения регрессии, полученного по плану ЦПФЭ.
13. Приведите возможные варианты дальнейших исследований, исходя из анализа уравнения регрессии первого порядка, полученного в результате ЦПФЭ.
14. Напишите уравнение для определения предельной абсолютной погрешности параметра Y, рассчитанного по многофакторному уравнению регрессии первого порядка.

Задачи для самостоятельного решения

Уровень 1

1. Исследуется зависимость выхода некоторого химического вещества в зависимости от температуры и давления в реакторе. Область варьирования температуры $x_1 \in [80,\ 140\ ^\circ\text{C}]$, давления $x_2 \in [0.1,\ 0.4\ \text{МПа}]$. Написать формулы перехода от натуральных значений факторов x_r к нормированным X_r ($r = 1, 2$). Найти основной уровень и интервалы варьирования для обоих факторов.

Ответ: $X_1 = (x_1 - 110)/30$, $x_{10} = 110\ ^\circ\text{C}$, $\Delta x_1 = 30\ ^\circ\text{C}$;
$X_2 = (x_2 - 0.25)/0.15$, $x_{20} = 0.25$ МПа, $\Delta x_2 = 0.15$ МПа.

Уровень 2

2. Выполнен ЦПФЭ, экспериментальные данные которого представлены в таблице. Построить уравнение регрессии первого порядка, проверить коэффициенты регрессии на значимость, а само уравнение – на адекватность.

N	X_0	X_1	X_2	Y_1	Y_2	Y_3	Y_4
1	1	–1	–1	20.3	19.9	19.3	18.9
2	1	1	–1	8.9	7.5	10.0	8.6
3	1	–1	1	29.6	31.7	32.2	30.2
4	1	1	1	21.9	22.5	21.2	21.3
5	1	0	0	19.7	21.2	20.9	20.7

Ответ: $Y = 20.33 - 5.01X_1 + 6.08X_2$. Все три коэффициента уравнения регрессии значимы, так как $\Delta b_0 = 0.41$, $\Delta b_1 = \Delta b_2 = 0.46$. Полученное уравнение адекватно, так как $F_э = 2.13 < F_{2,\,15,\,0.95} = 3.68$.

3. Выполнен ЦПФЭ, экспериментальные данные которого даны в таблице. Построить уравнение регрессии первого порядка, проверить коэффициенты регрессии на значимость, а уравнение – на адекватность.

N	X_0	X_1	X_2	Y_1	Y_2	Y_3	Y_4
1	1	–1	–1	2.7	2.3	1.7	1.3
2	1	1	–1	5.3	3.9	6.4	5.0
3	1	–1	1	–7.0	–4.9	–4.4	–6.4
4	1	1	1	–0.3	0.3	–1.0	–0.9
5	1	0	0	–0.3	1.2	0.9	0.7

Ответ: $Y = 2.09X_1 - 3.33X_2$. Линейные коэффициента уравнения регрессии значимы, так как $\Delta b_1 = \Delta b_2 = 0.46$. Свободный член уравнения $b_0 = 0.33$ – незначим, так как $\Delta b_0 = 0.41$. Полученное уравнение адекватно, так как $F_э = 2.08 < F_{3,\,15,\,0.95} = 3.29$.

4. Выполнен ЦПФЭ, экспериментальные данные которого даны в таблице. Построить уравнение регрессии первого порядка, проверить коэффициенты регрессии на значимость, а уравнение – на адекватность.

N	X_0	X_1	X_2	Y_1	Y_2	Y_3	Y_4
1	1	–1	–1	14.4	14.0	13.4	13.0
2	1	1	–1	14.8	13.4	15.9	14.5
3	1	–1	1	6.1	8.2	8.7	6.7
4	1	1	1	11.4	12.0	10.7	10.8
5	1	0	0	14.7	16.2	15.9	15.7

Ответ: $Y = 12.53 + 1.19X_1 - 2.43X_2$. Все три коэффициента уравнения регрессии значимы, так как $\Delta b_0 = 0.41$, $\Delta b_1 = \Delta b_2 = 0.46$. Полученное уравнение неадекватно, так как $F_э = 37.6 > F_{2,\,15,\,0.95} = 3.75$.

Уровень 3

5. Выполнен ЦПФЭ, экспериментальные данные которого даны в таблице. Построить уравнение регрессии первого порядка, проверить коэффициенты регрессии на значимость, а уравнение – на адекватность.

N	X_0	X_1	X_2	X_3	Y_1	Y_2	Y_3	Y_4
1	1	–1	–1	–1	17.1	18.6	15.7	14.9
2	1	1	–1	–1	25.8	27.3	26.5	24.3
3	1	–1	1	–1	4.4	4.6	6.3	4.2
4	1	1	1	–1	17.0	15.5	17.0	15.3
5	1	–1	–1	1	22.6	23.4	25.0	24.1
6	1	1	–1	1	33.8	33.3	33.3	33.3
7	1	–1	1	1	15.5	13.2	15.0	14.3
8	1	1	1	1	24.6	24.9	23.7	24.9
9	1	0	0	0	18.3	19.1	20.7	19.8

Ответ: $Y = 19.98 + 5.05X_1 - 4.96X_2 + 4.08X_3$. Все четыре коэффициента уравнения регрессии значимы, так как $\Delta b_0 = 0.31$, $\Delta b_1 = \Delta b_2 = \Delta b_3 = 0.37$. Полученное уравнение адекватно, так как $F_э = 2.03 < F_{5,\,27,\,0.95} = 2.57$.

6. Выполнен ПФЭ, экспериментальные данные которого даны в таблице. Построить уравнение регрессии первого порядка, проверить коэффициенты регрессии на значимость, а уравнение – на адекватность.

N	X_0	X_1	X_2	X_3	Y_1	Y_2	Y_3	Y_4
1	1	–1	–1	–1	13.3	14.8	11.9	11.1
2	1	1	–1	–1	12.8	14.3	13.5	11.3
3	1	–1	1	–1	–0.4	–0.2	1.5	–0.6
4	1	1	1	–1	3.8	2.3	3.8	2.1
5	1	–1	–1	1	18.8	19.6	21.2	20.3
6	1	1	–1	1	19.6	19.1	19.1	19.1
7	1	–1	1	1	11.9	9.6	11.4	10.7
8	1	1	1	1	11.4	11.7	10.5	11.7
9	1	0	0	0	18.3	19.1	20.7	19.8

Ответ: $Y = 12.19 - 4.96X_2 + 4.08X_3$. Коэффициенты b_0, b_2, b_3 значимы, так как $\Delta b_0 = 0.31$, $\Delta b_2 = \Delta b_3 = 0.37$. Линейный коэффициент $b_1 = 0.35$ – незначим, так как $\Delta b_1 = 0.37$. Полученное уравнение неадекватно, так как $F_э = 42.5 > F_{6,\,27,\,0.95} = 2.46$.

3.4. Центральный дробный факторный эксперимент (тема 8)

3.4.1. Теория

На начальном этапе исследований малоизученных объектов зачастую не хватает априорной информации о факторах, существенно влияющих на исследуемый параметр. В результате возникает соблазн включить в эксперимент как можно большее число факторов ($k \geq 4$) и оценить их влияние. Для этой цели можно использовать ПФЭ, однако количество экспериментов может оказаться чрезмерно большим. Например, даже по плану ПФЭ при $k = 4$ нужно провести $2^4 = 16$ опытов, а если учесть, например, четырехкратное дублирование, то получается $16 \times 4 = 64$ экспериментальные точки. Так как начальный этап исследований связан с построением уравнения регрессии первого порядка $Y = b_0X_0 + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4$ (5 коэффициентов), то оставшиеся степени свободы для проверки уравнения на адекватность $f_{ад} = 16 - 5 = 11$ являются в значительной степени избыточными, причем число «лишних» степеней свободы $f_{ад} = 2^k - k$ в ПФЭ стремительно возрастает с увеличением k.

«Лишние» степени свободы в ПФЭ можно использовать для оценки эффектов взаимодействия второго и более высокого порядка: X_rX_s, $X_rX_sX_v$, $X_rX_sX_vX_w$ и т. п. Например, план ПФЭ для построения 2-факторного уравнения регрессии неполного второго порядка $Y = b_0X_0 + b_1X_1 + b_2X_2 + b_{12}X_1X_2$ с единственно возможным эффектом взаимодействия, получаемый перемножением линейных факторов X_1X_2, представлен в таблице 17.

Таблица 17 – План ПФЭ для $k = 2$

N_{20}	X_{0j}	X_{1j}	X_{2j}	$X_{1j}X_{2j}$
1	+	–	–	+
2	+	+	–	–
3	+	–	+	–
4	+	+	+	+

План ПФЭ для построения 3-факторного уравнения регрессии неполного второго порядка $Y = b_0X_0 + b_1X_1 + b_2X_2 + b_3X_3 + b_{12}X_1X_2 + b_{13}X_1X_3 + b_{23}X_2X_3 + b_{123}X_1X_2X_3$ со всеми возможными эффектами взаимодействиями линейных факторов (X_1X_2, X_1X_3, X_2X_3) представлен в таблице 18.

Таблица 18 – План ПФЭ для $k = 3$

N_{30}	X_{0j}	X_{1j}	X_{2j}	X_{3j}	$X_{1j}X_{2j}$	$X_{1j}X_{2j}$	$X_{1j}X_{2j}$	$X_{1j}X_{2j}X_{3j}$
1	+	–	–	–	+	+	+	–
2	+	+	–	–	–	–	+	+
3	+	–	+	–	–	+	–	+
4	+	+	+	–	+	–	–	–
5	+	–	–	+	+	–	–	+
6	+	+	–	+	–	+	–	–
7	+	–	+	+	–	–	+	–
8	+	+	+	+	+	+	+	+

Легко показать, что все факторы планов ПФЭ, приведенных в таблицах 17 и 18, взаимно ортогональны (убедитесь в этом самостоятельно). Это означает, что система уравнений для расчета коэффициентов регрессии, полученная МНК, имеет диагональный вид, а сами коэффициенты независимо рассчитываются по следующим формулам:

$$b_{rs} = \frac{\sum_{j=1}^{N_k} (X_{rj} X_{sj}) \overline{Y}_j}{\sum_{j=1}^{N_k} (X_{rj} X_{sj})^2}, \quad \begin{matrix} r < s; \\ r = 1, ..., k-1; \;\; s = 2, ..., k; \end{matrix} \tag{161}$$

$$b_{rsv} = \frac{\sum_{j=1}^{N_k} (X_{rj} X_{sj} X_{vj}) \overline{Y}_j}{\sum_{j=1}^{N_k} (X_{rj} X_{sj} X_{vj})^2}, \quad \begin{matrix} r < s < v; \quad r = 1, ..., k-2; \\ s = 2, ..., k-1, v = 3, ..., k. \end{matrix} \tag{162}$$

Так как начальный этап изучения объектов связан с построением уравнения регрессии первого порядка, то «лишними» степенями свободы в плане ПФЭ типа 2^k можно распорядиться по-другому: вместо определения эффектов взаимодействия второго и более высокого порядка, определить коэффициенты при дополнительных линейных факторах.

Проиллюстрируем эту идею на примере плана ПФЭ типа 2^3. Вместо того, чтобы по этому плану искать регрессионные коэффициенты b_0, b_1, b_2, b_3, b_{12}, b_{13}, b_{23}, b_{123}, ограничимся поиском только свободного члена и линейных коэффициентов регрессии b_0, b_1, b_2, b_3, а вместо эффектов взаимодействия введем дополнительный линейный коэффициент b_4, причем уровни тройного взаимодействия $X_1 X_2 X_3$ примем за уровни дополнительного фактора X_4 (таблица 19):

$$X_4 = X_1 X_2 X_3 . \tag{163}$$

Равенство (163) называется генерирующим соотношением. Эксперимент, выполненный по такому плану, называют дробным факторным экспериментом (ДФЭ) и обозначают символом 2^{k-g}, где k – общее число линейных факторов, а g – число генерирующих соотношений.

Таблица 19 – План ДФЭ типа 2^{4-1}.

N_{30}	X_0	X_1	X_2	X_3	$X_4 = X_1X_2X_3$
1	+	–	–	–	–
2	+	+	–	–	+
3	+	–	+	–	+
4	+	+	+	–	–
5	+	–	–	+	+
6	+	+	–	+	–
7	+	–	+	+	–
8	+	+	+	+	+

Таким образом, для построения 4-факторного уравнения регрессии первого порядка $Y = b_0X_0 + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4$ вместо плана ПФЭ типа 2^4 (16 опытов) используется только его часть (дробная реплика). В общем случае можно использовать планы, в которых дополнительные линейные факторы приравнивают к различным эффектам взаимодействия. План ДФЭ типа 2^{k-g} строится на базе плана ПФЭ типа 2^{k_0}, где $k_0 = k - g$.

Анализируя данные таблицы 19, нетрудно убедиться, что новый линейный фактор X_4 ортогонален факторам X_0 X_1, X_2, X_3 (убедитесь в этом самостоятельно). Поэтому формулы для расчета регрессионных коэффициентов будут такие же, как и формулы для расчета коэффициентов уравнения регрессии, построенного по плану ПФЭ (см. уравнения (146), (147), заменив при этом N_{k0} на N_{kg}):

$$b_0 = \frac{\sum_{j=1}^{N_k} X_{0j} \overline{Y}_j}{\sum_{j=1}^{N_k} X_{0j}^2} = \frac{\sum_{j=1}^{N_k} X_{rj} \overline{Y}_j}{N_k} = \frac{\sum_{j=1}^{N_k} X_{rj} \overline{Y}_j}{N_{kg} + 1}; \tag{164}$$

$$b_r = \frac{\sum_{j=1}^{N_k} X_{rj}\overline{Y}_j}{\sum_{j=1}^{N_k} X_{rj}^2} = \frac{\sum_{j=1}^{N_k} X_{rj}\overline{Y}_j}{N_{kg}}, \; r = 1, ..., k. \tag{165}$$

Теория построения уравнений регрессии первого порядка с помощью планов ДФЭ, включающая расчет коэффициентов регрессии, проверку значимости этих коэффициентов, а так же самого уравнения регрессии на адекватность, ничем не отличается от аналогичной теории для ЦПФЭ (см. раздел 3.3.1), не забывая при этом, как уже было сказано выше, заменить символ N_{k0} на N_{kg}.

Следует отметить, что выбор генерирующих соотношений для плана ДФЭ на базе ПФЭ типа 2^{k_0} неоднозначен и в некотором роде произволен. Выбор того или иного генерирующего соотношения определяется априорной информацией, опытом и интуицией. Генерирующее соотношение $X_4 = -X_1X_2X_3$ может оказаться ничуть не хуже, чем соотношение $X_4 = +X_1X_2X_3$ (уравнение (163)). Кроме того, в качестве генерирующих соотношений можно было бы взять и другие: $X_4 = \pm X_1X_2$, $X_4 = \pm X_1X_3$, $X_4 = \pm X_2X_3$.

Возвратимся к рассмотренному выше примеру плана ДФЭ типа $2^{4\text{-}1}$. Из данных таблицы 19 видно, что значения столбцов X_4 и $X_1X_2X_3$ идентичны (впрочем, и без анализа содержимого таблицы 19 такой вывод очевиден и является следствием выбранного генерирующего соотношения $X_4 = X_1X_2X_3$ (уравнение (163)), и поэтому коэффициенты регрессии при них b_1 и b_{123} нельзя рассчитать раздельно, а можно определить только их сумму (принято говорить, что эти коэффициенты смешаны между собою):

$$b_4 + b_{123} = \frac{-\overline{Y}_1 + \overline{Y}_2 + \overline{Y}_3 - \overline{Y}_4 + \overline{Y}_5 - \overline{Y}_6 - \overline{Y}_7 + \overline{Y}_8}{8}. \tag{166}$$

Этот недостаток планов ДФЭ является своеобразной «платой» за уменьшение количества опытов по сравнению с планами ПФЭ. Можно предположить, что не только коэффициент b_4 смешан с b_{123}, но и другие коэффициенты b_0, b_1, b_2, b_3 в плане ДФЭ типа 2^{4-1} также смешаны с другими эффектами взаимодействия.

Опишем процедуру, позволяющую определить, какие факторы в ДФЭ смешаны между собой и каким образом. Но сначала сделаем

это не в общем виде, а на примере плана ДФЭ типа 2^{4-1} с генерирующим соотношением $X_4 = +X_1X_2X_3$.

Для преобразования уравнения (163) будем использовать следующие очевидные равенства:

$$X_r^2 = 1, \; r = 1, \ldots, k; \tag{167}$$

$$X_0X_r = X_r, \; r = 1, \ldots, k. \tag{168}$$

Умножая обе части генерирующего соотношения (163) на X_4, получим:

$$1 = X_1X_2X_3X_4. \tag{169}$$

Равенство (169) называют первичным определяющим контрастом. Умножая обе части первичного определяющего контраста (169) последовательно на факторы X_0, X_1, X_2, X_3, X_4 с учетом свойств (167), (168), оценим, какие из факторов тождественно совпадают.

$$X_0 = X_1X_2X_3X_4, \tag{170}$$

$$X_1 = X_2X_3X_4, \tag{171}$$

$$X_2 = X_1X_3X_4, \tag{172}$$

$$X_3 = X_1X_2X_4, \tag{173}$$

$$X_4 = X_1X_2X_3. \tag{174}$$

Заметим, что уравнение (174) уже встречалось в качестве генерирующего соотношения (уравнение (163)), остальные получены впервые.

Из уравнений (170)–(174) следует, что система оценок смешивания линейных коэффициентов уравнения регрессии имеет следующий вид:

$$b_0 \to \beta_0 + \beta_{1234}, \tag{175}$$

$$b_1 \to \beta_1 + \beta_{234}, \tag{176}$$

$$b_2 \to \beta_2 + \beta_{134}, \tag{177}$$

$$b_3 \to \beta_3 + \beta_{124}, \qquad (178)$$

$$b_4 \to \beta_4 + \beta_{123}. \qquad (179)$$

Из уравнений (176)–(179) следует, что свободный член и линейные коэффициенты смешаны с эффектами четвертого и третьего порядка и не смешаны с эффектами взаимодействия второго порядка. Этот факт позволяет надеяться на то, что план ДФЭ типа 2^{4-1}, состоящий из 8 опытов, позволит оценить свободный и линейные коэффициенты достаточно точно, так как практика показывает, что взаимодействия высших порядков, как правило, не оказывают статистически значимого влияния на исследуемый параметр.

В связи с этим при построении плана ДФЭ типа 2^{k-g} следует выбирать такие генерирующие соотношения, которые не приводили бы к смешиванию линейных эффектов со значимыми взаимодействиями, в первую очередь с бинарными. Поэтому при построении плана ДФЭ очень полезно исследовать возможности различных генерирующих соотношений. Однако еще раз отметим, что построение планов ДФЭ неформализуемо. В том случае, когда априорной информации нет, а интуиция «молчит», выбранное генерирующее соотношение может привести к смешиванию линейного эффекта со значимым взаимодействием. И тогда при интерпретации полученного уравнения регрессии может возникнуть искажение наших представлений о влиянии того или иного фактора на поведение исследуемого объекта. Поэтому при использовании ДФЭ, особенно большой дробности, например, типа 2^{5-2}, 2^{6-3}, 2^{6-2} и т. п., нужен тщательный анализ априорной информации, а иногда и постановка проверяющих внеплановых экспериментов.

Следует отметить, что целесообразно применять такие планы ДФЭ, в которых линейные эффекты смешиваются с взаимодействиями наивысшего порядка. Если на стадии составления плана ДФЭ, например, типа 2^{6-3}, по какой-то причине не удается избежать нежелательного смешивания значимых эффектов, то следует перейти к составлению плана ДФЭ меньшей дробности, например 2^{6-2}, то есть от плана с 8-ю опытами перейти к плану с 16-ю опытами.

Выше были рассмотрены планы ДФЭ типа 2^{k-1}, в которых было использовано только одно генерирующее соотношение. В планах ДФЭ типа 2^{k-2}, в которых используются 2 генерирующих соотношения, определение оценок смешивания строится следующим образом. Сначала на основе 2-х генерирующих соотношений строят

2 первичных определяющих контраста, а затем путем их перемножения получают еще один вторичный определяющий контраст. Для того чтобы охарактеризовать разрешающую способность плана ДФЭ типа 2^{k-2}, необходимо построить главный определяющий контраст, который состоит из 1 (единицы), двух первичных определяющих контрастов и одного вторичного контраста. В общем случае главный определяющий контраст состоит из 2^g элементов.

Проиллюстрируем построение системы оценок смешивания коэффициентов уравнения регрессии для планов ДФЭ типа 2^{k-2} на примере плана ДФЭ типа 2^{5-2}. В качестве генерирующих соотношений выберем $X_4 = -X_2X_3$ и $X_5 = X_1X_2X_3$ ($g = 2$). На их основе построим два первичных определяющих контраста: $1 = -X_2X_3X_4$ и $1 = X_1X_2X_3X_5$. Путем перемножения первичных определяющих контрастов с учетом свойства, математически выраженного уравнением (167), получим еще один вторичный определяющий контраст: $1 = -X_1X_4X_5$. Главный определяющий контраст, состоящий из 1 (единицы), двух первичных определяющих контрастов и одного вторичного контраста, в этом случае будет выглядеть следующим образом (содержит $2^g = 2^2 = 4$ элементов):

$$1 = -X_2X_3X_4 = X_1X_2X_3X_5 = -X_1X_4X_5. \qquad (180)$$

Для планов ДФЭ типа 2^{k-3}, в которых используется 3 генерирующих соотношения, определение оценок смешивания строится следующим образом. Сначала на основе 3-х генерирующих соотношений строят 3 первичных определяющих контраста A, B, C. Путем попарного перемножения 3-х первичных контрастов находят еще 3 вторичных определяющих контраста AB, AC, BC, а затем путем тройного перемножения 3-х первичных контрастов находят еще один вторичный определяющий контраст ABC. Главный определяющий контраст в этом случае вместе с 1 (единицей), 3-мя первичными определяющими контрастами и 4-мя вторичными определяющими контрастами состоит из 8-ми элементов ($2^g = 2^3 = 8$):

$$1 = A = B = C = AB = AC = BC = ABC. \qquad (181)$$

Для придания планам ДФЭ большей мощности его следует дополнить еще одним опытом в центре плана и по аналогии с ЦПФЭ такой план следует называть ЦДФЭ. Аргументы в пользу преимущества планов ЦДФЭ по сравнению с планами ДФЭ полностью совпадают с аргументами при переходе от планов ПФЭ к планам ЦПФЭ (см. раздел 3.3.1). Количество опытов плана ЦДФЭ рассчитывается по формуле $N_k = N_{kg} + 1 = 2^{k-g} + 1$.

При построении планов ЦДФЭ следует учитывать, что для выбранного числа факторов k существует максимальное число генерирующих соотношений g. Рассчитать максимальное значение g можно из очевидного утверждения: число опытов в ЦДФЭ должно быть больше числа коэффициентов уравнения регрессии. Для уравнения регрессии первого порядка (135) это утверждение записывается следующим образом:

$$N_k = 2^{k-g} + 1 > 1 + k, \quad (182)$$

или, как показано в приложении 6:

$$g < k - \frac{\lg k}{\lg 2}. \quad (183)$$

Для уравнения регрессии неполного второго порядка (136) это утверждение запишется следующим образом:

$$N_k = 2^{k-g} + 1 > 1 + k + C_k^2 = 1 + k + \frac{k(k-1)}{1 \cdot 2} = \frac{k^2 + k + 2}{2} \quad (184)$$

или, как показано в приложении 6:

$$g < 1 + k - \frac{\lg\left[k(k+1)\right]}{\lg 2}. \quad (185)$$

В таблице 20 приведены значения генерирующих соотношений g, которые рассчитаны по уравнениям (183), (185) для различного числа факторов, встречающихся на практике $k = 2, \ldots, 10$.

В заключение этого раздела приведем следующее замечание. Не располагая *априори* информацией о существенном значении какого-либо эффекта взаимодействия второго порядка, можно ошибочно считать его незначимым, использовать для построения генерирующего соотношения и тем самым исказить влияние дополнительно введенного фактора на исследуемый параметр. Обнаружить

это бывает достаточно трудно даже *апостериори*. Проведение некоторого дополнительного количества опытов, случайным образом выбранных внутри факторного пространства, может дать необходимую информацию для правильного принятия решения. Но без особой необходимости увлекаться планами ЦДФЭ большой дробности не следует.

Таблица 20 – Значения чисел g для планов ЦДФЭ

k	$Y = b_0 + \sum_{r=1}^{k} b_r X_r$	$Y = b_0 + \sum_{r=1}^{k} b_r X_r + \sum_{r<s}^{k} b_{rs} X_r X_s$
	g	
2	0	0
3	0, 1	0
4	0, 1	0
5	0, 1, 2	0, 1
6	0, 1, 2, 3	0, 1
7	0, 1, 2, 3, 4	0, 1, 2
8	0, 1, 2, 3, 4	0, 1, 2
9	0, 1, 2, 3, 4, 5	0, 1, 2, 3
10	0, 1, 2, 3, 4, 5, 6	0, 1, 2, 3, 4

3.4.2. Алгоритм решения прикладных задач

1. Выбрать объект исследования и на основе литературных и патентных источников собрать необходимую априорную информацию о факторах и параметрах, характеризующих исследуемый объект.
2. Выбрать параметр Y, который наиболее полно характеризует исследуемый объект, а также число дублей в каждом опыте. Рекомендуем число дублей брать не менее 3.
3. Выбрать k варьируемых факторов x_r, которые существенным образом влияют на изменение параметра, а также диапазон их варьирования $[x_{r\,\min}, x_{r\,\max}]$. С учетом того, что в планах ЦПФЭ и ЦДФЭ каждый фактор варьируется только на 3-х уровнях, натуральные значения факторов будут равняться максимальному, основному и минимальному значениям, а нормированные ± 1 и 0.
4. Из материально-технических соображений выбрать приемлемое число опытов базового плана ПФЭ типа 2^{k_0}.

5. Для выбранного числа варьируемых факторов k и базового значения k_0 проверить выполнение для генерирующих соотношений g ограничительных неравенств, где $k_0 = k - g$:

Для уравнения регрессии первого порядка на базе ЦДФЭ:

$$g < k - \frac{\lg k}{\lg 2}.$$

Для уравнения регрессии неполного второго порядка на базе ЦДФЭ:

$$g < 1 + k - \frac{\lg\left[k(k+1)\right]}{\lg 2}.$$

Если указанные неравенства для $g = k - k_0$ не выполняются, то следует уменьшить число генерирующих соотношений g плана ЦДФЭ, либо за счет уменьшения числа варьируемых факторов k либо за счет увеличения числа факторов k_0 базового плана ЦПФЭ типа 2^{k_0}.

6. Построить план ПФЭ типа 2^{k_0}, в который включить все факторы: X_0, X_r ($r = 1, \ldots, k$), $X_r X_s$, ($r < s$, $r = 1, \ldots, k-1$, $s = 2, \ldots, k$), $X_r X_s X_v$ ($r < s <$ v, $r = 1, \ldots, k-2$, $s = 2, \ldots, k-1$, $v = 3, \ldots, k$) и т. д.

7. Задать g генерирующих соотношений, для чего дополнительные линейные факторы X_r ($r = k_0 + 1, \ldots, k$) приравнять к эффектам взаимодействия наивысшего порядка в плане ПФЭ типа 2^{k_0}. Выбор эффектов высшего порядка для генерирующих соотношений осуществляется на основе априорной информации или интуиции. Построенный план ПФЭ типа 2^{k_0} после замены выбранных эффектов взаимодействия высшего порядка на g дополнительных линейных факторов становится планом ДФЭ типа 2^{k-g}. Для построения плана ЦДФЭ типа $2^{k-g} + 1$ дополнить план ДФЭ типа 2^{k-g} одним опытом в центре выбранного факторного пространства (0, …, 0).

8. На базе g генерирующих соотношений построить g первичных определяющих контрастов путем умножения обеих частей генерирующих соотношений на X_r ($r = k_0 + 1, \ldots, k$), не забывая при этом, что $X_r^2 = 1$.

9. На базе g первичных определяющих контрастов построить $C_g^2 = \dfrac{g(g-1)}{1 \cdot 2}$ вторичных определяющих контрастов путем попарного перемножения первичных контрастов, не забывая при этом, что $X_r^2 = 1$, где $r = 1, \ldots, k$.

10. На базе g первичных определяющих контрастов построить $C_g^3 = \dfrac{g(g-1)(g-2)}{1 \cdot 2 \cdot 3}$ вторичных определяющих контрастов путем тройного перемножения первичных контрастов, не забывая при этом, что $X_r^2 = 1$, где $r = 1, \ldots, k$ и т. п. до построения последнего вторичного определяющего контраста, получаемого перемножением всех первичных контрастов.

11. Чтобы охарактеризовать разрешающую способность выбранной реплики ДФЭ типа 2^{k-g}, следует построить главный определяющий контраст, который состоит из 1 (единицы), g первичных определяющих контрастов и $2^g - g - 1$ вторичных определяющих контрастов, всего 2^g элементов. Таким образом, для $g = 1$ главный определяющий контраст состоит из 2-х элементов, для $g = 2$ из 4-х элементов, для $g = 3$ из 8-ми элементов и т. д.

12. Построить систему оценок смешивания свободного члена и линейных коэффициентов уравнения регрессии путем умножения главного определяющего контраста на X_0, X_r ($r = 1, \ldots, k$), не забывая при этом, что $X_0 X_r = X_r$ и $X_r^2 = 1$. При построении системы оценок смешивания ограничимся только свободным и линейными коэффициентами, так как стратегия исследований на первом этапе связана с построением уравнения регрессии первого порядка.

13. После построения плана ДФЭ выполнить эксперимент по плану ЦДФЭ с числом опытов $N_k = 2^{k-g} + 1$, провести предварительный анализ экспериментальных данных, рассчитать коэффициенты уравнения регрессии первого порядка, проверить их на значимость, а само уравнение – на адекватность. Математический аппарат этого пункта подробно изложен в алгоритме 3.3.2 с учетом необходимой замены $N_{k0} \to N_{kg}$.

3.4.3. Базовая задача

Определить, с какими бинарными эффектами взаимодействиями второго порядка смешаны линейные эффекты в 7-ми факторном уравнении регрессии при использовании плана ДФЭ типа 2^{7-4} для $g = 4$ (ссылки на пункты алгоритма 3.4.2):

$$Y = b_0 X_0 + b_1 X_1 + b_2 X_2 + b_3 X_3 + b_4 X_4 + b_5 X_5 + b_6 X_6 + b_7 X_7 . \qquad (186)$$

Очевидно, что план ДФЭ типа 2^{7-4} строится на базе плана ПФЭ типа 2^3 ($k_0 = 3$), так как $7 - 4 = 3$ (пп. 4–6). Число генерирующих соотношений $g = 4$ допустимо, поскольку $4 < 7 - \lg 7/\lg 2 = 4.19$. При моделировании объекта 7-ми факторным уравнением регрессии первого порядка выберем следующие четыре генерирующих соотношения (п. 7):

$$X_4 = X_1 X_2 , \qquad (187)$$

$$X_5 = X_1 X_3 , \qquad (188)$$

$$X_6 = X_2 X_3 , \qquad (189)$$

$$X_7 = X_1 X_2 X_3 . \qquad (190)$$

Вопрос о целесообразности применения плана с такой, относительно большой степенью дробности, рассматриваться не будет, так как целью базовой задачи является иллюстрация алгоритма построения системы оценок эффектов смешивания линейных коэффициентов регрессии.

На базе 4-х выбранных генерирующих соотношений (187)–(190) построим 4 первичных определяющих контраста (п. 8):

$$1 = X_1 X_2 X_4 , \qquad (191)$$

$$1 = X_1 X_3 X_5 , \qquad (192)$$

$$1 = X_2 X_3 X_6 , \qquad (193)$$

$$1 = X_1 X_2 X_3 X_7 . \qquad (194)$$

Следующий этап – построение определяющих контрастов на основе 4-х первичных путем попарного перемножения их с учетом свойства $X_r^2 = 1$ (уравнение 167). Их количество будет равно:

$$C_4^2 = \frac{4 \cdot 3}{1 \cdot 2} = 6 , \qquad (195)$$

$$1 = (X_1 X_2 X_4)(X_1 X_3 X_5) = X_2 X_3 X_4 X_5 , \qquad (196)$$

$$1 = (X_1 X_2 X_4)(X_2 X_3 X_6) = X_1 X_3 X_4 X_6 , \qquad (197)$$

$$1 = (X_1 X_2 X_4)(X_1 X_2 X_3 X_7) = X_3 X_4 X_7 , \qquad (198)$$

$$1 = (X_1 X_3 X_5)(X_2 X_3 X_6) = X_1 X_2 X_5 X_6 , \qquad (199)$$

$$1 = (X_1 X_3 X_5)(X_1 X_2 X_3 X_7) = X_2 X_5 X_7 , \qquad (200)$$

$$1 = (X_2 X_3 X_6)(X_1 X_2 X_3 X_7) = X_1 X_6 X_7 . \qquad (201)$$

Очередной этап – построение определяющих контрастов на основе 4-х первичных путем тройного перемножения их с учетом свойства $X_r^2 = 1$ (уравнение 167). Их количество будет равно:

$$C_4^3 = \frac{4 \cdot 3 \cdot 2}{1 \cdot 2 \cdot 3} = 4 , \qquad (202)$$

$$1 = (X_1 X_2 X_4)(X_1 X_3 X_5)\left(X_2 X_3 X_6\right) = X_4 X_5 X_6 , \qquad (203)$$

$$1 = (X_1 X_2 X_4)\left(X_1 X_3 X_5\right)(X_1 X_2 X_3 X_7) = X_1 X_4 X_5 X_7 , \qquad (204)$$

$$1 = (X_1 X_2 X_4)\left(X_2 X_3 X_6\right)(X_1 X_2 X_3 X_7) = X_2 X_4 X_6 X_7 , \qquad (205)$$

$$1 = (X_1 X_3 X_5)\left(X_2 X_3 X_6\right)(X_1 X_2 X_3 X_7) = X_3 X_5 X_6 X_7 , \qquad (206)$$

И, наконец, еще один определяющий контраст получается путем четверного перемножения всех первичных:

$$1 = \left(X_1 X_2 X_4\right)(X_1 X_3 X_5)\left(X_2 X_3 X_6\right)(X_1 X_2 X_3 X_7) =$$

$$= X_1 X_2 X_3 X_4 X_5 X_6 X_7 . \qquad (207)$$

Главный определяющий будет состоять из 16-ти структурных единиц (2^g) – из 1 (единица) и 15-ти определяющих контрастов:

$$1 = X_1X_2X_4 = X_1X_3X_5 = X_2X_3X_6 = X_1X_2X_3X_7 =$$

$$= X_2X_3X_4X_5 = X_1X_3X_4X_6 = X_3X_4X_7 = X_1X_2X_5X_6 = X_2X_5X_7 =$$

$$= X_1X_6X_7 = X_4X_5X_6 = X_1X_4X_5X_7 = X_2X_4X_6X_7 =$$

$$X_3X_5X_6X_7 = X_1X_2X_3X_4X_5X_6X_7 . \qquad (208)$$

Умножая главный определяющий контраст поочередно на X_1, X_2, X_3, X_4, X_5, X_6, X_7 и оставляя только линейные и бинарные коэффициенты, получаем:

$$b_1 \to \beta_1 + \beta_{24} + \beta_{35} + \beta_{67} , \qquad (209)$$

$$b_2 \to \beta_2 + \beta_{14} + \beta_{36} + \beta_{57} , \qquad (210)$$

$$b_3 \to \beta_3 + \beta_{15} + \beta_{26} + \beta_{47} , \qquad (211)$$

$$b_4 \to \beta_4 + \beta_{12} + \beta_{37} + \beta_{56} , \qquad (212)$$

$$b_5 \to \beta_5 + \beta_{13} + \beta_{27} + \beta_{46} , \qquad (213)$$

$$b_6 \to \beta_6 + \beta_{23} + \beta_{17} + \beta_{45} , \qquad (214)$$

$$b_7 \to \beta_7 + \beta_{34} + \beta_{25} + \beta_{16} . \qquad (215)$$

Из уравнения (208) следует, что свободный член b_0 смешан только с тройными и более высокого порядка взаимодействиями $b_0 \to \beta_0$. Если гипотеза о незначимости эффектов взаимодействий второго и более высоких порядков подтвердится, то экономия материальных, энергетических и временных ресурсов будет очевидна: 8 опытов плана ДФЭ типа 2^{7-4} вместо 128 опытов плана ПФЭ типа 2^7 (!).

После проведения эксперимента по плану ЦДФЭ типа $2^{7-4} + 1$ необходимо провести предварительную обработку экспериментальных данных, построить уравнение регрессии первого порядка и проверить его на качество методами математической статистики (см. алгоритм 3.3.2).

3.4.4. Материалы для самостоятельной работы, лабораторных и практических занятий

Тесты

1. План эксперимента ДФЭ 2^{4-1} при генерирующем соотношении $X_4 = -X_1X_2X_3$.

а)

N_{k0}	X_0	X_1	X_2	X_3	X_4
1	+	–	–	–	–
2	+	+	–	–	+
3	+	–	+	–	+
4	+	+	+	–	–
5	+	–	–	+	+
6	+	+	–	+	–
7	+	–	+	+	–
8	+	+	+	+	+

б)

N_{k0}	X_0	X_1	X_2	X_3	X_4
1	+	–	–	–	+
2	+	+	–	–	–
3	+	–	+	–	–
4	+	+	+	–	+
5	+	–	–	+	–
6	+	+	–	+	+
7	+	–	+	+	+
8	+	+	+	+	–

в)

N_{k0}	X_0	X_1	X_2	X_3	X_4
1	+	–	–	–	–
2	+	+	–	–	+
3	+	–	+	–	+
4	+	+	+	–	–
5	+	–	–	+	–
6	+	+	–	+	+
7	+	–	+	+	+
8	+	+	+	+	–

2. Найти определяющие контрасты для ДФЭ 2^{5-2} при построении уравнения регрессии $Y = b_0X_0 + b_1X_1 + b_2X_2 + b_3X_4 + b_4X_4 + b_5X_5$, если в качестве регенерирующих соотношений выбраны следующие выражения: $X_4 = -X_1X_2$ и $X_5 = X_1X_2X_3$:

а) $1 = -X_1X_2X_3 = X_1X_2X_3X_5 = -X_3X_4X_5$;

б) $1 = -X_1X_2X_4 = X_1X_2X_3X_4 = -X_3X_4X_5$;

в) $1 = -X_1X_2X_4 = X_1X_2X_3X_5 = -X_3X_4X_5$;

г) $1 = -X_1X_2X_4 = X_1X_2X_3X_5 = X_3X_4X_5$.

Вопросы для самопроверки

1. Что такое дробная реплика 2^{k-g}? Что обозначают символом g?
2. Что такое генерирующее соотношение?
3. Что такое определяющий контраст?
4. Что такое главный определяющий контраст? Как он строится?
5. Если генерирующее соотношение одно, то сколько определяющих контрастов?
6. Если генерирующих соотношений два, то сколько определяющих контрастов?
7. Если генерирующих соотношений три, то сколько определяющих контрастов?
8. Проводится эксперимент ДФЭ типа 2^{5-2}. Выбраны генерирующие соотношения $X_4 = -X_1X_2$ и $X_5 = -X_1X_3$. Напишите систему оценок смешивания для свободного и линейных коэффициентов уравнения регрессии $Y = b_0X_0 + b_1X_1 + b_2X_2 + b_3X_4 + b_4X_4 + b_5X_5$.
9. Напишите формулы для расчета свободного и линейных коэффициентов уравнения регрессии в ДФЭ для k факторов.
10. Напишите формулы для расчета дисперсий значимости свободного и линейных коэффициентов уравнения регрессии в ДФЭ для k факторов.
11. Сформулируйте алгоритм проверки на адекватность уравнения регрессии, полученного в результате проведения опытов по плану ДФЭ типа 2^{k-g}.
12. Что такое ЦДФЭ и как строится его план?
13. Приведите варианты дальнейших исследований уравнения регрессии первого порядка, полученного в результате ЦДФЭ.

Задачи для самостоятельного решения

Уровень 1

1. Сколько опытов в дробной реплике ДФЭ типа 2^{8-4}?
Ответ: $N_{kg} = 16$.

Уровень 2

2. Принято решение провести эксперимент по дробной реплике типа 2^{6-2}, в качестве генерирующих соотношений взяли следующие: $X_5 = X_1X_2X_3$; $X_6 = X_1X_2X_4$. Как будут смешаны между собой все парные взаимодействия?

Ответ: всего парных взаимодействий 15, так как $C_6^2 = \dfrac{6\cdot 5}{1\cdot 2} = 15$,

и между собою они смешаны следующим образом:

$$b_{12} \to \beta_{12} + \beta_{35} + \beta_{46},$$
$$b_{13} \to \beta_{13} + \beta_{25},$$
$$b_{14} \to \beta_{14} + \beta_{26},$$
$$b_{15} \to \beta_{15} + \beta_{23},$$
$$b_{16} \to \beta_{16} + \beta_{24},$$
$$b_{34} \to \beta_{34} + \beta_{56},$$
$$b_{36} \to \beta_{36} + \beta_{45}.$$

3. Принято решение провести эксперимент по дробной реплике типа 2^{6-3}, в качестве генерирующих соотношений взяли следующие: $X_4 = X_1X_2$; $X_5 = X_1X_3$; $X_6 = X_2X_3$. С какими парными взаимодействиями будут смешаны линейные коэффициенты?

Ответ:

$$b_1 \to \beta_1 + \beta_{24} + \beta_{35},$$
$$b_2 \to \beta_2 + \beta_{14} + \beta_{36},$$
$$b_3 \to \beta_3 + \beta_{15} + \beta_{26},$$
$$b_4 \to \beta_4 + \beta_{12} + \beta_{56},$$
$$b_5 \to \beta_5 + \beta_{13} + \beta_{46},$$
$$b_6 \to \beta_6 + \beta_{23} + \beta_{45}.$$

Уровень 3

4. Сколько регенерирующих соотношений и определяющих контрастов в плане ДФЭ типа 2^{9-5}?
Ответ: 5 и 31, соответственно.

3.5. Крутое восхождение (тема 9)

3.5.1. Теория

Часто исследуемый объект описывается нелинейными моделями, имеющими максимум или минимум. Это относится и к задачам откорма скота, и к задачам получения максимальной урожайности, и к задачам оптимальной настройки сельскохозяйственной техники. Математически это означает, что в некоторой области факторного пространства изучаемый объект следует моделировать регрессионным уравнением второго порядка (см. уравнение (137)). Область, в которой изучаемый объект описывается нелинейным уравнением регрессии, будем называть стационарной. Как попасть в стационарную область?

Напомним, что изучение исследуемого объекта всегда следует начинать с простейшей модели – уравнения регрессии первого порядка (см. уравнение (135)). Если уравнение регрессии первого порядка окажется адекватным, то в теории планирования эксперимента известен прием, который позволяет за относительно небольшое количество опытов достичь стационарной области. Этот прием называется методом крутого восхождения (спуска). Суть его заключается в проведении дополнительных опытов, которые позволяют реализовать движение к оптимуму по кратчайшему пути – по градиенту.

Поясним смысл движения по градиенту. Представьте себе холм, вершина которого находится на высоте 100 м над уровнем океана. Мысленно проведем через холм секущую плоскость параллельно поверхности океана на высоте 90 м над уровнем океана. Пересечение поверхности склона холма с этой плоскостью образует линию, повторяющую очертание склона на этой высоте. Эту линию называют линией постоянного уровня, в данном случае на высоте 90 м. Повторим мысленное рассечение холма плоскостями на высоте 80, 70, 60, 50, 40 30, 20, 10 и 0 м над уровнем океана. Получим топографическую карту холма, на которой линии постоянного уровня позволяют нам представить объемную геометрию холма (рисунок 3). Например, если в каком-то месте линии постоянного уровня располагаются очень близко друг к другу («густо»), то в этом месте в реальном холме наблюдается большая крутизна склона. И наоборот, если расстояние между линиями постоянного уровня большое («редко»), то в этом месте холм имеет пологий склон. Очевидно,

что если мы оказались в некоторой точке холма, например в точке A_0 (рисунок 3), и хотим кратчайшим путем взобраться на вершину, то нужно идти по направлению, перпендикулярному линиям постоянного уровня. Это и есть движение, которое в математике называют движением по градиенту (крутое восхождение).

На рисунке 3 представлены линии постоянного уровня некоторой двухфакторной функции $Y = Y(X_1, X_2)$. Градиент непрерывной функции, как следует из математического анализа, есть вектор

$$\overline{\text{grad } Y} = \frac{\partial Y}{\partial X_1}\vec{i} + \frac{\partial Y}{\partial X_2}\vec{j}\,, \qquad (216)$$

где $\dfrac{\partial Y}{\partial X_1}, \dfrac{\partial Y}{\partial X_2}$ – частные производные функции Y по факторам X_1 и X_2; $\vec{i}, \vec{j}$ – единичные векторы в направлении координатных осей факторного пространства OX_1 и OX_2.

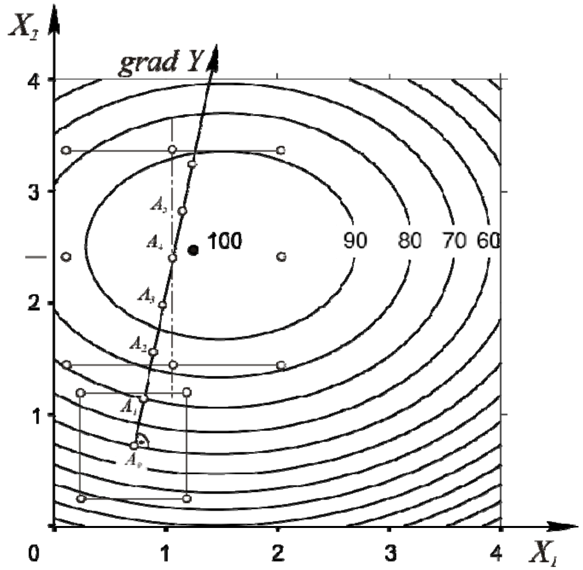

Рисунок 3 – Схема движения в стационарную область по градиенту.

Градиент линейной функции $Y = b_0 X_0 + b_1 X_1 + \ldots + b_k X_k$, как следует из уравнения (216), есть вектор, который представим в следующем виде:

$$\overrightarrow{\text{grad }Y} = \begin{Vmatrix} b_1 \\ \vdots \\ b_r \\ \vdots \\ b_k \end{Vmatrix}. \qquad (217)$$

Параметрическое уравнение прямой линии, проходящей через начало координат факторного пространства (0, 0, …, 0) параллельно градиенту функции, имеет следующий вид:

$$\begin{cases} X_1 = \xi b_1 \\ \quad\vdots \\ X_r = \xi b_r\,, \quad r = 1, ..., k\,, \\ \quad\vdots \\ X_k = \xi b_k \end{cases} \qquad (218)$$

где ξ – числовой параметр, который пробегает значения от 0 до $+\infty$ при крутом восхождении, и от 0 до $-\infty$ при крутом спуске.

Если уравнение регрессии первого порядка адекватно, то можно рассчитать направление крутого восхождения (спуска) и приступить к движению по градиенту. Начинать движение следует из центра факторного пространства A_0, в котором оно адекватно. Напомним, что нормированные значения факторов изменяются в интервале $1 \le X_r \le +1$. На рисунке 3 прямой соответствующей grad$Y$ обозначено направление крутого восхождения из точки A_0. Следующую точку на прямой крутого восхождении A_1 желательно выбрать так, чтобы она лежала не на границе исследованного факторного пространства, а внутри его, близко к границе, а последующая точка A_2 выходила бы за его пределы. Сравнение экспериментального значения параметра Y в точке A_1 с рассчитанным по уравнению регрессии даст дополнительную информацию об адекватности используемого линейного уравнения регрессии. Следует отметить, что выбор шага при крутом восхождении (спуске) является неформализуемой операцией. Если шаг будет небольшим, то может потребоваться относительно большое количество опытов для достижения стационарной области. Если шаг будет большим, то можно проскочить максимум (минимум) исследуемого объекта.

Так как уравнение прямой, по которому осуществляется крутое восхождение (спуск) параметрическое, то нужно задать величину шага в изменении числового параметра ξ и рассчитать координаты точек, по которым будет совершаться крутое восхождение (спуск) сначала в нормированных, а потом в натуральных значениях факторов.

Для этого сначала находят s-й фактор, для которого линейный коэффициент регрессии b_s максимальный по абсолютной величине, и обозначают его буквой b:

$$b = \max\left|b_s\right|. \qquad (219)$$

Затем выбирают значение параметра $\xi = \xi_1$ для первого шага из условия, чтобы 1-я точка при крутом восхождении (спуске) лежала бы внутри исследованного факторного пространства, то есть

$$\xi_1 b = \mu < 1 \text{ или } \xi_1 = \mu / b\,, \qquad (220)$$

где $0 \le \mu \le 1$ (при крутом восхождении), $-1 \le \mu \le 0$ (при крутом спуске).

Если принять $\mu = 0$, то из уравнения (220) получим $\xi_1 = 0$, а из уравнения (218) следует, что все $X_r = 0$ (точки A_0 в нормированных значениях факторов). Если принять $\mu = +1$, то из уравнения (220) получим $\xi_1 = 1/b$, а из уравнения (218) следует, что для s-го фактора в $X_s = 1$, и все остальные значения факторов X_r ($r \ne s$) по модулю будут меньше единицы, так как после подстановки уравнения (220) в (218) имеем $X_r = \left|b_r\right| / b < 1$. Поэтому точка A_1 лежит на границе выбранного факторного пространства. При осуществлении крутого восхождения (спуска) желательно, как было сказано выше, чтобы точка A_1 все-таки лежала внутри факторного пространства, и поэтому μ целесообразно брать в диапазоне $0.7 \le \mu \le 0.9$ при крутом восхождении ($-0.9 \le \mu\ -0.7$ при крутом спуске).

После выбора первого шага рассчитывают координаты последующих точек крутого восхождения (спуска) с шагом h:

$$X_r^h = h \xi_1 b_r\,, \qquad r = 1, ..., k \qquad (221)$$

или с учетом уравнения (218):

$$X_r^h = h \mu \frac{b_r}{b}\,,\ r = 1, ..., k\,, \qquad (222)$$

где $h = 1, 2, 3, 4, \ldots$ – номер шага в направлении крутого восхождения $0.7 \le \mu \le 0.9$ (спуске $-0.9 \le \mu \le -0.7$).

Для проведения опытов по крутому восхождению (спуску) необходимо перейти от нормированных к натуральным значениям факторов. Подставляя в уравнение (131) уравнение (222), получим:

$$x_r^h = x_{ro} + h\mu \frac{b_r}{b} \Delta x_r, \; r = 1, \ldots, k. \qquad (223)$$

После расчета условий опытов по крутому восхождению (спуску) осуществляют эксперимент по этому плану. Опыт, в котором параметр будет иметь наибольшее (наименьшее) значение, выбрать за новый центр факторного пространства для дальнейших исследований.

В новом факторном пространстве проводят исследования по плану ЦПФЭ (ЦДФЭ) с целью построения нового уравнения регрессии первого порядка. Если оно окажется адекватным, то процедуру крутого восхождения повторяют. Если же полученное уравнение регрессии первого порядка окажется неадекватным, это означает, что достигнута стационарная область и переходят к построению уравнения регрессии второго порядка с соответствующим планированием эксперимента (см. разделы 3.6.1 и 3.6.2).

На рисунке 3 наибольшее значение параметра Y при крутом восхождении наблюдается в точке A_4 (она ближе всех точек, лежащих на прямой крутого восхождения, расположена к максимуму параметра Y, равного 100). В данном случае из рисунка 3 следует, что достигнута стационарная область, и точку A_4 можно принять за новый центр факторного пространства.

Следует отметить, что при крутом восхождении (спуске) в каждом шаге можно взять минимальное число дублей $n = 2$, так как ошибка в этом случае может только несколько исказить положение центра нового факторного пространства. Но поскольку новое факторное пространство имеет значительные размеры, построенное на нем уравнение регрессии второго порядка «накроет» максимум (минимум) и позволит найти его точное значение.

3.5.2. Алгоритм решения прикладных задач

1. Если многофакторное уравнение регрессии первого порядка, полученное по алгоритму 3.3.2 или 3.4.2 адекватно, перейти к процедуре крутого восхождения.

2. Найти s-й фактор, для которого коэффициент регрессии b_s будет максимальным по абсолютной величине, и обозначить $b = |b_s|$.

3. Выбрать значение параметра $\xi_1 = \mu / b$. Для первого шага крутого восхождения рекомендуем взять $0.7 \le \mu \le 0.9$; для крутого спуска $-0.9 \le \mu \le -0.7$.

4. Рассчитать координаты опытов крутого восхождения (спуска) в натуральных значениях факторов:

$$x_r^h = x_{ro} + h\, \mu \Delta x_r \frac{b_r}{b}, \; r = 1, \ldots, k,$$

где $h = 1, 2, 3, 4, \ldots$ – номер шага в направлении крутого восхождения при $0.7 \le \mu \le 0.9$ или спуска при $-0.9 \le \mu \le -0.7$.

5. Число дублей в каждом опыте принять $n = 2$.

6. Из реализованных опытов выбрать тот, для которого параметр Y будет иметь наибольшее (наименьшее) значение. Натуральные значения факторов наилучшего опыта принять за основной уровень нового плана ЦПФЭ или ЦДФЭ. Повторить процедуру построения и анализа уравнения регрессии первого порядка по алгоритму 3.3.2. Если полученное уравнение регрессии первого порядка окажется адекватным повторить процедуру крутого восхождения, если неадекватным – перейти к построению уравнения регрессии второго порядка (см. разделы 3.6.1 и 3.6.2).

3.5.3. Базовая задача

Продолжим решение базовой задачи 3.3.3. Так как в полученном уравнении $Y = 28.45 + 6.77X_1 + 5.62X_2$ все коэффициенты регрессии значимы и оно адекватно, реализуем процедуру крутого восхождения (п. 1). Ссылки **на алгоритм 3.5.2.**

Так как для фактора X_1 коэффициент регрессии максимальный по абсолютной величине (п. 2), то рассчитаем значение параметра $\xi = \xi_1$ для первого шага при $\mu = 0.75$, то есть $\xi_1 = \mu / |b_1| = 0.75/6.77 = 0.111$ (п. 3).

Рассчитаем координаты опытов крутого восхождения (основные уровни и интервалы варьирования факторов взяты из таблицы 15) (п. 4):

$$x_1^h = x_{10} + h\xi_1 (b_1 \Delta x_1) = 1 + h \cdot 0.111 \cdot (6.77 \cdot 0.5) = 1 + 0.375h,$$

$$x_2^h = x_{20} + h\xi_1 (b_2 \Delta x_2) = 0.5 + h \cdot 0.111 \cdot (5.62 \cdot 0.25) = 0.5 + 0.156h,$$

где $h = 1, 2, 3, 4, \ldots$ – номер шага в направлении крутого восхождения. Результаты расчета внесем в таблицу 21, в которой продолжена нумерация опытов таблицы 16. Для вхождения в стационарную область продублируем каждый опыт только два раза, $n = 2$ (п. 5).

Таблица 21 – План крутого восхождения

N	h	x_1, ц/га	x_2, ц/га	Y_1^h, ц/га	Y_2^h, ц/га	$\overline{Y^h}$, ц/га
6	1	1.375	0.656	36.9	37.5	37.2
7	2	1.750	0.812	43.5	44.1	43.8
8	3	2.125	0.968	49.0	49.2	49.1
9	4	2.500	1.124	53.2	52.8	53.0
10	5	2.875	1.280	56.4	56.0	56.2
11	6	3.250	1.436	56.4	56.5	56.6
12	**7**	**3.625**	**1.592**	**56.9**	**57.5**	**57.2**
13	8	4.000	1.748	55.7	55.1	55.4
14	9	4.375	1.904	51.8	52.2	52.5

Из реализованных 9-ти опытов (№№ 6 – 14) максимальное значение параметра $Y = 57.2$ ц/га достигается в опыте № 12. Координаты наилучшего опыта с некоторым округлением принимаются за центр нового факторного пространства (в данном случае $x_{10} = 3.5$ ц/га, $x_{20} = 1.5$ ц/га) (п. 6).

Далее для каждого фактора выбирают новые значения интервалов варьирования, строят новый план ЦПФЭ (ЦДФЭ), находят новое уравнение регрессии первого порядка. Если оно окажется адекватным, то повторяют крутое восхождение, если нет, то это означает, что достигнута стационарная область и следует приступить к построению уравнения регрессии второго порядка (см. базовую задачу 3.6.3).

3.5.4. Материалы для самостоятельной работы, лабораторных и практических занятий

Тесты

1. В результате исследований получено адекватное уравнение регрессии первого порядка: $Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3$. Градиент этой функции выражается следующим уравнением:

$$\text{а)}\ \overline{\text{grad}\,Y} = b_0\vec{l} + b_1\vec{i} + b_2\vec{j} + b_3\vec{k}\,,$$

где $\vec{l}, \vec{i}, \vec{j}, \vec{k}$ в направлении осей $0Y$, $0X_1$; $0X_2$, $0X_3$;

$$\text{б)}\ \text{grad}\,Y = b_1\vec{i} + b_2\vec{j} + b_3\vec{k}\,,$$

где $\vec{i}, \vec{j}, \vec{k}$ в направлении осей $0X_1$; $0X_2$, $0X_3$;

$$\text{в)}\ \text{grad}\,Y = \begin{Vmatrix} b_1 \\ b_2 \\ b_3 \end{Vmatrix};\ \text{г)}\ \text{grad}\,Y = \begin{Vmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{Vmatrix}.$$

2. Параметрическое уравнение прямой линии для уравнения регрессии, приведенного в п. 1, проходящее через центр факторного пространства параллельно градиенту функции, в нормированных значениях факторов равно:

$$\text{а)}\ \begin{cases} X_1 = \xi b_1 \\ X_2 = \mu b_2 \\ X_3 = \zeta b_3 \end{cases};\ \text{б)}\ \begin{cases} X_1 = \xi b_1 \\ X_2 = \xi b_2 \\ X_3 = \xi b_3 \end{cases};\ \text{в)}\ \begin{cases} b_1 = \xi X_1 \\ b_2 = \xi X_2 \\ b_3 = \xi X_3 \end{cases};\ \text{г)}\ \begin{cases} X_1 = -\xi b_1 \\ X_2 = -\xi b_2 \\ X_3 = -\xi b_3 \end{cases}.$$

3. Параметрическое уравнение прямой линии для уравнения регрессии, приведенного в п. 1, проходящее через центр факторного пространства параллельно градиенту функции, в натуральных значениях факторов равно:

$$\text{а)}\ \begin{cases} x_1 = x_{10} + \xi b_1 \Delta x_1 \\ x_2 = x_{20} + \xi b_2 \Delta x_2 \\ x_3 = x_{30} + \xi b_3 \Delta x_3 \end{cases};\quad \text{б)}\ \begin{cases} x_1 = \Delta x_1 + \xi b_1 x_{10} \\ x_2 = \Delta x_2 + \xi b_2 x_{20} \\ x_3 = \Delta x_3 + \xi b_3 x_{30} \end{cases};$$

$$\text{в)}\ \begin{cases} x_1 = b_1 \Delta x_1 + \xi x_{10} \\ x_2 = b_2 \Delta x_2 + \xi x_{20} \\ x_3 = b_3 \Delta x_3 + \xi x_{30} \end{cases};\quad \text{г)}\ \begin{cases} x_1 = x_{10} + \xi b_1 \Delta x_1 \\ x_2 = x_{20} + \mu b_2 \Delta x_2 \\ x_3 = x_{30} + \zeta b_3 \Delta x_3 \end{cases}.$$

4. Натуральные значения факторов опытов при крутом восхождении равны:

а) $\begin{cases} x_1^h = x_{10} + h\mu \dfrac{b_1}{b} \Delta x_1 \\ x_2^h = x_{20} + h\mu \dfrac{b_2}{b} \Delta x_2 \\ x_3^h = x_{30} + h\mu \dfrac{b_3}{b} \Delta x_3 \end{cases}$;   б) $\begin{cases} x_1^h = x_{10} + h\mu \Delta x_1 \\ x_2^h = x_{20} + h\mu \Delta x_2 \\ x_3^h = x_{30} + h\mu \Delta x_3 \end{cases}$;

в) $\begin{cases} x_1^h = x_{10} + h\mu b b_1 \Delta x_1 \\ x_2^h = x_{20} + h\mu b b_2 \Delta x_2 \\ x_3^h = x_{30} + h\mu b b_3 \Delta x_3 \end{cases}$,

где $0.7 \le \mu \le 0.9$, $b = \max|b_r|$, $r = 1, 2, 3$.

Вопросы для самопроверки

1. Что такое стационарная область факторного пространства?
2. На каком этапе исследований целесообразно применять метод крутого восхождения?
3. Напишите параметрическое уравнение прямой крутого восхождения, проходящей через начало координат факторного пространства, в нормированных значениях факторов.
4. Чем отличается крутое восхождение от крутого спуска?
5. Напишите параметрическое уравнение прямой крутого спуска, проходящей через начало координат факторного пространства, в натуральных значениях факторов.
6. Сформулируйте алгоритм расчета первого шага при движении в факторном пространстве в направлении крутого восхождения (спуска).
7. Когда следует прекратить крутое восхождение и какова стратегия дальнейшего исследования?

Задачи для самостоятельного решения

Уровень 2

1. Влажность зерна $Y(\%)$, подвергаемого сушке, зависит от температуры воздуха x_1 и его скорости x_2. Был проведен эксперимент, в котором температура воздуха варьировалась в диапазоне от 40 до 70 ºС, а скорость потока от 2 до 5 м/с. Адекватное уравнение регрессии первого порядка имеет следующий вид: $Y(\%) = 60 - 12X_1 - 8X_2$. Рассчитать условия 5-ти опытов по крутому спуску (взять $\mu = 0.75$).

Ответ: при $\mu = -0.75$; $x_{1о} = 55$ ºС, $\Delta x_1 = 15$ ºС; $x_{2о} = 3.5$ м/с, $\Delta x_2 = 1.5$ м/с:

N	x_1, ºС	x_2, м/с
1	66.3	4.3
2	77.5	5.0
3	88.8	5.8
4	100.0	6.5
5	111.3	7.3

2. Привес цыплят Y(г/сут.) зависит от количества витаминной добавки x_1 и калорийности корма x_2. Проведен эксперимент, в котором витаминная добавка x_1 варьировалось в диапазоне от 0 до 12 г/сут., а калорийность корма x_2 от 8 до 14 ккал/сут. Адекватное уравнение регрессии первого порядка имеет следующий вид: $Y(\text{г/сут}) = 17 + 4.0X_1 + 3.0X_2$. Рассчитать условия 4-х опытов по крутому восхождению (взять $\mu = 0.8$).

Ответ: при $\mu = 0.8$; $x_{10} = 6$ г/сут., $\Delta x_1 = 6$ г/сут.; $x_{20} = 11$ ккал/сут., $\Delta x_2 = 4$ ккал/сут.:

N	x_1, г/сут.	x_2, ккал/сут.
1	10.80	13.4
2	15.6	15.8
3	20.4	18.2
4	25.2	20.6
5	30.0	23.0

3.6. Ортогональное центральное композиционное планирование (тема 10)

Уравнение регрессии первого порядка, моделирующее поведение изучаемого объекта в стационарной области, будет неадекватно вследствие существенной кривизны истинной функции, описывающей поведение изучаемого объекта. Поэтому для моделирования исследуемого объекта в стационарной области следует использовать уравнением регрессии второго порядка (см. уравнение (145))

$$Y = b_0 + \sum_{r=1}^{k} b_r X_r + \sum_{\substack{r=1 \\ r<s}}^{k} b_{rs} X_r X_s + \sum_{r=1}^{k} b_{rr} X_r^2 . \qquad (224)$$

Для построения уравнения регрессии второго порядка известно большое количество планов экспериментов, но все они должны удовлетворять следующим требованиям:

– число опытов N_k должно быть больше количества коэффициентов регрессии уравнения второго порядка, равного B_k (см. уравнение (137)):

$$B_k = 1 + k + \frac{k(k-1)}{2} + k = 1 + 2k + \frac{k(k-1)}{2} = \frac{(k+1)(k+2)}{2} ; \qquad (225)$$

– для оценки квадратичные эффектов число уровней варьирования каждого фактора должно быть не менее трех.

3.6.1. Теория

Рассмотрим только один тип плана второго порядка, называемый центральным композиционным планом (ЦКП). ЦКП состоит из двух структурных блоков:

– ЦДФЭ с числом опытов $N_k = N_{kg} + 1 = 2^{k-g} + 1$, где g – число генерирующих соотношений (при $g = 0$ план ЦДФЭ переходит в план ЦПФЭ с числом опытов $N_k = N_{k0} + 1 = 2^{k-0} + 1 = 2^k + 1$);

– опыты в «звездных» точках, по 2 опыта на каждый фактор. Общее число опытов в «звездных» точках $2k$, координаты которых в факторном пространстве: $(\pm\alpha_k, 0, \ldots, 0)$, $(0, \pm\alpha_k, \ldots, 0)$, $(0, 0, \ldots, \pm\alpha_k)$, где α_k – звездное плечо, величина которого, как будет показано ниже, зависит от числа факторов;

Общее количество опытов ЦКП равно:

$$N_k = 2^{k-g} + 2k + 1 . \qquad (226)$$

Так как число опытов должно быть больше количества коэффициентов регрессии уравнения второго порядка (см. уравнение (225)), то

$$2^{k-g} + 2k + 1 > 1 + 2k + \frac{k(k-1)}{2} . \qquad (227)$$

Из неравенства (228) следует (вывод приведен в приложении 6):

$$g < k + 1 - \frac{\lg[k(k-1)]}{\lg 2} . \qquad (228)$$

В таблице 22 приведены значения числа генерирующих соотношений g, рассчитанных по уравнению (228) для числа факторов k, которые могут быть на практике $k = 2, \ldots, 10$.

Таблица 22 – Значения числа g для ЦКП

k	$Y = b_0 + \sum_{r=1}^{k} b_r X_r + \sum_{r<s}^{k} b_{rs} X_r X_s + \sum_{r=1}^{k} b_{rr} X_r^2$
	g
2	0
3	0, 1
4	0, 1
5	0, 1
6	0, 1, 2
7	0, 1, 2
8	0, 1, 2, 3
9	0, 1, 2, 3
10	0, 1, 2, 3, 4

Приведенный способ построения ЦКП оправдывает свое название:

Ц – центральный, так как содержит опыты в центре плана, а все остальные опыты симметричны относительно центра факторного пространства;

К – композиционный, так как к плану ЦДФЭ (ЦПФЭ) добавляются опыты в «звездных» точках.

Композиционность ЦКП делает его достаточно экономным, так как новые опыты добавляются к уже реализованным по плану ЦДФЭ (ЦПФЭ).

В таблице 23 приведен ЦКП для построения двухфакторного уравнения регрессии второго порядка (уравнение (224)).

Таблица 23 – ЦКП для $k = 2$

N	X_0	X_1	X_2	X_1X_2	X_1^2	X_2^2	Примечание
1	+	–	–	+	+	+	Опыты ЦПФЭ
2	+	+	–	–	+	+	
3	+	–	+	–	+	+	
4	+	+	+	+	+	+	
5	+	0	0	0	0	0	
6	+	$+\alpha_2$	0	0	α_2^2	0	Опыты в «звездных» точках
7	+	$-\alpha_2$	0	0	α_2^2	0	
8	+	0	$+\alpha_2$	0	0	α_2^2	
9	+	0	$-\alpha_2$	0	0	α_2^2	
$\sum X^2$	9	0	0	0	$4+2\alpha_2^2$	$4+2\alpha_2^2$	

Анализ ЦКП (таблица 23) показывает, что не все факторы этого плана взаимно ортогональны. Факторы X_0, X_1, X_2, X_1X_2 взаимно ортогональны, а факторы X_1^2, X_2^2 не ортогональны ни между собой, ни с фактором X_0, так как $\sum_{j=1}^{N_k} X_{0j}X_{rj}^2 = 4 + 2\alpha_2^2 \neq 0$, $\sum_{j=1}^{N_k} X_{1j}^2 X_{2j}^2 = 4 \neq 0$.

Поэтому система уравнений для расчета коэффициентов обычного уравнения регрессии второго порядка, полученная МНК, не будет иметь диагональный вид. Это существенно усложняет расчеты коэффициентов уравнения регрессии, а самое главное, не позволит оценить их независимо друг от друга. Для того чтобы коэффициенты уравнения регрессии можно было вычислить независимо друг от друга по относительно простым формулам, как для планов ЦПФЭ (см. уравнения (146)–(147)), необходимо ортогонализировать квадратичные факторы (см., например, аналогичную процедуру для квадратичного члена однофакторного уравнения второго порядка (101)).

Для ортогонализации k квадратичных членов введем следующее преобразование: $X_{rj}^2 \to X_{rj}^2 - \lambda_k$, где λ_k – ортогонализирующий коэффициент, который, так же как и α_k, зависит от количества факторов k. План эксперимента ЦКП, в котором квадратичные члены ортогонализированы, в дальнейшем будем называть ортогональным центральным композиционным планом (ОЦКП). План ОЦКП в общем виде приведен в таблице 24 (число факторов k, число опытов ЦДФЭ (ЦПФЭ) обозначено единым символом N_{kg} (при $g = 0$ справедливо $N_{kg} = N_{k0}$).

Таблица 24 – ОЦКП для k факторов

N_k	X_0	X_1	…	X_k	X_1X_2	…	$X_{k-1}X_k$	$X_1^2-\lambda_k$	…	$X_k^2-\lambda_k$
1	+	–	–	–	+	+	+	$1-\lambda_k$	$1-\lambda_k$	$1-\lambda_k$
…	+	…	…	…	…	…	…	$1-\lambda_k$	$1-\lambda_k$	$1-\lambda_k$
N_{kg}	+	+	+	+	+	+	+	$1-\lambda_k$	$1-\lambda_k$	$1-\lambda_k$
$N_{kg}+1$	+	0	0	0	0	0	0	$-\lambda_k$	$-\lambda_k$	$-\lambda_k$
$N_{kg}+2$	+	$-\alpha_k$	0	0	0	0	0	$\alpha_k^2-\lambda_k$	$-\lambda_k$	$-\lambda_k$
$N_{kg}+3$	+	$+\alpha_k$	0	0	0	0	0	$\alpha_k^2-\lambda_k$	$-\lambda_k$	$-\lambda_k$
…	+	0	…	0	0	0	0	$-\lambda_k$	…	$-\lambda_k$
N_k-1	+	0	0	$-\alpha_k$	0	0	0	$-\lambda_k$	$-\lambda_k$	$\alpha_k^2-\lambda_k$
N_k	+	0	0	$+\alpha_k$	0	0	0	$-\lambda_k$	$-\lambda_k$	$\alpha_k^2-\lambda_k$
$\sum_{j=1}^{N_k} X^2$	N_k	$\sqrt{N_{kg}N_k}$	$\sqrt{N_{kg}N_k}$	$\sqrt{N_{kg}N_k}$	N_{kg}	N_{kg}	N_{kg}	$2\alpha_k^4$	$2\alpha_k^4$	$2\alpha_k^4$

Требование ортогональности всех факторов - необходимое условие независимого определения коэффициентов многофакторного уравнения регрессии второго порядка, а также дисперсий их значимости по относительно простым формулам. Условия ортогональности всех факторов имеют следующий вид (см. таблицу 24):

$$\sum_{j=1}^{N_k} X_{0j} X_{rj} = 0, \; r = 1, ..., k; \tag{229}$$

$$\sum_{j=1}^{N_k} X_{0j} X_{rsj} = 0, \; r < s; \; r = 1, ..., k-1; \; s = 2, ..., k; \tag{230}$$

$$\sum_{j=1}^{N_k} X_{rj} X_{sj} = 0, \; r, s = 1, ..., k; \tag{231}$$

$$\sum_{\substack{j=1 \\ v<w}}^{N_k} X_{rj} X_{vwj} = 0, \; r = 1, ..., k \;\; v = 1, ..., k-1; \; w = 2, ..., k; \tag{232}$$

$$\sum_{j=1}^{N_k} X_{rj}(X_{sj}^2 - \lambda_k) = 0\,, \quad r, s = 1, ..., k\,; \qquad (233)$$

$$\sum_{\substack{j=1 \\ r<s,\ v<w}}^{N_k} X_{rsj} X_{vwj} = 0\,, \quad r, v = 1, ..., k-1;\ \ s, w = 2, ..., k\,; \qquad (234)$$

$$\sum_{\substack{j=1 \\ r<s}}^{N_k} X_{rsj}(X_{vj}^2 - \lambda_k) = 0\,, \quad \begin{array}{l} r = 1, ..., k-1;\ \ s = 2, ..., k; \\ v = 1, ..., k; \end{array} \qquad (235)$$

$$\sum_{j=1}^{N_k} X_{0j}^2(X_{rj}^2 - \lambda_k) = 0\,, \quad r = 1, ..., k\,; \qquad (236)$$

$$\sum_{j=1}^{N_k} (X_{rj}^2 - \lambda_k)(X_{sj}^2 - \lambda_k) = 0\,, \quad r, s = 1, ..., k\,. \qquad (237)$$

Нетрудно показать, что уравнения (229)–(235) выполняются при любых α_k и λ_k (убедитесь в этом самостоятельно, используя данные плана ОЦКП, приведенные в таблице 24). А вот уравнения (236) и (237) выполняются только при определенных значениях α_k и λ_k. Решение системы уравнений (236), (237) с учетом данных таблицы 24, имеет следующий вид (вывод приведен в приложении 6):

$$\alpha_k = \sqrt{\frac{\sqrt{N_k \cdot N_{kg}} - N_{kg}}{2}}\,, \qquad (238)$$

$$\lambda_k = \sqrt{\frac{N_{kg}}{N_k}}\,. \qquad (239)$$

Для ОЦКП при $k = 2$; $g = 0$; $N_{k0} = N_{20} = 2^k = 2^2 = 4$; $N_k = N_2 = N_{20} + 2k + 1 = 4 + 2 \cdot 2 + 1 = 9$ по уравнениям (238), (239) получаем:

$$\alpha_2 = 1,$$

$$\lambda_2 = 0.6667.$$

Двухфакторный ОЦКП приведен в таблице 25.

Таблица 25 – Двухфакторный ОЦКП ($k = 2$)

N_2	X_0	X_1	X_2	X_1X_2	$X_1^2 - 0.6667$	$X_2^2 - 0.6667$	Примечание
1	+	–	–	+	0.3333	0.3333	Опыты ЦПФЭ $N_{20} + 1 = 5$
2	+	+	–	–	0.3333	0.3333	
3	+	–	+	–	0.3333	0.3333	
4	+	+	+	+	0.3333	0.3333	
5	+	0	0	0	– 0.6667	– 0.6667	
6	+	–	0	0	0.3333	– 0.6667	«Звездные» опыты $2k = 4$
7	+	+	0	0	0.3333	– 0.6667	
8	+	0	–	0	– 0.6667	0.3333	
9	+	0	+	0	– 0.6667	0.3333	
$\sum_{j=1}^{9} X_j$	9	0	0	0	0	0	
$\sum_{j=1}^{9} X_j^2$	9	6	6	4	2	2	

Легко убедиться в том, что все факторы плана, приведенного в таблице 25, взаимно ортогональны (убедитесь в этом самостоятельно). Кроме того, для каждого фактора в таблице 25 приведена сумма квадратов его элементов, знание которых, как будет показано ниже, необходимо для расчетов коэффициентов регрессии и дисперсий их значимости.

Для ОЦКП при $k = 3$; $g = 0$; $N_{k0} = N_{30} = 2^k = 2^3 = 8$; $N_k = N_3 = N_{k0} + 2k + 1 = 8 + 2 \cdot 3 + 1 = 15$ по уравнениям (238), (239) получаем:

$$\alpha_2 = 1.2154,$$

$$\lambda_2 = 0.7303.$$

Трехфакторный ОЦКП приведен в таблице 26.

И в этом случае легко убедиться в том, что все факторы плана, приведенного в таблице 26, взаимно ортогональны (убедитесь в этом самостоятельно). Для каждого фактора таблицы 26 приведена сумма квадратов его элементов, знание которых необходимо для расчетов коэффициентов регрессии и дисперсий их значимости.

Таблица 26 – Трехфакторный ОЦКП (k = 3)

N_3	X_0	X_1	X_2	X_3	X_1X_2	X_1X_3	X_2X_3	$X_1^2 - 0.7303$	$X_2^2 - 0.7303$	$X_3^2 - 0.7303$
1	+	–	–	–	+	+	+	0.2697	0.2697	0.2697
2	+	+	–	–	–	–	+	0.2697	0.2697	0.2697
3	+	–	+	–	–	+	–	0.2697	0.2697	0.2697
4	+	+	+	–	+	–	–	0.2697	0.2697	0.2697
5	+	–	–	+	+	–	–	0.2697	0.2697	0.2697
6	+	+	–	+	–	+	–	0.2697	0.2697	0.2697
7	+	–	+	+	–	–	+	0.2697	0.2697	0.2697
8	+	+	+	+	+	+	+	0.2697	0.2697	0.2697
9	+	0	0	0	0	0	0	–0.7303	–0.7303	–0.7303
10	+	–1.2154	0	0	0	0	0	0.7469	–0.7303	–0.7303
11	+	+1.2154	0	0	0	0	0	0.7469	–0.7303	–0.7303
12	+	0	–1.2154	0	0	0	0	–0.7303	0.7469	–0.7303
13	+	0	+1.2154	0	0	0	0	–0.7303	0.7469	–0.7303
14	+	0	0	–1.2154	0	0	0	–0.7303	–0.7303	0.7469
15	+	0	0	+1.2154	0	0	0	–0.7303	–0.7303	0.7469
$\sum_{j=1}^{15} X_j$	15	0	0	0	0	0	0	0	0	0
$\sum_{j=1}^{15} X_j^2$	15	10.9545	10.9545	10.9545	8	8	8	4.3644	4.3644	4.3644

В ОЦКП все факторы ортогональны и поэтому уравнение регрессии второго порядка следует строить в ортогонализированной форме, а не в обычной виде уравнения (224), в которой часть факторов не ортогональна:

$$Y = b_0^* + b_1^* X_1 + \ldots + b_k^* X_k + b_{12}^* X_1 X_2 + \ldots + b_{(k-1)k}^* X_{k-1} X_k +$$

$$+ b_{11}^* (X_1^2 - \lambda_k) + \ldots + b_{kk}^* (X_k^2 - \lambda_k). \qquad (240)$$

Так как система факторов ортогонализированного уравнения регрессии (240) несколько отличается от системы факторов уравнения регрессии в обычной форме (224), то коэффициенты регрессии ортогонализированного уравнения помечены знаком (*), так как они могут отличаться от коэффициентов уравнения (240).

Благодаря тому, что все факторы ОЦКП ортогональны (см. уравнения (229)–(237)), коэффициенты ортогонализированного уравнения регрессии второго порядка (по аналогии с формулами (107), (154), (155), (161)) рассчитываются по относительно простым формулам (приводятся без доказательства):

$$b_0^* = \frac{\sum_{j=1}^{N_k} X_{0j} \overline{Y}_j}{\sum_{j=1}^{N_k} X_{0j}^2} = \frac{\sum_{j=1}^{N_k} X_{0j} \overline{Y}_j}{B_0}; \qquad (241)$$

$$b_r^* = \frac{\sum_{j=1}^{N_k} X_{rj} \overline{Y}_j}{\sum_{j=1}^{N_k} X_{rj}^2} = \frac{\sum_{j=1}^{N_k} X_{rj} \overline{Y}_j}{B_r}, \qquad r = 1, \ldots, k; \qquad (242)$$

$$b_{rs}^* = \frac{\sum_{\substack{j=1 \\ r<s}}^{N_k} \left(X_{rj} X_{sj}\right) \overline{Y}_j}{\sum_{\substack{j=1 \\ r<s}}^{N_k} \left(X_{rj} X_{sj}\right)^2} = \frac{\sum_{\substack{j=1 \\ r<s}}^{N_k} \left(X_{rj} X_{sj}\right) \overline{Y}_j}{B_{rs}}, \; r = 1, \ldots, k-1; \; s = 2, \ldots, k; \qquad (243)$$

$$b_{rr}^* = \frac{\sum_{j=1}^{N_k} \left(X_{rj}^2 - \lambda_k\right) \overline{Y}_j}{\sum_{j=1}^{N_k} \left(X_{rj}^2 - \lambda_k\right)^2} = \frac{\sum_{j=1}^{N_k} \left(X_{rj}^2 - \lambda_k\right) \overline{Y}_j}{B_{rr}}, \; r = 1, \ldots, k, \qquad (244)$$

где B_0, B_r, B_{rs}, B_{rr} – константы, введенные для упрощения записи формул (241)–(244), равны суммам квадратов элементов столбцов соответствующих факторов. Эти константы, зависящие от количества факторов k и структуры плана ОЦКП (на базе ЦПФЭ или ЦДФЭ), равны следующим значениям (вывод приведен в приложении 6):

$$B_0 = \sum_{j=1}^{N_k} X_{0j}^2 = N_k; \qquad (245)$$

$$B_r = \sum_{j=1}^{N_k} X_{rj}^2 = N_{kg} + 2\alpha_k^2 = \sqrt{N_{kg} N_k}, \; r = 1, \ldots, k; \qquad (246)$$

$$B_{rs} = \sum_{j=1}^{N_k} (X_{rj} X_{sj})^2 = N_{kg}, \quad r < s, \; r = 1, \ldots, k-1, \; s = 2, \ldots, k; \qquad (247)$$

$$B_{rr} = \sum_{j=1}^{N_k} (X_{rj}^2 - \lambda_k)^2 = 2\alpha_k^4, \quad r = 1, \ldots, k. \qquad (248)$$

Из уравнения (246) следует, что $B_1 = B_2 = \ldots = B_k$.

Из уравнения (247) следует, что $B_{12} = B_{13} = \ldots = B_{k-1,\,k}$.

Из уравнения (248) следует, что $B_{11} = B_{22} = \ldots = B_{kk}$.

Сравнивая уравнения регрессии второго порядка в обычном виде (224) с аналогичным уравнением в ортогонализированной форме (240), можно определить взаимосвязь их коэффициентов:

$$b_r = b_r^*, \quad r = 1, \ldots, k; \qquad (249)$$

$$b_{rs} = b_{rs}^*, \; r < s, \quad r = 1, \ldots, k-1, \; s = 2, \ldots, k; \qquad (250)$$

$$b_{rr} = b_{rr}^*, \quad r = 1, \ldots, k; \qquad (251)$$

$$b_0 = b_0^* - \lambda_k \left(b_{11}^* + \ldots + b_{kk}^*\right). \qquad (252)$$

Для построения ОЦКП для различных k и g рекомендуем использовать приведенные в таблице 27 формулы и константы.

Таблица 27 – Константы для построения ОЦКП и уравнения регрессии второго порядка

Константы	Количество факторов k						
	2	3	3	4	4	5	6
$B_k = (k+2)(k+2)/2$	6	10	10	15	15	21	28
g	0	1	0	1	0	1	2
$N_{kg} = 2^{k-g}$	4	4	8	8	16	16	16
$2k$	4	6	6	8	8	10	12
$N_k = N_{kg} + 2k + 1$	9	11	15	17	25	27	29
$\alpha_k^2 = \left(\sqrt{N_{kg}N_k} - N_{kg}\right)\Big/2$	1	1.3166	1.4772	1.8310	2	2.3923	2.7703
α_k	1	1.1474	1.2154	1.3531	1.4142	1.5467	1.6644
$\lambda_k = \sqrt{N_{kg}/N_k}$	0.6667	0.6030	0.7303	0.6860	0.8	0.7698	0.7428
$1 - \lambda_k$	0.3333	0.3970	0.2697	0.3140	0.2	0.2302	0.2575
$\alpha_k^2 - \lambda_k$	0.3333	0.7136	0.7469	1.1415	1.2	1.6225	2.0275
$B_0 = \sum\limits_{j=1}^{N_k} X_{0j}^2 = N_k$	9	11	15	17	25	27	29
$B_r = \sum\limits_{j=1}^{N_k} X_{rj}^2 = \sqrt{N_{kg}N_k}$	6	6.6332	10.9545	11.6619	20	20.7846	21.5407
$B_{rs} = \sum\limits_{j=1}^{N_k} (X_{rj}X_{sj})^2 = N_{kg}$	4	4	8	8	16	16	16
$B_{rr} = \sum\limits_{j=1}^{N_k} \left(X_{rj}^2 - \lambda_k\right)^2 = 2\alpha_k^4$	2	3.4670	4.3644	6.7048	8	11.4462	15.3495
$B_{00} = \dfrac{2\alpha_k^4 N_k}{2\alpha_k^4 + kN_{kg}}$	1.8	2.4657	2.3080	2.9449	2.7778	3.3796	3.9976

Формулы для расчета коэффициентов уравнения регрессии второго порядка обычного вида (224), рассчитанные по уравнениям (249)-(252) с учетом (241) – (248), имеют следующий вид:

$$b_0 = \frac{\sum_{j=1}^{N_k} X_{0j} \overline{Y}_j}{N_k} - \lambda_k \sum_{r=1}^{k} b_{rr}\,, \tag{253}$$

$$b_r = \frac{\sum_{j=1}^{N_k} X_{rj} \overline{Y}_j}{\sqrt{N_{kg} N_k}}\,, \quad r = 1, ..., k\,, \tag{254}$$

$$b_{rs} = \frac{\sum_{\substack{j=1 \\ r<s}}^{N_k} \left(X_{rj} X_{sj} \right) \overline{Y}_j}{N_{kg}}\,, \qquad r = 1, ..., k-1; \;\; s = 2, ..., k\,, \tag{255}$$

$$b_{rr} = \frac{\sum_{j=1}^{N_k} \left(X_{rj}^2 - \lambda_k \right) \overline{Y}_j}{2\alpha_k^4}\,, \quad r = 1, ..., k\,, \tag{256}$$

где
$$\alpha_k^2 = \frac{\sqrt{N_k \cdot N_{kg}} - N_{kg}}{2}\,, \quad \lambda_k = \sqrt{\frac{N_{kg}}{N_k}}\,. \tag{257}$$

После расчета коэффициентов регрессии необходимо произвести статистическую оценку качества полученной модели: проверить коэффициенты уравнения регрессии на значимость, а само уравнение регрессии на адекватность.

Проверка коэффициентов уравнения регрессии на значимость по критерию Стьюдента. Расчет дисперсии воспроизводимости производится по всем опытам ОЦКП при числе опытов $N_k = 2^{k-g} + 2k + 1$ (см., например, уравнение (148)):

$$S_{\text{воспр}}^2 = \frac{\sum_{j=1}^{N_k} S_j^2}{N_k}\,, \tag{258}$$

Расчеты дисперсий значимости коэффициентов ортогонализированного уравнения регрессии (240) производится по следующим формулам (см., например, уравнения (112) и (149)–(150)):

$$S^2(b_0^*) = \frac{S_{\text{воспр}}^2}{n\sum_{j=1}^{N_k} X_{0j}^2} = \frac{S_{\text{воспр}}^2}{nB_0}\,; \tag{259}$$

$$S^2(b_r^*) = \frac{S_{\text{воспр}}^2}{n\sum_{j=1}^{N_k} X_{rj}^2} = \frac{S_{\text{воспр}}^2}{nB_r}\,, \qquad r = 1, \ldots, k; \tag{260}$$

$$S^2(b_{rs}^*) = \frac{S_{\text{воспр}}^2}{n\sum_{\substack{j=1 \\ r<s}}^{N_k} \left(X_{rj} X_{sj} \right)^2} = \frac{S_{\text{воспр}}^2}{nB_{rs}}\,, \; r = 1, ..., k-1; \; s = 2, ..., k\,; \tag{261}$$

$$S^2(b_{rr}^*) = \frac{S_{\text{воспр}}^2}{n\sum_{j=1}^{N_k} \left(X_{rj}^2 - \lambda_k \right)^2} = \frac{S_{\text{воспр}}^2}{nB_{rr}}\,, \quad r = 1, \ldots, k. \tag{262}$$

В формулах (259) (262) использованы ранее введенные константы B_0, B_r, B_{rs}, B_{rr} (см. уравнения (245)–(248)), значения которых для различных значений k и g приведены в таблице 27.

Дисперсии значимости $S^2(b_0), S^2(b_r), S^2(b_{rs}), S^2(b_{rr})$ коэффициентов b_0, b_r, b_{rs}, b_{rr} уравнения регрессии второго порядка в обычном виде (224) могут быть рассчитаны с учетом установленной ранее взаимосвязи коэффициентов обычного и ортогонализированного уравнений регрессии (см. уравнения (249)–(252)), а также уравнений (36), (37) и равны следующим значениям:

$$S^2(b_r) = S^2(b_r^*), \tag{263}$$

$$S^2(b_{rs}) = S^2(b_{rs}^*), \tag{264}$$

$$S^2(b_{rr}) = S^2(b_{rr}^*)\,, \tag{265}$$

$$S^2(b_0) = S^2(b_0^*) + \lambda_k^2 \left[S^2(b_{11}^*) + \ldots + S^2(b_{kk}^*) \right]. \tag{266}$$

В явном виде формулы для расчета дисперсий значимости коэффициентов уравнения регрессии второго порядка в обычном виде, рассчитанные по уравнениям (263)–(266) с учетом (245)–(248) имеют следующий вид:

$$S^2(b_r)=\frac{S^2_{\text{воспр}}}{n\sum\limits_{j=1}^{N_k}X_{rj}^2}=\frac{S^2_{\text{воспр}}}{n\sqrt{N_{kg}N_k}},\quad r=1,...,k;\tag{267}$$

$$S^2(b_{rs})=\frac{S^2_{\text{воспр}}}{n\sum\limits_{\substack{j=1\\r<s}}^{N_k}\left(X_{rj}X_{sj}\right)^2}=\frac{S^2_{\text{воспр}}}{nN_{kg}},\quad r=1,...,k-1;\; s=2,...,k;\tag{268}$$

$$S^2(b_{rr})=\frac{S^2_{\text{воспр}}}{n\sum\limits_{j=1}^{N_k}\left(X_{rj}^2-\lambda_k\right)^2}=\frac{S^2_{\text{воспр}}}{2n\alpha_k^4},\quad r=1,\ldots,k;\tag{269}$$

$$S^2(b_0)=\frac{S^2_{\text{воспр}}}{n\sum\limits_{j=1}^{N_k}X_{0j}^2}+\frac{\lambda^2kS^2_{\text{воспр}}}{n\sum\limits_{j=1}^{N_k}\left(X_{rj}^2-\lambda_k\right)^2}=\frac{S^2_{\text{воспр}}}{nB_{00}},\tag{270}$$

где
$$B_{00}=\frac{2\alpha_k^4N_k}{2\alpha_k^4+kN_{kg}}.\tag{271}$$

Вывод уравнений (270), (271) приведен в приложении 6, а численные значения вновь введенной константы B_{00} для различных значений k и g приведены в таблице 27.

Доверительные интервалы коэффициентов регрессии Δb_r для обычного уравнения регрессии второго порядка, построенного по ОЦКП, рассчитываются по критерию Стьюдента (уравнение (78)), число степеней свободы которого $f=N_k\left(n-1\right)$.

Процедура проверки многофакторного уравнения регрессии второго порядка на адекватность полностью совпадает с аналогичной процедурой для однофакторного уравнения регрессии второго порядка (см. уравнения (115)–(121)).

Уравнение регрессии второго порядка позволяет оценить предельную абсолютную погрешность параметра Y, рассчитанного по уравнению регрессии второго порядка в ортогонализированной форме. Факторы X_0, X_r, X_rX_s, X_r^2 в уравнении регрессии второго порядка в обычном виде (224) не ортогональны (см. таблицу 23) и поэтому регрессионные коэффициенты этого уравнения b_0, b_r, b_{rs} b_{rr} взаимозависимы (см. уравнения (249)–(252)). Следовательно, применить уравнения (37) для расчета дисперсии $S^2(Y)$ по уравнению регрессии второго порядка в обычном виде нельзя. А вот факторы X_0, X_r, X_rX_s, $(X_r^2-\lambda_k)$ ортогонализированного уравнения регрессии второго порядка (240) ортогональны, поэтому регрессионные коэффициенты $b_0^*, b_r^*, b_{rs}^*, b_{rr}^*$ независимы. Применяя к нему уравнения (36), (37) получим

$$\Delta Y(X_1,...,X_k)=t_{N_k(n-1),\,p}\cdot S(Y),\tag{272}$$

где
$$S^2(Y)=S^2(b_0^*)+S^2(b_r^*)\sum_{r=1}^{k}X_r^2+S^2(b_{rs}^*)\sum_{\substack{r=1\\r<s}}^{k}(X_rX_s)^2+$$

$$+S^2(b_{rr}^*)\sum_{r=1}^{k}(X_r^2-\lambda_k)^2.\tag{273}$$

С учетом уравнений (238), (239), (245)–(248), (259)–(262) уравнение (273) имеет следующий вид (вывод приведен в приложении 6):

$$S^2(Y)=\frac{S^2_{\text{воспр}}}{nN_k}\left(1+\frac{\sum\limits_{r=1}^{k}X_r^2}{\lambda_k}+\frac{\sum\limits_{r=1,\,r<s}^{k}(X_rX_s)^2}{\lambda_k^2}+\frac{2\sum\limits_{r=1}^{k}(X_r^2-\lambda_k)^2}{N_{kg}(1-\lambda_k)^2}\right).\tag{274}$$

Подставляя уравнение (274) в (272), получим

$$\Delta Y(X_1,...,X_k)=$$

$$=t_{N_k\,(n-1),\,p}\frac{S_{\text{воспр}}}{\sqrt{nN_k}}\sqrt{1+\frac{\sum\limits_{r=1}^{k}X_r^2}{\lambda_k}+\frac{\sum\limits_{r=1,\,r<s}^{k}(X_rX_s)^2}{\lambda_k^2}+\frac{2\sum\limits_{r=1}^{k}(X_r^2-\lambda_k)^2}{N_{kg}(1-\lambda_k)^2}}.\tag{275}$$

Анализ уравнения (275) позволяет сделать несколько выводов:

а) минимальная предельная абсолютная погрешность достигается не в центре плана (как для уравнения регрессии первого порядка, см. уравнение (157)), а в точках (вывод приведен в приложении 6):

$$X_1^2 = ... = X_k^2 = \lambda_k \frac{4\lambda_k^2 - N_{kg}(1-\lambda_k)^2}{4\lambda_k^2 + (k-1)N_{kg}(1-\lambda_k)^2} = \gamma_k \lambda_k , \qquad (276)$$

где

$$\gamma_k = \frac{4\lambda_k^2 - N_{kg}(1-\lambda_k)^2}{4\lambda_k^2 + (k-1)N_{kg}(1-\lambda_k)^2} , \qquad (277)$$

$$\Delta Y(\sqrt{\gamma_k \lambda_k}, ..., \sqrt{\gamma_k \lambda_k}) = \frac{t_{N_k(n-1),\, p} \cdot S_{\text{воспр}}}{\sqrt{nN_k}} \sqrt{1 + k\gamma + \frac{k(k\text{-}1)}{2}\gamma^2 + \frac{2k\lambda_k^2(1-\gamma)^2}{N_{kg}(1-\lambda_k)^2}} ; \qquad (278)$$

б) предельная абсолютная погрешность в центре факторного плана $X_1 = \ldots X_k = 0$ равна:

$$\Delta Y(0, ..., 0) = \frac{t_{N_k(n-1),\, p} \cdot S_{\text{воспр}}}{\sqrt{nN_k}} \sqrt{1 + \frac{2k\lambda_k^2}{N_{kg}(1-\lambda_k)^2}} ; \qquad (279)$$

в) предельная абсолютная погрешность на границах факторного плана $X_1 = \ldots X_k = 1$ равна:

$$\Delta Y(1, ..., 1) = \frac{t_{N_k(n-1),\, p} \cdot S_{\text{воспр}}}{\sqrt{nN_k}} \sqrt{1 + \frac{k}{\lambda_k} + \frac{k(k\text{-}1)}{2\lambda_k^2} + \frac{2k}{N_{kg}}} ; \qquad (280)$$

В таблице 28 приведены значения предельных абсолютных погрешностей параметра Y, рассчитанного по многофакторному уравнению регрессии второго порядка для различных точек исследованного факторного пространства и различных значений факторов k (использованы данные таблицы 27 и уравнения (278)-(280)).

г) предельная абсолютная погрешность параметра Y, рассчитанного по адекватному многофакторному уравнению регрессии второго порядка, так же как и для уравнения регрессии первого порядка, значительно меньше предельной абсолютной погрешности выборочного среднего в каждом опыте.

Подтвердим количественно это утверждение. Предельная абсолютная погрешность параметра Y в каждом опыте определяет так же, как и для однофакторного уравнения регрессии первого порядка (см. уравнение (95)). Соотношение предельных экспериментальных и расчетных ошибок, например, в центре 4-факторного плана $X_1 = X_2 = X_3 = X_4$ 0, используя данные таблицы 28 и уравнения (95), равно:

$$\frac{\Delta \overline{Y}_j}{\Delta Y(0)} \approx \frac{t_{n-1,\, p} \cdot S_{\text{воспр}}}{\sqrt{n} \cdot 0.598 \cdot S_{\text{воспр}}} = \frac{t_{n-1,\, p}}{0.598 \cdot \sqrt{n}} . \qquad (281)$$

Так как в экспериментах $n \approx 4$, $t_{3,\, 0.95} \approx 3$ (см. приложение 2), то из уравнения (281) следует, что в центре плана

$$\frac{\Delta Y_j}{\Delta Y(0)} = \frac{3}{0.598 \cdot 2} \approx 2.5 . \qquad (282)$$

То есть, предельная абсолютная погрешность рассчитанного значения параметра Y в центре плана по 4-х факторному уравнению регрессии второго порядка, которое построено для ОЦКП, примерно в 2.5 раза меньше предельной абсолютной погрешности средних значений параметров найденных экспериментально в каждом опыте.

Таблица 28 – Предельная абсолютная погрешность параметра Y, рассчитанного по многофакторному уравнению регрессии второго порядка, построенного по ОЦКП.

k	$\Delta Y(\sqrt{\gamma\lambda_k}, ..., \sqrt{\gamma\lambda_k})$	$\Delta Y(0, ..., 0)$	$\Delta Y(1, ..., 1)$
2	$0.612 \cdot S_{\text{воспр}}$	$0.765 \cdot S_{\text{воспр}}$	$0.921 \cdot S_{\text{воспр}}$
3	$0.559 \cdot S_{\text{воспр}}$	$0.663 \cdot S_{\text{воспр}}$	$0.881 \cdot S_{\text{воспр}}$
4	$0.505 \cdot S_{\text{воспр}}$	$0.598 \cdot S_{\text{воспр}}$	$0.794 \cdot S_{\text{воспр}}$
5	$0.453 \cdot S_{\text{воспр}}$	$0.551 \cdot S_{\text{воспр}}$	$0.684 \cdot S_{\text{воспр}}$

Уравнение регрессии второго порядка позволяет определить оптимальные значения факторов, при которых исследованный параметр достигает экстремум. В отличие от однофакторных уравнений регрессии второго порядка, в которых всегда существует абсолютный экстремум, в многофакторных уравнениях регрессии второго порядка возможно отсутствие абсолютного экстремума.

Если уравнение регрессии второго порядка имеет абсолютный экстремум, то необходимое условием этого – равенство нулю частных производных $\partial Y / \partial X_r = 0$ ($r = 1, \ldots, k$), а достаточное условие - $\partial^2 Y / \partial X_r^2 < 0$ (максимум), $\partial^2 Y / \partial X_r^2 > 0$ (минимум). Достаточное условие для максимума – все $b_{rr} < 0$, а минимума – все $b_{rr} > 0$. Для уравнения $Y = b_0 + \sum_{r=1}^{k} b_r X_r + \sum_{\substack{r=1 \\ r<s}}^{k} b_{rs} X_r X_s + \sum_{r=1}^{k} b_{rr} X_r^2$ необходимое условие экстремума выражается следующим образом:

$$\begin{cases} \dfrac{\partial Y}{\partial X_1} = 0 \;\rightarrow\; b_1 + b_{12} X_2 + \ldots + b_{1k} X_k + 2b_{11} X_1 = 0 \quad; \\ \dfrac{\partial Y}{\partial X_2} = 0 \;\rightarrow\; b_2 + b_{12} X_1 + \ldots + b_{2k} X_k + 2b_{22} X_2 = 0 \quad; \\ \qquad\qquad \ldots \\ \dfrac{\partial Y}{\partial X_k} = 0 \;\rightarrow\; b_k + b_{1k} X_k + \ldots + b_{k-1,k} X_{k-1} + 2b_{kk} X_k = 0. \end{cases} \tag{283}$$

Таким образом, если все коэффициенты регрессии при квадратичных членах отрицательные (или положительные), то полученное уравнение регрессии имеет абсолютный максимум (или минимум) $Y_{\max\,(\min)} = Y(X^*_{1,\,\text{опт}}, \ldots, X^*_{k,\,\text{опт}})$, где $X^*_{1,\,\text{опт}}, \ldots, X^*_{k,\,\text{опт}}$ – корни системы (283). Оптимальные значения факторов в натуральном выражении находятся по известным уравнениям (137)–(139):

$$x_{r,\,\text{опт}} = x_{r\text{o}} + X^*_{r,\,\text{опт}} \Delta x_r, \; r = 1, \ldots, k. \tag{284}$$

Рассчитав оптимальные значения факторов, при которых уравнение регрессии достигает абсолютного экстремума, полезно поставить проверочный эксперимент в точке максимума (минимума) с n дублями. Следует убедиться в том, что экспериментальное среднее значение параметра в оптимальной точке попадает в доверительный интервал расчетного значения (см. уравнение (275)).

Если максимум (минимум) лежит вне исследованного факторного пространства, возможно существенное отличие рассчитанного и экспериментального значений экстремума. В этом случае следует провести новый эксперимент по ОЦКП с центром в точке, в которой был достигнут наилучший экспериментальный результат.

Если часть коэффициентов $b_{rr} > 0$, а другая часть $b_{rr} < 0$, то уравнение регрессии второго порядка не имеет абсолютного экстремума. Тем не менее, можно найти значение локального экстремума, который находится на границе факторного пространства. Эта задача решается численными методами, которые в данном учебно-методическом комплексе не рассматриваются. Можно найти максимальное (минимальное) значение параметра, которое лежит на границе исследованного факторного пространства, варьируя значениями всех факторов в выбранной области, с помощью, например, компьютерной программы Microsoft Excel «Поиск решения».

3.6.2. Алгоритм решения прикладных задач

1. Эксперимент для построения уравнения регрессии второго порядка по плану ОЦКП базируется на результатах предшествующего эксперимента по плану ЦДФЭ (ЦПФЭ), когда уравнение регрессии, построенное по этим планам, оказывается неадекватным. Поэтому все параметры плана ЦДФЭ (ЦПФЭ), такие, как k, g, а также факторы $x_1, \ldots, x_k$ (натуральные значения), основные уровни и интервалы их варьирования $[x_{k\,\min}, x_{k\,\max}]$, $X_1, \ldots, X_k$ (нормированные значения), исследуемый параметр Y и число дублей остаются без изменения при достроении композиционного плана ОЦКП.

2. Записать уравнение регрессии второго порядка в обычном виде для описания стационарной зоны:

$$Y = b_0 + \sum_{r=1}^{k} b_r X_r + \sum_{\substack{r=1 \\ r<s}}^{k} b_{rs} X_r X_s + \sum_{r=1}^{k} b_{rr} X_r^2 .$$

3. Рассчитать по заданным k и g значения звездного плеча α_k и ортогонализирующего коэффициента λ_k:

$$\alpha_k = \sqrt{\frac{\sqrt{N_k N_{kg}} - N_{kg}}{2}} ,$$

$$\lambda_k = \sqrt{\frac{N_{kg}}{N_k}}\,.$$

Проверить выполнение неравенства:

$$g < k + 1 - \frac{\lg[k(k-1)]}{\lg 2}\,.$$

4. Перевести нормированные значения факторов в «звездных» опытах в натуральные:

$$x_{\alpha r} = x_{r\text{o}} + \Delta x_r \alpha_k\,,$$

где $x_{r\text{o}} = \dfrac{x_{r\max} + x_{r\min}}{2}$; $\Delta x_r = \dfrac{x_{r\max} - x_{r\min}}{2}$.

5. Построить план ОЦКП эксперимента с числом опытов N_k, включающий следующие опыты:

– опыты плана ЦДФЭ, в количестве $N_{kg} = 2^{k-g} + 1$;

– звездные точки $(\pm\alpha_k, 0, \ldots, 0)$, $(0, \pm\alpha_k, \ldots, 0)$, $(0, 0, \ldots, \pm\alpha_k)$ в количестве $2k$.

Значения X_0, X_r $(r = 1, \ldots, k)$, X_rX_s $(r < s$, $r = 1, \ldots, k-1$, $s = 2, \ldots, k)$, $(X_r^2 - \lambda_k)$ $(r = 1, \ldots, k)$ внести в таблицу ОЦКП.

6. Выполнить эксперимент. Экспериментальные данные с числом дублей n в каждом опыте внести в таблицу ОЦКП.

7. Провести предварительную обработку экспериментальных данных.

7.1. Рассчитать выборочное среднее в каждом опыте:

$$\overline{Y}_j = \frac{\sum_{i=1}^{n} Y_{ji}}{n}\,,\ j = 1, \ldots, N_k.$$

7.2. Рассчитать выборочную дисперсию в каждом опыте:

$$S_j^2 = \frac{\sum_{i=1}^{n}\left(Y_j - \overline{Y}_j\right)^2}{n-1}\,,\quad j = 1, \ldots, N_k.$$

7.3. Проверить экспериментальные данные каждого опыта на промахи по критерию Смирнова–Граббса (приложение 1):

$$\tau_{\text{э}j} = \max\left|\frac{Y_{ji} - \overline{Y}_j}{S_j}\right| > \tau_{n-1,\,p}\,,\quad j = 1, \ldots, N_k.$$

7.4. Проверить экспериментальные данные каждого опыта на принадлежность их к нормальному закону распределения:

$$\left|\frac{\text{САО}}{S_j} - 0.7979\right| = \left|\frac{\sum_{i=1}^{n}\left|Y_{ji} - \overline{Y}_j\right|}{nS_j} - 0.7979\right| < \frac{0.4}{\sqrt{n}}\,,\quad j = 1, \ldots, N_k.$$

7.5. Проверить дисперсии всех опытов на однородность по критерию Кохрена (приложение 5):

$$\frac{\max S_j^2}{\sum_{j=1}^{N_k} S_j^2} < G_{n-1,\,N_k,\,p}\,,$$

где $G_{n-1,\,N_{k0}+1,\,p}$ – критическое значение критерия Кохрена (приложение 5).

Внимание! Дальнейшая обработка экспериментальных данных с целью построения уравнения регрессии возможна только в том случае, если в каждом опыте нет промахов, данные каждого опыта подчиняются нормальному закону распределения, а дисперсии всех опытов однородны. Все опыты, в которых эти требования не выполнены, следует переделать.

8. В матрице ОЦКП построить столбцы $X_{0j}\overline{Y}_j$, $X_{rj}\overline{Y}_j$ $(r = 1, \ldots, k)$, $X_{rj}X_{sj}\overline{Y}_j$ $(r < s$, $r = 1, \ldots, k–1$, $s = 2, \ldots, k)$, $(X_{rj}^2 - \lambda_k)\overline{Y}_j$ $(r = 1, \ldots, k)$ и рассчитать их суммы.

9. Рассчитать коэффициенты обычного уравнения регрессии второго порядка b_0, b_r, b_{rs}, b_{rr}:

$$b_r = \frac{\sum_{j=1}^{N_k} X_{rj} \overline{Y}_j}{\sum_{j=1}^{N_k} X_{rj}^2} = \frac{\sum_{j=1}^{N_k} X_{rj} \overline{Y}_j}{\sqrt{N_{kg} N_k}},\ r = 1, \ldots, k;$$

$$b_{rs} = \frac{\sum_{\substack{j=1 \\ r<s}}^{N_k} \left(X_{rj} X_{sj} \right) \overline{Y}_j}{\sum_{\substack{j=1 \\ r<s}}^{N_k} \left(X_{rj} X_{sj} \right)^2} = \frac{\sum_{\substack{j=1 \\ r<s}}^{N_k} \left(X_{rj} X_{sj} \right) \overline{Y}_j}{N_{kg}}, \quad r = 1, \ldots, k-1;\ s = 2, \ldots, k;$$

$$b_{rr} = \frac{\sum_{j=1}^{N_k} \left(X_{rj}^2 - \lambda_k \right) \overline{Y}_j}{\sum_{j=1}^{N_k} \left(X_{rj}^2 - \lambda_k \right)^2} = \frac{\sum_{j=1}^{N_k} \left(X_{rj}^2 - \lambda_k \right) \overline{Y}_j}{2\alpha_k^4}, \quad r = 1, \ldots, k;$$

$$b_0 = \frac{\sum_{j=1}^{N_k} X_{0j} \overline{Y}_j}{\sum_{j=1}^{N_k} X_{0j}^2} - \lambda_k \sum_{r=1}^{k} b_{rr} = \frac{\sum_{j=1}^{N_k} X_{0j} \overline{Y}_j}{N_k} - \lambda_k \sum_{r=1}^{k} b_{rr}.$$

10. Проверить коэффициенты регрессии b_0, b_r, b_{rs}, b_{rr} на значимость.

10.1. Рассчитать дисперсию воспроизводимости $S_{\text{воспр}}^2$:

$$S_{\text{воспр}}^2 = \frac{\sum_{j=1}^{N_k} S_j^2}{N_k}.$$

10.2. Рассчитать дисперсии значимости коэффициентов регрессии b_0, b_r, b_{rs}, b_{rr}:

$$S^2(b_r) = \frac{S_{\text{воспр}}^2}{n \sum_{j=1}^{N_k} X_{rj}^2} = \frac{S_{\text{воспр}}^2}{n\sqrt{N_{kg} N_k}}, \quad r = 1, \ldots, k;$$

$$S^2(b_{rs}) = \frac{S_{\text{воспр}}^2}{n \sum_{\substack{j=1 \\ r<s}}^{N_k} \left(X_{rj} X_{sj} \right)^2} = \frac{S_{\text{воспр}}^2}{n N_{kg}}, \quad r = 1, \ldots, k-1,\ s = 2, \ldots, k;$$

$$S^2(b_{rr}) = \frac{S_{\text{воспр}}^2}{n \sum_{j=1}^{N_k} \left(X_{rj}^2 - \lambda_k \right)^2} = \frac{S_{\text{воспр}}^2}{2n\alpha_k^4}, \quad r = 1, \ldots, k;$$

$$S^2(b_0) = \frac{S_{\text{воспр}}^2}{n \sum_{j=1}^{N_k} X_{0j}^2} + \frac{\lambda^2 k S_{\text{воспр}}^2}{n \sum_{j=1}^{N_k} \left(X_{rj}^2 - \lambda_k \right)^2} = \frac{(2\alpha_k^4 + k N_{kg}) S_{\text{воспр}}^2}{2\alpha_k^4 N_k}.$$

10.3. Рассчитать доверительные интервалы коэффициентов регрессии b_0, b_r, b_{rs}, b_{rr}:

$$\Delta b = t_{N_k(n-1),\, p} \cdot S(b),$$

где $t_{N_k(n-1),\, p}$ – критическое значение коэффициент Стьюдента (приложение 2).

Если выполняется неравенство $\Delta b < |b|$, то соответствующий коэффициент регрессии b является статистически значимым. В противном случае, если $\Delta b > |b|$ коэффициент регрессии b статистически незначим. Незначимые коэффициенты из уравнения регрессии следует исключить. Если некоторые коэффициенты b_{rr} окажутся незначимы, то этот факт следует учесть при расчете коэффициента b_0.

11. Проверить полученное уравнение регрессии на адекватность.

11.1. Рассчитать Y_j^p в каждом опыте $j = 1, \ldots, N_k$ по уравнению регрессии второго порядка.

11.2. Рассчитать квадрат разности экспериментальных $\overline{Y}_j$ и расчетных Y_j^p значений $(\overline{Y}_j - Y_j^p)^2$ в каждом опыте $j = 1, \ldots, N_k$, а также остаточную сумму квадратов φ:

$$\varphi = \sum_{j=1}^{N_k} (\overline{Y}_j - Y_j^p)^2.$$

11.3. Рассчитать дисперсию адекватности:

$$S^2_{\text{ад}} = \frac{n\varphi}{N_k - B}, f = N_k - B,$$

где B – число значимых коэффициентов регрессии.

11.4. Рассчитать экспериментальное значение критерия Фишера:

$$F_э = \frac{\max\,(S^2_{\text{ад}}, S^2_{\text{воспр}})}{\min\,(S^2_{\text{ад}}, S^2_{\text{воспр}})}.$$

11.5. Если $F_э = \dfrac{S^2_{\text{ад}}}{S^2_{\text{воспр}}} < F_{N-B,\,N(n-1),\,p}$ при $S^2_{\text{ад}} > S^2_{\text{воспр}}$,

где $F_{N-B,\,N(n-1),\,p}$ – критическое значение критерия Фишера (приложение 4), то $\sigma^2_{\text{ад}} = \sigma^2_{\text{воспр}}$ и уравнение регрессии второго порядка адекватно. Следует перейти к процедуре оптимизации исследуемого объекта (см. п. 13 данного алгоритма).

11.6. Если $F_э = \dfrac{S^2_{\text{ад}}}{S^2_{\text{воспр}}} > F_{N-B,\,N(n-1),\,p}$ при $S^2_{\text{ад}} > S^2_{\text{воспр}}$,

то $\sigma^2_{\text{ад}} > \sigma^2_{\text{воспр}}$ и уравнение регрессии второго порядка неадекватно. Следует перейти к построению уравнения регрессии третьего порядка (в данном учебно-методическом комплексе процедура построения таких моделей не рассматривается).

11.7. Если $F_э = \dfrac{S^2_{\text{воспр}}}{S^2_{\text{ад}}} < F_{N(n-1),\,N-B,\,p}$ при $S^2_{\text{ад}} < S^2_{\text{воспр}}$,

где $F_{N(n-1),\,N-B,\,p}$ – критическое значение критерия Фишера (приложение 4), то $\sigma^2_{\text{ад}} = \sigma^2_{\text{воспр}}$ и уравнение регрессии второго порядка адекватно. Следует перейти к процедуре оптимизации исследуемого объекта (см. п. 13 данного алгоритма).

11.8. Если $F_э = \dfrac{S^2_{\text{воспр}}}{S^2_{\text{ад}}} > F_{N(n-1),\,N-B,\,p}$ при $S^2_{\text{ад}} < S^2_{\text{воспр}}$,

то $\sigma^2_{\text{воспр}} > \sigma^2_{\text{ад}}$. Вероятность такого случая на практике ничтожно мала. Такой случай является следствием либо плохой организации эксперимента, в котором нарушена идея случайности (нарушены основные статистические предпосылки), либо недостатка степеней свободы для оценки дисперсии адекватности. Необходимо повторить эксперимент, причем желательно увеличить число дублей n.

12. Предельная абсолютная погрешность параметра Y, рассчитанного по уравнению регрессии второго порядка, определяется следующим уравнением:

$$\Delta Y(X_1, ..., X_k) =$$

$$= t_{N_k\,(n-1),\,p}\frac{S_{\text{воспр}}}{\sqrt{nN_k}}\sqrt{1 + \frac{\sum_{r=1}^{k} X_r^2}{\lambda_k} + \frac{\sum_{r=1,\,r<s}^{k}(X_r X_s)^2}{\lambda_k^2} + \frac{2\sum_{r=1}^{k}(X_r^2 - \lambda_k)^2}{N_{kg}(1-\lambda_k)^2}}.$$

13. Если обычное уравнение регрессии адекватно $Y = b_0 + \sum_{r=1}^{k} b_r X_r + \sum_{\substack{r=1\\ r<s}}^{k} b_{rs} X_r X_s + \sum_{r=1}^{k} b_{rr} X_r^2$ и все коэффициенты регрессии при квадратичных членах отрицательные (положительные), то уравнение регрессии имеет абсолютный максимум (минимум) $Y_{\max\,(\min)} = Y(X^*_{1,\,\text{опт}}, ..., X^*_{k,\,\text{опт}})$, где $X^*_{1,\,\text{опт}}, ..., X^*_{k,\,\text{опт}}$ – оптимальные значения факторов, являющиеся корнями системы уравнений:

$$\begin{cases} 2b_{11}X_1 + b_{12}X_2 + b_{13}X_3 + ... + b_{1(k-2)}X_{k-2} + b_{1(k-1)}X_{k-1} + b_{1k}X_k = -b_1, \\ b_{12}X_1 + 2b_{22}X_2 + b_{23}X_3 + ... + b_{2(k-2)}X_{k-2} + b_{2(k-1)}X_{k-1} + b_{2k}X_k = -b_2, \\ ... \\ b_{1k}X_1 + b_{2k}X_2 + b_{3k}X_3 + ... + b_{(k-2)k}X_{k-2} + b_{(k-1)k}X_{k-1} + 2b_{kk}X_k = -b_k. \end{cases}$$

В случае, если максимум (минимум) лежит вне факторного пространства ОЦКП, и рассчитанное значение $Y_{\max\,(\min)}$ статистически значимо отличается от экспериментального, то следует провести новый эксперимент по ОЦКП, а за центр плана взять наилучшую экспериментальную точку.

В случае, когда одна часть коэффициентов $b_{rr} > 0$, а другая часть $b_{rr} < 0$, то уравнение регрессии второго порядка не имеет абсолютного экстремума. Найти локальный экстремум в исследованном факторном пространстве, например, с помощью компьютерной программы Microsoft Excel «Поиск решения».

14. В заключение, если в этом есть необходимость, построить многофакторное уравнение регрессии второго порядка в натуральных значениях факторов. Для выполнения этой процедуры в уравнении регрессии второго порядка нормированные значения факторов X_r необходимо заменить на натуральные значения x_r по формуле

$$X_r = \frac{x_r - x_{r0}}{\Delta x_r}.$$

3.6.3. Базовая задача

Продолжим решение задачи, начатой в базовых задачах (3.3.3) и (3.5.3).

Планирование эксперимента.

В результате крутого восхождения наилучшие результаты по урожайности $Y = 57.2$ ц/га получены в опыте со значениями факторов $x_1 = 3.61$ ц/га и $x_2 = 1.58$ ц/га (см. базовую задачу (3.5.3)). Значения факторов этого опыте с некоторым округлением принимм за центр нового факторного пространства: $x_{10} = 3.5$ ц/га и $x_{20} = 1.5$ ц/га.

Параметр. В качестве эффективного параметра, описывающего исследуемый объект, оставим ту же величину – урожайность пшеницы Y (ц/га), количество дублей в каждом опыте так же принимаем равным 4-м ($n = 4$).

Факторы. Так как крутое восхождение показало, что вне исследованной области факторного пространства (см. базовую задачу 3.4.3) зависимость параметра от факторов существенно отличается от линейной (см. таблицу 21), то для более надежного выявления квадратичного эффекта и гарантированного включения максимума функции в новую область факторного пространства несколько увеличим интервалы варьирования (таблица 29).

План эксперимента. Так же как и в базовой задаче 3.3.3. исследование изучаемого объекта начнем с построения многофакторного уравнения регрессии первого порядка. Построим для этого план ЦПФЭ, который для двух факторов ($k = 2$) состоит из 5 опытов ($N_k = N_{k0} + 1 = 2^k + 1 = 5$). Матрица планирования ЦПФЭ представ-

лена в таблице 30. Ссылки при построении многофакторного уравнения регрессии первого порядка на 3.3.2.

Таблица 29 – Основные уровни и интервалы варьирования факторов в новой области факторного пространства

Факторы	1-й фактор (семена)		2-й фактор (удобрение)	
	x_1, ц/га	X_1	x_2, ц/га	X_2
Основной уровень	$x_{10} = 3.5$	0	$x_{20} = 1.5$	0
Интервал варьирования	$\Delta x_1 = 1$	1	$\Delta x_2 = 1$	1
Верхний уровень	$x_{1\,max} = 4.5$	+1	$X_{2\,max} = 2.5$	+1
Нижний уровень	$x_{1\,min} = 2.5$	–1	$X_{2\,min} = 0.5$	–1
Формулы перевода натуральных значений факторов x_r в нормированные X_r и обратно	$X_{1j} = \frac{x_{1j} - 3.5}{1}$; $x_{1j} = 3.5 + X_{1j}$ $X_{2j} = \frac{x_{1j} - 1.5}{1}$; $x_{2j} = 1.5 + X_{2j}$			

Методика эксперимента повторяет методику, описанную в базовой задаче 3.3.3. Значения факторов X_{0j}, X_{1j}, X_{2j} и результаты эксперимента Y_{ji} внесены в соответствующие столбцы, таблицы 30.

Предварительная обработка экспериментальных данных.

Вычисление выборочного среднего в каждом опыте (п. 6.1):

$$\overline{Y}_j = \frac{\sum_{i=1}^{n} Y_{ji}}{n}, \quad j = 1, ..., 5.$$

Например, $\overline{Y}_1 = (43.2 + 42.9 + 42.9 + 42.7)/4 = 42.93$.

Вычисление выборочной дисперсии в каждом опыте (п. 6.2):

$$S_j^2 = \frac{\sum_{i=1}^{n} (Y_{ji} - \overline{Y}_j)^2}{n-1}, \quad j = 1, ..., 5.$$

Например, $S_1^2 = \frac{(43.2 - 42.93)^2 + (42.9 - 42.93)^2}{4-1} +$

$$+ \frac{(42.9 - 42.93)^2 + (42.7 - 42.93)^2}{4-1} = 0.0425, \quad S_1 = \sqrt{0.0425} = 0.206.$$

Таблица 30 – Матрица планирования и результаты анализа экспериментальных данных

План эксперимента				Результаты опытов и обработки эксперимента										
N	X_{0j}	X_{1j}	X_{2j}	Y_{j1}	Y_{j2}	Y_{j3}	Y_{j4}	$\overline{Y}_j$	S_j^2	$X_{0j}\overline{Y}_j$	$X_{1j}\overline{Y}_j$	$X_{2j}\overline{Y}_j$	Y_j^p	$(\overline{Y}_j - Y_j^p)^2$
1	1	–	–	43.2	42.9	42.9	42.7	42.93	0.0425	42.93	–42.93	–42.93	44.90	3.88
2	1	+	–	35.3	35.2	35.5	35.5	35.38	0.0225	35.38	35.38	–35.38	37.64	5.11
3	1	–	+	57.7	57.6	57.5	57.4	57.55	0.0167	57.55	–57.55	57.55	59.82	5.15
4	1	+	+	50.7	50.2	50.7	50.7	50.58	0.0625	50.58	50.58	50.58	52.56	3.92
5	1	0	0	57.1	57.6	57.0	57.1	57.20	0.0733	57.20	0	0	48.73	71.74
$\sum_{j=1}^{5} X_{rj}^2$	5	4	4					$\sum_{j=1}^{5} S_j^2 = 0.2536$		243.64	–14.52	29.82	$\varphi = 89.8$	
b	48.73	–3.63	7.46	$t_{15,\,0.95} = 2.131$				$S^2_{\text{восп}} = 0.0507$					$S^2_{\text{ад}} = 179.6$	
$S^2(b)$	0.00254	0.00317	0.0032					$G_э = 0.289$					$F_э = 3542$ (!)	
$S(b)$	0.0504	0.0563	0.056					$G_{3,\,5,\,0.95} = 0.598$					$F_{2,\,15,\,0.95} = 3.682$	
Δb	0.11	0.12	0.12										**Модель неадекватна**	

Анализ данных каждого опыта на промахи (п. 6.3) показал, что во всех опытах №№ 1–5 промахов нет, Например, для опыта № 1 наиболее выделяющимся является значение 43.2. Критерий Смирнова–Граббса (п. 6.3):

$$\tau_{э1} = \max\left|Y_{1i} - \overline{Y}_1\right| / S_1 = |43.2 - 42.93| / 0.206 = 1.31 .$$

В опыте № 1 промахов нет, так как критическое значение критерия Смирнова–Граббса, найденное из приложения 1 при числе степеней свободы $f = n - 2 = 4 - 2 = 2$, равно $\tau_{2,\ 0.95} = 1.69$ ($\tau_{э1} = 1.31 < \tau_{2,\,0.95} = 1.69$).

Анализ на принадлежность случайных значений параметра $\boldsymbol{Y_{ji}}$ в каждом опыте к нормальному закону распределения (п. 6.4) показал, что экспериментальные данные подчиняются нормальному закону распределения. Например, значения опыта № 1 Y_{1i} подчиняются нормальному закону распределения, так как:

$$|\text{САО}_1 / S_1 - 0.7979| < 0.4/\sqrt{n} \rightarrow |0.140/0.206 - 0.7979| < 0.4/\sqrt{4} \rightarrow$$

$$\rightarrow 0.1183 < 0.2 \text{, где } \text{САО}_1 = \frac{\sum_{i=1}^{n}\left|Y_{1i} - \overline{Y}_1\right|}{n} = \frac{|43.2 - 42.93|}{4} +$$

$$+\frac{|42.9 - 42.93| + |42.9 - 42.93| + |42.7 - 42.93|}{4} = 0.140 .$$

Проверка дисперсий всех опытов на однородность по критерию Кохрена (п. 6.5) показала, что все дисперсии однородны, так как $G_э < G_{n-1,\ N_{k0}+1,\ p} = G_{3,\,5,\,0.94} = 0.598$ (приложение 5).

$$G_э = \frac{\max S_j^2}{\sum_{j=1}^{5} S_j^2} = \frac{0.0733}{0.0425 + 0.0225 + 0.0167 + 0.0625 + 0.0733} =$$

$$= \frac{0.0733}{0.2536} = 0.289 .$$

Поскольку предварительная обработка экспериментальных данных в каждом опыте не выявила отклонений от основных предпосылок регрессионного анализа (отсутствие промахов, нормальный закон распределения и однородность дисперсий), переходим к следующему этапу – построению уравнения регрессии первого порядка.

Построение уравнения регрессии первого порядка. Моделирование исследуемого объекта в новой области факторного пространства начнем, как было сказано выше, с построения уравнения регрессии первого порядка: $Y = b_0X_0 + b_1X_1 + b_2X_2$.

Образуем факторы $X_{0j}\overline{Y}_j$, $X_{1j}\overline{Y}_j$, $X_{2j}\overline{Y}_j$ и рассчитаем их суммы (п. 7):

$$\sum_{j=1}^{5} X_{0j}\overline{Y}_j = 1\cdot 42.93 + 1\cdot 35.38 + 1\cdot 57.55 + 1\cdot 50.58 + 1\cdot 57.20 = 243.64 ,$$

$$\sum_{j=1}^{5} X_{0j}\overline{Y}_j = (-1)\cdot 42.93 + 1\cdot 35.38 + (-1)\cdot 57.55 +$$

$$+ 1\cdot 50.58 + 0\cdot 57.20 = -14.52 ,$$

$$\sum_{j=1}^{5} X_{2j}\overline{Y}_j = (-1)\cdot 42.93 + (-1)\cdot 35.38 + 1\cdot 57.55 +$$

$$+ 1\cdot 50.58 + 0\cdot 57.20 = 29.82 .$$

Рассчитаем коэффициенты уравнения регрессии первого порядка b_0, b_1, b_2 (п. 9):

$$b_0 = \frac{\sum_{j=1}^{N_k} X_{0j}\overline{Y}_j}{\sum_{j=1}^{N_k} X_{0j}^2} = \frac{\sum_{j=1}^{N_k} X_{0j}\overline{Y}_j}{N_k} = \frac{243.64}{5} = 48.73 ,$$

$$b_1 = \frac{\sum_{j=1}^{N_k} X_{1j}\overline{Y}_j}{\sum_{j=1}^{N_k} X_{1j}^2} = \frac{\sum_{j=1}^{N_k} X_{1j}\overline{Y}_j}{N_{k0}} = \frac{-14.52}{4} = -3.63 ,$$

$$b_2 = \frac{\sum_{j=1}^{N_k} X_{2j}\overline{Y}_j}{\sum_{j=1}^{N_k} X_{2j}^2} = \frac{\sum_{j=1}^{N_k} X_{2j}\overline{Y}_j}{N_{k0}} = \frac{29.82}{4} = 7.46 .$$

Проверка коэффициентов уравнения регрессии первого порядка на значимость.

Дисперсия воспроизводимости (п. 10.1):

$$S^2_{\text{воспр}} = \frac{\sum_{j=1}^{N_k} S_j^2}{N_k} = \frac{0.2536}{5} = 0.0507 .$$

Дисперсии значимости коэффициентов регрессии b_0, b_1, b_2 (п. 10.2):

$$S^2(b_0) = \frac{S^2_{\text{воспр}}}{nN_k} = \frac{0.0507}{4 \cdot 5} = 0.00254 \to S(b_0) = 0.0504 ,$$

$$S^2(b_1) = S^2(b_2) = \frac{S^2_{\text{воспр}}}{nN_{k0}} = \frac{0.0507}{4 \cdot 4} = 0.00317 \to S\left(b_1\right) = S(b_2) = 0.0563 .$$

Расчет доверительных интервалов коэффициентов регрессии (п. 10.3):

$$\Delta b_0 = t_{N_k(n-1),\, p} S(b_0) = t_{15,\, 0.95} S(b_0) = 2.131 \cdot 0.0504 = 0.1074 \approx 0.11 ,$$

$$\Delta b_1 = \Delta b_2 = t_{N_k(n-1),\, p} S(b_1) = t_{15,\, 0{,}95} S(b_1) = 2.131 \cdot 0.0563 = 0.1200 \approx 0.12 ,$$

где $t_{N_k(n-1),\, p} = t_{15,\, 0.95} = 2.131$ – критическое значение критерия Стьюдента (приложение 2).

Все коэффициенты регрессии b_0, b_1, b_2 значимы, так как $\Delta b_0 < |b_0|$ (0.11 < 48.73); $\Delta b_1 < |b_1|$ (0.12 < 3.63); и $\Delta b_2 < |b_2|$ (0.12 < 7.46) (п. 10.4).

Получено уравнение регрессии первого порядка $Y = 48.73 — 3.63X_1 + 7.46X_2$, в котором все коэффициенты регрессии значимы.

Анализ уравнения регрессии первого порядка на адекватность.

Расчетные значения параметра Y_j^p определяются по уравнению $Y = 48.73 - 3.63X_1 + 7.46X_2$ (п. 11.1).

Например, $Y_1^p = 48.73 - 3.63 \cdot (-1) + 7.46 \cdot (-1) = 44.90$.

Остаточная сумма квадратов φ (п. 11.2):

$$\varphi = \sum_{j=1}^{5} (\overline{Y}_j - Y_j^p)^2 = (42.93 - 44.90)^2 + (35.38 - 37.64)^2 +$$

$$(57.55 - 59.82)^2 + (50.58 - 52.56)^2 + (57.20 - 48.73)^2 =$$

$$= 3.88 + 5.11 + 5.15 + 3.92 + 71.74 = 89.8 .$$

Дисперсия адекватности (п. 11.3):

$$S^2_{\text{ад}} = \frac{n\varphi}{N_k - B} = \frac{4 \cdot 89.8}{5 - 3} = 179.6 ,\ f = N_{k0} + 1 - B = 4 + 1 - 3 = 2 .$$

Экспериментальное значение критерия Фишера (п. 11.4):

$$F_{\text{э}} = \frac{S^2_{\text{ад}}}{S^2_{\text{воспр}}} = \frac{179.8}{0.0507} = 3542 \text{, так как } S^2_{\text{ад}} > S^2_{\text{воспр}} .$$

Так как $F_{\text{э}} = 3542 > F_{(N_k - B),\, N_k(n-1),\, p} = F_{2,\, 15,\, 0.95} = 3.682$ (приложение 4), то $\sigma^2_{\text{ад}} > \sigma^2_{\text{воспр}}$. Это означает, что полученное уравнение регрессии первого порядка $Y = 48.73 - 3.63X_1 + 7.46X_2$ неадекватно (п. 11.6). Поэтому для описания исследуемого объекта следует перейти к построению уравнения регрессии второго порядка $Y = b_0 + b_1X_1 + b_2X_2 + b_{12}X_1X_2 + b_{11}X_1^2 + b_{22}X_2^2$.

Для построения двухфакторного уравнения регрессии второго порядка (k = 2) используем ОЦКП. Для этого к уже проведенным 5-ти опытам ЦПФЭ добавим еще 4 «звездных» опыта (п. 5). При построении двухфакторного уравнения регрессии второго порядка ссылки даны пункты алгоритма 3.6.2.

Рассчитаем величину звездного плеча α_2 и величину ортогонализирующего коэффициента λ_2 (N_{k0} = 4, N_k = 9) (п. 3):

$$\alpha_2 = \sqrt{\frac{\sqrt{N_k \cdot N_{k0}} - N_{k0}}{2}} = \sqrt{\frac{\sqrt{9 \cdot 4} - 4}{2}} = 1 ,$$

$$\lambda_2 = \sqrt{\frac{N_{k0}}{N_k}} = \sqrt{\frac{4}{9}} = \frac{2}{3} = 0.6667 .$$

Построим план эксперимента ОЦКП (таблица 31): столбцы X_{0j}, X_{1j}, X_{2j}, X_{1j}, X_{2j}, $(X_{1j}^2 - 0.6667)$, $(X_{2j}^2 - 0.6667)$.

Методика эксперимента. Так как количество дополнительных опытов равно $2k = 4$ ($k = 2$), а число дублей в каждом опыте $n = 4$, организуем дополнительно 16 делянок по 1 м^2. Методика проведения эксперимента та же, что и в базовой задаче (3.3.3). Результаты проведенных опытов Y_{ji} №№ 1–5, а также дополнительных опытов №№ 6–9 внесены в таблицу 31 (п. 6).

Предварительная обработка экспериментальных данных в дополнительных опытах №№ 6 – 9 (данные опытов №№ 1 – 5 уже обработаны и результаты, которые приведены в таблице 30, перенесены в таблицу 31).

Вычисление выборочного среднего в каждом дополнительном опыте (п. 7.1):

$$\overline{Y}_j = \frac{\sum_{i=1}^{n} Y_{ji}}{n}, \ j = 6, ..., 9.$$

Например, $\overline{Y}_9 = (57.5 + 57.4 + 57.6 + 57.6)/4 = 57.53$.

Вычисление выборочной дисперсии в каждом дополнительном опыте (п. 7.2).

$$S_j^2 = \frac{\sum_{i=1}^{n} (Y_{ji} - \overline{Y}_j)^2}{n-1}, \quad j = 6, ..., 9.$$

Например, $S_9^2 = \dfrac{(57.5 - 57.53)^2 + (57.4 - 57.53)^2}{4-1} +$

$$+\frac{(57.6 - 57.53)^2 + (57.6 - 57.53)^2}{4-1} = 0.00920, \ S_9 = \sqrt{0.0092} = 0.0959.$$

Таблица 31 – План ОЦКП и результаты анализа экспериментальных данных

План эксперимента и результаты опытов										
N	X_{0j}	X_{1j}	X_{2j}	X_{12j}	$X_{1j}^2-\lambda_2$	$X_{2j}^2-\lambda_2$	Y_{1j}	Y_{2j}	Y_{3j}	Y_{4j}
1	1	–	–	+	0.3333	0.3333	43.2	42.9	42.9	42.7
2	1	+	–	–	0.3333	0.3333	35.3	35.2	35.5	35.5
3	1	–	+	–	0.3333	0.3333	57.7	57.6	57.5	57.4
4	1	+	+	+	0.3333	0.3333	50.7	50.2	50.7	50.7
5	1	0	0	0	–0.6667	–0.6667	57.1	57.6	57.0	57.1
6	1	–	0	0	0.3333	–0.6667	57.5	57.1	57.2	57.0
7	1	+	0	0	0.3333	–0.6667	50.0	50.5	50.4	50.5
8	1	0	–	0	–0.6667	0.3333	42.7	42.6	42.5	42.8
9	1	0	+	0	–0.6667	0.3333	57.5	57.4	57.6	57.6
$\sum_{j=1}^{9} X_{rj}^2$	9	6	6	4	2	2				
b	57.24	–3.56	7.45	0.15	–3.46	–7.16				
$S^2(b)$	0.0053	0.0016	0.0016	0.0024	0.0048	0.0048	$t_{27,\,0.95} = 2.052$			
$S(b)$	0.073	0.040	0.040	0.049	0.069	0.069				
Δb	0.15	0.08	0.08	0.10	0.14	0.14				

Окончание таблицы 31.

Результаты обработки эксперимента										
N	$\overline{Y}_j$	S_j^2	$X_{0j}\overline{Y}_j$	$X_{1j}\overline{Y}_j$	$X_{2j}\overline{Y}_j$	$X_{1j}X_{2j}\overline{Y}_j$	$(X_{1j}^2-\lambda_2)\overline{Y}_j$	$(X_{1j}^2-\lambda_2)\overline{Y}$	Y_j^p	$(\overline{Y}_j-Y_j^p)^2$
1	42.93	0.0425	42.93	– 42.93	– 42.93	42.93	14.31	14.31	42.87	0.0036
2	35.38	0.0225	35.38	35.38	– 35.38	– 35.38	11.79	11.79	35.46	0.0064
3	57.55	0.0167	57.55	– 57.55	57.55	– 57.55	19.18	19.18	57.48	0.0049
4	50.58	0.0625	50.58	50.58	50.58	50.58	16.86	16.86	50.65	0.0049
5	57.20	0.0733	57.20	0	0	0	– 38.13	– 38.14	57.24	0.0016
6	57.20	0.0467	57.20	– 57.20	0	0	19.06	– 38.14	57.33	0.0169
7	50.35	0.0567	50.35	50.35	0	0	16.78	– 33.57	50.21	0.0196
8	42.65	0.0167	42.65	0	42.65	0	– 28.43	14.22	42.63	0.0004
9	57.53	0.0092	57.53	0	57.53	0	– 38.36	19.17	57.53	0.0000
$\sum_{j=1}^{9} S_j^2 = 0.3468$			451.37	–21.37	44.70	0.58	–6.94	–14.32	$\varphi = 0.0583$	
$S^2_{воспр} = 0.0385$								$S^2_{ад} = 0.0777$		
$G_э = 0.211$								$F_э = 2.018$		
$G_{3,\,9,\,0.95} = 0.403$								$F_{3,\,27,\,0.95} = 2.960$		
Дисперсии однородны								Модель адекватна		

Анализ результатов на промахи (п. 7.3) показал, что во всех дополнительных опытах промахов нет. Например, для опыта № 9 наиболее выделяющимся является значение 57.4. Критерий Смирнова–Граббса равен (п. 7.3):

$$\tau_{э9} = \max\left|Y_{9i} - \overline{Y}_9\right| / S_9 = \left|57.4 - 57.53\right| / 0.0959 = 1.36 .$$

В опыте № 9 промахов нет, так как критическое значение критерия Смирнова–Граббса, найденное из приложения 1 при числе степеней свободы $f = n - 2 = 4 - 2 = 2$, равно $\tau_{2,\ 0.95} = 1.69$ ($\tau_{э9} = 1.36 < \tau_{2,\,0.95} = 1.69$).

Анализ на принадлежность случайных значений параметра *Y_{ji}* в каждом дополнительном опыте к нормальному закону распределения (п. 6.4) показал, что экспериментальные данные подчиняются нормальному закону распределения. Например, для Y_{9i}:

$$\left|\text{САО}_9 / S_9 - 0.7979\right| < 0.4/\sqrt{n} \rightarrow \left|0.075/0.0959 - 0.7979\right| < 0.4/\sqrt{4} \rightarrow$$

$$\rightarrow 0.0158 < 0.2 ,$$

где
$$\text{САО}_9 = \frac{\sum\limits_{i=1}^{n}\left|Y_{1i} - \overline{Y}_9\right|}{n} = \frac{\left|57.5 - 57.53\right|}{4} +$$

$$+ \frac{\left|57.4 - 57.53\right| + \left|57.6 - 57.53\right| + \left|57.6 - 57.53\right|}{4} = 0.075$$

Проверка дисперсий всех опытов на однородность по критерию Кохрена (п. 7.5) показала, что все дисперсии однородны, так как $G_э < G_{n-1,\ N_k,\ p} = G_{3,\,9,\,0.94} = 0.403$ (приложение 5).

$$G_э = \frac{\max S_j^2}{\sum\limits_{j=1}^{9} S_j^2} = \frac{0.0733}{0.0425 + 0.0225 + 0.0167 + 0.0625 + 0.0733} +$$

$$+ \frac{0.0733}{0.0467 + 0.0567 + 0.0167 + 0.0092} = \frac{0.0733}{0.3468} = 0.211 .$$

Поскольку предварительная обработка экспериментальных данных в каждом опыте не выявила отклонений от основных предпосылок регрессионного анализа, переходим к следующему этапу – расчету коэффициентов уравнения регрессии второго порядка.

Расчет коэффициентов уравнения регрессии второго порядка.

Образуем столбцы (п. 8) $X_{0j}\overline{Y}_j$, $X_{1j}\overline{Y}_j$, $X_{2j}\overline{Y}$, $X_{1j}X_{2j}\overline{Y}_j$, $(X_{1j}^2 - \lambda_k)\overline{Y}_j$, $(X_{2j}^2 - \lambda_k)\overline{Y}_j$ рассчитаем их суммы по 9-ти опытам:

$$\sum_{j=1}^{9} X_{0j}\overline{Y}_j = 1\cdot 42.93 + 1\cdot 35.38 + 1\cdot 57.55 + 1\cdot 50.58 + 1\cdot 57.20 +$$

$$+ 1\cdot 57.20 + 1\cdot 50.35 + 1\cdot 42.65 + 1\cdot 57.53 = 451.37 ;$$

$$\sum_{j=1}^{9} X_{1j}\overline{Y}_j = (-1)\cdot 42.93 + 1\cdot 35.38 + (-1)\cdot 57.55 + 1\cdot 50.58 + 0\cdot 57.20 +$$

$$+ (-1)\cdot 57.20 + 1\cdot 50.35 + 0\cdot 42.65 + 0\cdot 57.53 = -21.37 ;$$

$$\sum_{j=1}^{9} X_{2j}\overline{Y}_j = (-1)\cdot 42.93 + (-1)\cdot 35.38 + 1\cdot 57.55 + 1\cdot 50.58 + 0\cdot 57.20 +$$

$$+ 0\cdot 57.20 + 0\cdot 50.35 + (-1)\cdot 42.65 + 1\cdot 57.53 = 44.70 ;$$

$$\sum_{j=1}^{9} X_{1j}X_{2j}\overline{Y}_j = 1\cdot 42.93 + (-1)\cdot 35.38 + (-1)\cdot 57.55 + 1\cdot 50.58 + 0\cdot 57.20 +$$

$$+ 0\cdot 57.20 + 0\cdot 50.35 + 0\cdot 42.65 + 0\cdot 57.53 = 0.58 ;$$

$$\sum_{j=1}^{9}(X_{1j}^2 - 0.667)\overline{Y}_j = 0.3333\cdot 42.93 + 0.3333\cdot 35.38 + 0.3333\cdot 57.55 +$$

$$+ 0.3333\cdot 50.58 + (-0.6667)\cdot 57.20 + 0.3333\cdot 57.20 + 0.3333\cdot 50.35 +$$

$$+ (-0.6667)\cdot 42.65 + (-0.6667)\cdot 57.53 = -6.94 ;$$

$$\sum_{j=1}^{9}(X_{2j}^2 - 0.667)\overline{Y}_j = 0.3333\cdot 42.93 + 0.3333\cdot 35.38 + 0.3333\cdot 57.55 +$$

$$+ 0.3333\cdot 50.58 + (-0.6667)\cdot 57.20 + (-0.6667)\cdot 57.20 +$$

$$+ (-0.6667)\cdot 50.35 + 0.3333\cdot 42.65 + 0.3333\cdot 57.53 = -14.32 .$$

Коэффициенты уравнения регрессии (п. 9):

$$b_1 = \frac{\sum_{j=1}^{9} X_{1j}\overline{Y}_j}{\sqrt{N_{k0}N_k}} = \frac{-21.37}{\sqrt{4\cdot 9}} = \frac{-21.37}{6} = -3.56\,;$$

$$b_2 = \frac{\sum_{j=1}^{9} X_{2j}\overline{Y}_j}{\sqrt{N_{k0}N_k}} = \frac{44.70}{\sqrt{4\cdot 9}} = \frac{44.70}{6} = 7.45\,;$$

$$b_{12} = \frac{\sum_{j=1}^{9} X_{12j}\overline{Y}_j}{N_{k0}} = \frac{0.58}{4} = 0.1450 \approx 0.14\,;$$

$$b_{11} = \frac{\sum_{j=1}^{N_k}\left(X_{1j}^2 - 0.667\right)\overline{Y}_j}{2\alpha_2^4} = \frac{-6.94}{2\cdot 1^4} = -3.47\,;$$

$$b_{22} = \frac{\sum_{j=1}^{N_k}\left(X_{2j}^2 - 0.667\right)\overline{Y}_j}{2\alpha_2^4} = \frac{-14.32}{2\cdot 1^4} = -7.16\,;$$

$$b_0 = \frac{\sum_{j=1}^{9} X_{0j}\overline{Y}_j}{N_k} - \lambda_2(B_{11}+B_{22}) = \frac{451.37}{9} - 0.6667\cdot(-3.47-7.16) = 57.24\,.$$

Проверка коэффициентов уравнения регрессии второго порядка на значимость.

Дисперсия воспроизводимости (п. 10.1):

$$S^2_{\text{воспр}} = \frac{\sum_{j=1}^{9} S_j^2}{N_k} = \frac{0.3468}{9} = 0.0385\,.$$

Дисперсии значимости коэффициентов регрессии (п. 10.2):

$$S^2(b_0) = \frac{(2\alpha_k^4 + kN_{k0})S^2_{\text{воспр}}}{2n\alpha_k^4 N_k} = \frac{(2\cdot 1^4 + 2\cdot 4)\cdot 0.0385}{2\cdot 4\cdot 1^4\cdot 9} = 0.00535 \;\rightarrow$$

$$\rightarrow S(b_0) = 0.0731\,;$$

$$S^2(b_1) = S^2(b_2) = \frac{S^2_{\text{воспр}}}{n\sqrt{N_{k0}N_k}} = \frac{0.0385}{4\cdot 6} = 0.00160 \rightarrow S(b_1) = S(b_2) = 0.0400\,;$$

$$S^2(b_{12}) = \frac{S^2_{\text{воспр}}}{n\cdot N_{k0}} = \frac{0.0385}{4\cdot 4} = 0.00241 \;\rightarrow\; S(b_{12}) = 0.0491\,;$$

$$S^2_{b_{11}} = S^2_{b_{22}} = \frac{S^2_{\text{воспр}}}{2n\alpha_2^4} = \frac{0.0385}{4\cdot 2\cdot 1^4} = 0.00481 \rightarrow S(b_{11}) = S(b_{22}) = 0.0694\,.$$

Расчет доверительных интервалов коэффициентов регрессии (п. 10.3):

$$\Delta b_0 = t_{N_k(n-1),\,0.95}\cdot S(b_0) = 2.052\cdot 0.0731 = 0.1500 \approx 0.15\,,$$

$$\Delta b_1 = \Delta b_2 = t_{N_k(n-1),\,0.95}\cdot S(b_1) = 2.052\cdot 0.040 = 0.0821 \approx 0.08\,,$$

$$\Delta b_{12} = t_{N_k(n-1),\,0.95}\cdot S(b_{12}) = 2.052\cdot 0.049 = 0.1005 \approx 0.10\,,$$

$$\Delta b_{11} = \Delta b_{22} = t_{N_k(n-1),\,0.95}\cdot S(b_{11}) = 2.052\cdot 0.069 = 0.1416 \approx 0.14\,.$$

где $t_{N_k(n-1),\,0.95} = t_{27,\,0.95} = 2.052$ – критическое значение критерия Стьюдента (приложение 2).

Все коэффициенты регрессии b_0, b_1, b_2, b_{12}, b_{11}, b_{22}, значимы, так как $\Delta b_0 < |b_0|$ (0.15 < 57.24); $\Delta b_1 < |b_1|$ (0.08 < 3.56), $\Delta b_2 < |b_2|$ (0.08 < 7.45), $\Delta b_{12} < |b_{12}|$ (0.10 < 0.14), $\Delta b_{11} < |b_{11}|$ (0.14 < 3.40); $\Delta b_{22} < |b_{22}|$ (0.14 < 7.28). В полученном уравнении регрессии второго порядка все коэффициенты регрессии значимы (п. 10.3):

$$Y = 57.24 - 3.56X_1 + 7.45X_2 + 0.14X_1X_2 - 3.46X_1^2 - 7.15X_2^2\,.$$

Анализ уравнения регрессии второго порядка на адекватность.

Расчетные значения параметра Y_j^p определяются по уравнению (п. 11.1):

$$Y = 57.24 - 3.56X_1 + 7.45X_2 + 0.14X_1X_2 - 3.46X_1^2 - 7.15X_2^2.$$

Например, $Y_1^p = 57.24 - 3.56\cdot(-1) + 7.45\cdot(-1) + 0.14\cdot(-1)\cdot(-1) -$

$$-3.46\cdot(-1)^2 - 7.15\cdot(-1)^2 = 42.87.$$

Расчетные значения параметра Y_j^p внесены в таблицу 31.

Остаточная сумма квадратов φ (п. 11.2):

$$\varphi = \sum_{j=1}^{9}(\overline{Y}_j - Y_j^p)^2 = (42.93-42.87)^2 + (35.38-35.46)^2 + (57.55-57.48)^2 +$$

$$+(50.58-50.65)^2 + (57.20-57.24)^2 + (57.20-57.33)^2 + (50.35-50.21)^2 +$$

$$+(42.65-42.63)^2 + (57.53-57.53)^2 = 0.0036 + 0.0064 + 0.0049 + 0.0049 +$$

$$+0.0016 + 0.0169 + 0.0196 + 0.0004 + 0.0000 = 0.0583.$$

Дисперсия адекватности (п. 11.3):

$$S_{\text{ад}}^2 = \frac{n\varphi}{N_k - B} = \frac{4\cdot 0.0583}{9-6} = 0.0777,\ f = N_k - B = 9 - 6 = 3.$$

Экспериментальное значение критерия Фишера (п. 11.4):

$$F_{\text{э}} = \frac{S_{\text{ад}}^2}{S_{\text{воспр}}^2} = \frac{0.0777}{0.0385} = 2.018,\ \text{так как}\ S_{\text{ад}}^2 > S_{\text{воспр}}^2.$$

где $F_{(N_k-B),\,N_k(n-1),\,0.95} = F_{3,\,27,\,0.95} = 2.960$ – критическое значение критерия Фишера (приложение 4).

Выполнение неравенства $F_{\text{э}} < F_{3,\,27,\,0.95}$ (3.33 < 4.75) означает, что $\sigma_{\text{ад}}^2 = \sigma_{\text{воспр}}^2$ и полученное уравнение регрессии второго порядка $Y = 57.24 - 3.56X_1 + 7.45X_2 + 0.14X_1X_2 - 3.46X_1^2 - 7.15X_2^2$ адекватно (п. 11.5).

Нахождение максимума (п. 13). Так как оба коэффициента регрессии при квадратичных членах отрицательные, то полученное уравнение имеет абсолютный максимум. Координаты точки максимума находим, решая систему уравнений:

$$\begin{cases} \dfrac{\partial Y}{\partial X_1} = 0 \ \rightarrow\ b_1 + b_{12}X_2 + 2b_{11}X_1 = 0 \\ \dfrac{\partial Y}{\partial X_2} = 0 \ \rightarrow\ b_2 + b_{12}X_1 + 2b_{22}X_2 = 0 \end{cases} \rightarrow$$

$$\begin{cases} -3.56 + 0.14X_2 - 2\cdot 3.46X_1 = 0 \\ 7.45 + 0.14X_1 - 2\cdot 7.15X_2 = 0 \end{cases} \rightarrow \left| \begin{array}{l} X_{1,\,\text{опт}} = -0.504 \\ X_{2,\,\text{опт}} = 0.516. \end{array} \right.$$

Расчетное значение $Y_{\max}$ равно:

$$Y_{\max} = 57.24 - 3.56\cdot(-0.504) + 7.45\cdot 0.516 + 0.14\cdot(-0.504)\cdot 0.516 -$$

$$-3.46\cdot(-0.504)^2 - 7.15\cdot 0.516^2 = 60.06.$$

Рассчитаем натуральные значения оптимальных значений факторов (п. 14) (для перехода от нормированных значений факторов к натуральным используются данные таблицы 26):

$$x_{1,\,\text{опт}} = x_{1\text{о}} + \Delta x_1 X_{1,\,\text{опт}} = 3.5 + 1\cdot(-0.504) = 3.00,$$
$$x_{2,\,\text{опт}} = x_{2\text{о}} + \Delta x_2 X_{2,\,\text{опт}} = 1.5 + 1\cdot 0.516 = 2.02.$$

Таким образом, максимальная урожайность пшеницы в 60.06 ц/га может быть получена при внесении посевного материала в количестве 3.0 ц/га и минерального удобрения в количестве 2.0 ц/га.

Рассчитаем предельную абсолютную погрешность $Y_{\max}$, полученную из уравнения регрессии второго порядка (п. 12).

$$\Delta Y(X_1, ..., X_k) =$$

$$= t_{N_k(n-1),\,p}\frac{S_{\text{воспр}}}{\sqrt{nN_k}}\sqrt{1 + \frac{\sum_{r=1}^{k}X_r^2}{\lambda_k} + \frac{\sum_{r=1,\,r<s}^{k}(X_rX_s)^2}{\lambda_k^2} + \frac{2\sum_{r=1}^{k}(X_r^2 - \lambda_k)^2}{N_{k0}(1-\lambda_k)^2}}.$$

Подставляя в выше написанную формулу $k = 2$; $n = 4$; $N_k = 9$; $N_{k0} = 4$; $t_{N_k(n-1),\,p} = t_{27,\,0.95} = 2.052$ (приложение 2); $\lambda_k = 0.6667$; $S_{\text{воспр}} = \sqrt{S_{\text{воспр}}^2} = \sqrt{0.0385} = 0.196$; $X_{1,\,\text{опт}}^2 = (-0.504)^2 = 0.254$;

$X_{2,\,\text{опт}}^2 = 0.516^2 = 0.266$; $\quad X_{1,\,\text{опт}}^2 X_{2,\,\text{опт}}^2 = (-0.504)^2 \cdot 0.516^2 = 0.0676$;

$(X_{1,\,\text{опт}}^2 - \lambda_k)^2 = \left[(-0.504)^2 - 0.6667\right]^2 = 0.170$;

$(X_{2,\,\text{опт}}^2 - \lambda_k)^2 = (0.516^2 - 0.6667)^2 = 0.160$, получаем:

$$\Delta Y_{\max} = \frac{2.052 \cdot 0.196}{\sqrt{4 \cdot 9}} \times$$

$$\times \sqrt{1 + \frac{0.254 + 0.266}{0.6667} + \frac{0.0676}{0.6667^2} + \frac{2 \cdot (0.170 + 0.160)}{4 \cdot (1 - 0.6667)^2}} = 0.124 \approx 0.12\,.$$

Таким образом, полученное двухфакторное уравнение регрессии второго порядка позволило рассчитать оптимальные значения факторов ($x_{1,\text{опт}} = 3.0$ ц/га; $x_{1,\text{опт}} = 3.0$ ц/га), при которых достигается максимальная урожайность пшеницы $Y_{\max} = 60.06 \pm 0.12$ ц/га (см. правила округления в алгоритме 1.3.2). Как и следовало ожидать, расчетная погрешность в определении максимальной урожайности $\Delta Y_{\max} = 0.12$ меньше экспериментальной погрешности определения той же величины по 4-м дублям $\overline{\Delta Y}_{\max} = 0.31$, так как

$$\overline{\Delta Y}_{\max} = \frac{t_{3,\,0.95} \cdot S_{\text{воспр}}}{\sqrt{n}} = \frac{3.182 \cdot 0.196}{\sqrt{4}} = 0.312 \approx 0.31 \text{ ц/га},$$

где $t_{n-1} = 3.182$ – критическое значение критерия Стьюдента при степенях свободы $f = 3$ (приложение 2). Для того чтобы экспериментальная погрешность определения параметра $Y_{\max}$ равнялась бы 0.12, число дублей n должно было быть не менее 13 (убедитесь в этом самостоятельно).

3.6.4. Материалы для самостоятельной работы, лабораторных и практических занятий

Тесты

1. Общее количество опытов ОЦКП равно:

а) $N_k = N_{kg} + 2k + 2 = 2^{k-g} + 2k + 2$;

б) $N_k = N_{kg} + 2k + 1 = 2^{k-g} + 2k + 1$;

в) $N_k = N_{k0} + 2k + 1 = 2^k + 2k + 1$.

2. Ортогонализирующий коэффициент равен:

а) $\lambda_k = \sqrt{\dfrac{N_{kg}}{N_k}}$; $\quad$ б) $\lambda_k = \sqrt{\dfrac{N_{k0}}{N_k}}$;

в) $\lambda_k = \sqrt{\dfrac{N_k}{N_{kg}}}$; $\quad$ г) $\lambda_k = \sqrt{\dfrac{N_k}{N_{k0}}}$.

3. Величина звездного плеча равна:

а) $\alpha_k = \sqrt{\dfrac{\sqrt{N_k \cdot N_{kg}} - N_{kg}}{2}}$; $\quad$ б) $\alpha_k = \sqrt{\dfrac{\sqrt{N_k \cdot N_{k0}} - N_{k0}}{2}}$;

в) $\alpha_k = \sqrt{\dfrac{\sqrt{N_k} - N_{k0}}{2}}$; $\quad$ г) $\alpha_k = \dfrac{\sqrt{N_k \cdot N_{k0}} - N_{k0}}{2}$.

4. ОЦКП для $k = 2$ имеет следующий вид.

а)

N_2	X_0	X_1	X_2	X_1X_2	$X_1^2 - 0.333$	$X_2^2 - 0.333$	Примечание
1	+	–	–	+	0.667	0.667	ПФЭ $N_{k0} = 4$
2	+	+	–	–	0.667	0.667	
3	+	–	+	–	0.667	0.667	
4	+	+	+	+	0.667	0.667	
5	+	–	0	0	0.667	– 0.667	«Звездные» опыты $2k = 4$
6	+	+	0	0	0.667	– 0.667	
7	+	0	–	0	– 0.667	0.667	
8	+	0	+	0	– 0.667	0.667	
9	+	0	0	0	– 0.667	– 0.667	Центр плана $N_0 = 1$
$\sum_{j=1}^{9} X_j$	9	0	0	0	0	0	
$\sum_{j=1}^{9} X_j^2$	9	6	6	4	2	2	

б)

N_2	X_0	X_1	X_2	X_1X_2	$X_1^2-0.667$	$X_2^2-0.667$	Примечание
1	+	–	–	+	0.333	0.333	ПФЭ $N_{k0}=4$
2	+	+	–	–	0.333	0.333	
3	+	–	+	–	0.333	0.333	
4	+	+	+	+	0.333	0.333	
5	+	–	0	0	0.333	– 0.667	«Звездные» опыты $2k=4$
6	+	+	0	0	0.333	– 0.667	
7	+	0	–	0	– 0.667	0.333	
8	+	0	+	0	– 0.667	0.333	
9	+	0	0	0	– 0.667	– 0.667	Центр плана $N_0=1$
$\sum_{j=1}^{9} X_j$	9	0	0	0	0	0	
$\sum_{j=1}^{9} X_j^2$	9	6	6	4	2	2	

в)

N_2	X_0	X_1	X_2	X_1X_2	$X_1^2-0.667$	$X_2^2-0.667$	Примечание
1	+	–	–	+	0.333	0.333	ПФЭ $N_{k0}=4$
2	+	+	–	–	0.333	0.333	
3	+	–	+	–	0.333	0.333	
4	+	+	+	+	0.333	0.333	
5	+	–	0	0	0.667	– 0.667	«Звездные» опыты $2k=4$
6	+	+	0	0	0.667	– 0.667	
7	+	0	–	0	– 0.667	0.667	
8	+	0	+	0	– 0.667	0.667	
9	+	0	0	0	– 0.667	– 0.667	Центр плана $N_0=1$
$\sum_{j=1}^{9} X_j$	9	0	0	0	0	0	
$\sum_{j=1}^{9} X_j^2$	9	6	6	4	2	2	

5. ОЦКП для k = 3 имеет следующий вид:

а)

N_3	X_0	X_1	X_2	X_3	X_1X_2	X_1X_2	X_1X_2	$X_1^2-0.7303$	$X_2^2-0.7303$	$X_3^2-0.7303$
1	+	–	–	–	+	+	+	0.2697	0.2697	0.2697
2	+	+	–	–	–	–	+	0.2697	0.2697	0.2697
3	+	–	+	–	–	+	–	0.2697	0.2697	0.2697
4	+	+	+	–	+	–	–	0.2697	0.2697	0.2697
5	+	–	–	+	+	–	–	0.2697	0.2697	0.2697
6	+	+	–	+	–	+	–	0.2697	0.2697	0.2697
7	+	–	+	+	–	–	+	0.2697	0.2697	0.2697
8	+	+	+	+	+	+	+	0.2697	0.2697	0.2697
9	+	–	0	0	0	0	0	0.7469	–0.7303	–0.7303
10	+	+	0	0	0	0	0	0.7469	–0.7303	–0.7303
11	+	0	–	0	0	0	0	–0.7303	0.7469	–0.7303
12	+	0	+	0	0	0	0	–0.7303	0.7469	–0.7303
13	+	0	0	–	0	0	0	–0.7303	–0.7303	0.7469
14	+	0	0	+	0	0	0	–0.7303	–0.7303	0.7469
15	+	0	0	0	0	0	0	–0.7303	–0.7303	–0.7303
$\sum_{j=1}^{15} X_j$	15	0	0	0	0	0	0	0	0	0
$\sum_{j=1}^{15} X_j^2$	15	10.9545	10.9545	10.9545	8	8	8	4.3644	4.3644	4.3644

б)

N_3	X_0	X_1	X_2	X_3	X_1X_2	X_1X_2	X_1X_2	$X_1^2-0.7303$	$X_2^2-0.7303$	$X_3^2-0.7303$
1	+	–	–	–	+	+	+	0.2697	0.2697	0.2697
2	+	+	–	–	–	–	+	0.2697	0.2697	0.2697
3	+	–	+	–	–	+	–	0.2697	0.2697	0.2697
4	+	+	+	–	+	–	–	0.2697	0.2697	0.2697
5	+	–	–	+	+	–	–	0.2697	0.2697	0.2697
6	+	+	–	+	–	+	–	0.2697	0.2697	0.2697
7	+	–	+	+	–	–	+	0.2697	0.2697	0.2697
8	+	+	+	+	+	+	+	0.2697	0.2697	0.2697
9	+	–1.2154	0	0	0	0	0	0.7469	–0.7303	–0.7303
10	+	+1.2154	0	0	0	0	0	0.7469	–0.7303	–0.7303
11	+	0	–1.2154	0	0	0	0	–0.7303	0.7469	–0.7303
12	+	0	+1.2154	0	0	0	0	–0.7303	0.7469	–0.7303
13	+	0	0	–1.2154	0	0	0	–0.7303	–0.7303	0.7469
14	+	0	0	+1.2154	0	0	0	–0.7303	–0.7303	0.7469
15	+	0	0	0	0	0	0	–0.7303	–0.7303	–0.7303
$\sum_{j=1}^{15} X_j$	15	0	0	0	0	0	0	0	0	0
$\sum_{j=1}^{15} X_j^2$	15	10.9545	10.9545	10.9545	8	8	8	4.3644	4.3644	4.3644

в)

N_3	X_0	X_1	X_2	X_3	X_1X_2	X_1X_2	X_1X_2	$X_1^2-0.7303$	$X_2^2-0.7303$	$X_3^2-0.7303$
1	+	–	–	–	+	+	+	0.2697	0.2697	0.2697
2	+	+	–	–	–	–	+	0.2697	0.2697	0.2697
3	+	–	+	–	–	+	–	0.2697	0.2697	0.2697
4	+	+	+	–	+	–	–	0.2697	0.2697	0.2697
5	+	–	–	+	+	–	–	0.2697	0.2697	0.2697
6	+	+	–	+	–	+	–	0.2697	0.2697	0.2697
7	+	–	+	+	–	–	+	0.2697	0.2697	0.2697
8	+	+	+	+	+	+	+	0.2697	0.2697	0.2697
9	+	–1.2154	0	0	0	0	0	0.7469	0.7303	0.7303
10	+	+1.2154	0	0	0	0	0	0.7469	0.7303	0.7303
11	+	0	–1.2154	0	0	0	0	0.7303	0.7469	0.7303
12	+	0	+1.2154	0	0	0	0	0.7303	0.7469	0.7303
13	+	0	0	–1.2154	0	0	0	0.7303	0.7303	0.7469
14	+	0	0	+1.2154	0	0	0	0.7303	0.7303	0.7469
15	+	0	0	0	0	0	0	0.7303	0.7303	0.7303
$\sum_{j=1}^{15} X_j$	15	0	0	0	0	0	0	0	0	0
$\sum_{j=1}^{15} X_j^2$	15	10.9545	10.9545	10.9545	8	8	8	4.3644	4.3644	4.3644

6. Для ортогонализированного уравнения регрессии второго порядка

$$Y = b_o^* + b_1^* X_1 + \ldots + b_k^* X_k + b_{12}^* X_1 X_2 + \ldots + b_{(k-1)k}^* X_{k-1} X_k +$$

$$+ b_{11}^* (X_1^2 - \lambda_k) + \ldots + b_{kk}^* (X_k^2 - \lambda_k)$$

коэффициент регрессии b_{rr}^* рассчитывается следующим образом:

$$\text{а) } b_{rr}^* = \frac{\sum_{j=1}^{N_k} \left(X_{rj}^2 - \lambda_k \right) \overline{Y}_j}{\sum_{j=1}^{N_k} \left(X_{rj}^2 - \lambda_k \right)^2};$$

$$\text{б) } b_{rr}^* = \frac{\sum_{j=1}^{N_k} \left(X_{rj}^2 - \lambda_k \right) \overline{Y}_j}{\sum_{j=1}^{N_k} \left(X_{rj}^2 - \lambda_k \right)};$$

$$\text{в) } b_{rr}^* = \frac{\sum_{j=1}^{N_k} X_{rj}^2 \overline{Y}_j}{\sum_{j=1}^{N_k} \left(X_{rj}^2 - \lambda_k \right)^2}.$$

7. Дисперсии значимости коэффициентов b_{rs} уравнения регрессии рассчитываются следующим образом:

$$\text{а) } S(b_{rs}) = \frac{S_{\text{воспр}}}{n N_{k0}};$$

$$\text{б) } S^2(b_{rs}) = \frac{S_{\text{воспр}}^2}{n N_{k0}};$$

$$\text{в) } S^2(b_{rs}) = \frac{S_{\text{воспр}}^2}{n \sum_{j=1,\, r<s}^{N_{k0}} (X_{rj} X_{sj})^2}.$$

8. Между коэффициентами b_0 и b_0^* уравнений регрессии второго порядка в разных формах существует следующая связь:

$$\text{а) } b_0 = b_0^* - \lambda_k \left(b_{11}^* + \ldots + b_{kk}^* \right);$$

$$\text{б) } b_0 = b_0^* - \left(b_{11}^* + \ldots + b_{kk}^* \right);$$

$$\text{в) } b_0 = b_0^* - \lambda_k^2 \left(b_{11}^* + \ldots + b_{kk}^* \right).$$

9. Найдите правильные формулы:

$$\text{а) } B_0 = \sum_{j=1}^{N_k} X_{0j}^2 = N_k + 1;$$

$$\text{б) } B_r = \sum_{j=1}^{N_k} X_{rj}^2 = \sqrt{N_{k0} N_k};$$

$$\text{в) } B_{rs} = \sum_{j=1}^{N_k} (X_{rj} X_{sj})^2 = \sqrt{N_{k0}};$$

$$\text{г) } B_{rr} = 2\alpha_k^2;$$

$$\text{д) } B_{00} = N_k \Big/ (1 + k N_{kg} \Big/ 2\alpha_k^4).$$

10. Обычное уравнение регрессии второго порядка $Y = b_0 + b_1 X_1 + b_2 X_2 + b_{12} X_1 X_2 + b_{11} X_1^2 + b_{22} X_2^2$ имеет максимум, если

а) $b_{11} > 0, \quad b_{22} > 0$, б) $b_{11} < 0, \quad b_{22} < 0$,

в) $b_{11} = 0, \quad b_{22} = 0$, г) $b_{11} > 0, \quad b_{22} < 0$.

Вопросы для самопроверки

1. Напишите уравнение регрессии второго порядка для 4-х факторов в обычном и ортогонализированном видах.

2. Чему равны величины звездного плеча и ортогонализирующего коэффициента для ОЦКП при 4-х факторах?

3. Составьте план ОЦКП для 4-х факторов.

4. Приведите формулу расчета величины звездного плеча в общем случае ОЦКП для k факторов.

5. Приведите формулу расчета величины ортогонализирующего коэффициента в общем случае ОЦКП для k факторов.

6. Напишите формулы для расчетов суммы квадратов факторов X_{0j}^2, X_{rj}^2, $(X_{rj}X_{sj})^2$, $(X_{rj}^2-\lambda_k)^2$ в ОЦКП.

7. Напишите формулы расчета коэффициентов b_0^*, b_r^*, b_{rs}^*, b_{rr}^* для уравнения регрессии второго порядка в ортогонализированном виде и b_0, b_r, b_{rs}, b_{rr} для уравнения регрессии второго порядка в обычном виде.

8 Напишите формулы расчета дисперсий значимости коэффициентов b_0^*, b_r^*, b_{rs}^*, b_{rr}^* для уравнения регрессии второго порядка в ортогонализированном виде и b_0, b_r, b_{rs}, b_{rr} для уравнения регрессии второго порядка в обычном виде.

9. Сформулируйте алгоритм проверки на адекватность уравнения регрессии второго порядка.

10. В каких случаях объект исследований, описываемый уравнением регрессии второго порядка, имеет максимум (минимум)? Как найти оптимальные значения факторов?

11. Выведите формулу для расчета предельной абсолютной погрешности параметра Y(X) по двухфакторному уравнению регрессии второго порядка.

Задачи для самостоятельного решения

Уровень 1

1. Доказать, что $B_0=\sum\limits_{j=1}^{N_k}X_{0j}^2=N_k$.

2. Доказать, что $B_r=\sum\limits_{j=1}^{N_k}X_{rj}^2=\sqrt{N_{k0}N_k}$.

3. Доказать, что $B_{rs}=\sum\limits_{j=1}^{N_k}(X_{rj}X_{sj})^2=N_{k0}$.

Уровень 2

4. Доказать, что $B_{rr}=2\alpha_k^4$.

5. Доказать, что $B_{00}=\dfrac{2\alpha_k^4N_k}{2\alpha_k^4+kN_{kg}}$.

Уровень 3

6. Доказать, что

$$\lambda_k=\sqrt{\frac{N_{kg}}{N_k}},$$

$$\alpha_k=\sqrt{\frac{\sqrt{N_k\cdot N_{kg}}-N_{kg}}{2}}.$$

7. Построить адекватное уравнение регрессии второго порядка, описывающее урожайность пшеницы от содержания посевного материала X_1 ($x_{10}=3.5$ ц/га, $\Delta x=1$ ц/га) и количества неорганического удобрения X_2 ($x_{20}=1.5$ ц/га, $\Delta x=1$ ц/га). Результаты эксперимента по ОЦКП приведены в таблице. Найти максимальное значение урожайности пшеницы и оптимальные значения факторов.

N	X_{1j}	X_{2j}	Y_{1j}	Y_{2j}	Y_{3j}
1	–1	–1	42.3	43.4	43.1
2	1	–1	35.0	35.4	35.5
3	–1	1	57.3	57.5	57.7
4	1	1	50.1	50.3	50.5
5	–1	0	57.0	57.2	57.3
6	1	0	50.2	50.5	50.7
7	0	–1	42.6	42.6	42.8
8	0	1	57.0	57.2	57.4
9	0	0	57.0	57.4	57.6

Ответ:

$Y=57.29-3.59X_1+7.35X_2-3.46X_1^2-7.34X_2^2$, $S_{\text{воспр}}^2=0.0785$, $S_{\text{ад}}^2=0.182$, $F_э=2.318$, $F_{4,18,\ 0.95}=2.928$. $Y_{\max}=60.06$ ц/га при $x_1=2.98$ ц/га и $x_2=2.00$ ц/га.

Модуль R
ИТОГОВОЕ ЗАНЯТИЕ
(тема 11)

В результате изучения модуля студент должен:

– **знать** цели и задачи дисциплины, базовые проблемы курса;

– **понимать** перспективы применения полученных знаний для решения актуальных задач в области АПК;

– **формировать** стремление к овладению новыми знаниями, расширению кругозора.

R.1. Перечень и содержание лекций

Лекция 1. Стратегия прикладных научных исследований: рынок – наука – производство – рынок. Рыночные критерии оптимизации. Цель прикладных исследований – разработка изделий с требуемым комплексом эксплуатационных параметров, а не материалов с заданным комплексом физико-механических свойств. Логика исследований и внедрения: рыночные требования → эксплуатационные параметры изделия → физико-механические свойства материалов → макро- и микроструктурные параметры материалов → конструкционные размеры изделий → модель эксплуатации → целевая функция → оптимальное изделие → технологические параметры процесса получения изделий → детерминистическая и стохастическая модели технологии серийного производства → технико-экономические параметры конкретного производства → целевая функция → оптимальная технология производства изделий с повышенной стабильностью свойств при минимально возможной стоимости. Критерий работоспособности изделия и программа управления качеством.

R.2. Материалы для самостоятельной работы и практических занятий

1. Сформулируйте несколько предложений, которые, на ваш взгляд, будут способствовать повышению эффективности научных исследований.

2. Прокомментируйте проблему: наука – технологии – инвестиции.

3. Что, на ваш взгляд, надо сделать, чтобы повысить уровень отдачи прикладных исследований народному хозяйству?

ЗАКЛЮЧЕНИЕ

В данном учебно-методическом комплексе изложены теория и практика для ограниченного класса задач: статистический анализ стохастических систем, описываемых непрерывными количественными параметрами, подчиняющимися нормальному закону распределения, при воздействии на них количественных детерминированных факторов. Моделирование таких систем базируется на аппарате регрессионного анализа, основы которого подробно изложены в данном издании применительно к однофакторным и многофакторным моделям первого и второго порядка. Регрессионные модели более высоких порядков не содержат принципиальных отличий в методах расчета и статистического анализа параметров, поэтому при необходимости вполне могут быть освоены самостоятельно на базе данного курса.

Существует большое количество задач других классов, анализ которых базируется на широком спектре различных статистических методов. В их числе дисперсионный анализ (количественный параметр – качественные факторы), ковариационный анализ (количественный параметр – количественные и качественные факторы), дискриминантный анализ (качественный параметр – количественные факторы), модели распознавания образов (качественные параметры – качественные факторы) и многие другие. Рассмотрение этих методов лежит за рамками начального курса основ научных исследований и моделирования. В литературных источниках можно найти работы, рассматривающие подобные задачи. Необходимо принять во внимание и то, что широкое использование средств вычислительной техники в современных научных исследованиях привело к разработке большого числа специализированных программных пакетов математической статистики, начиная от пакета простейшего статистического анализа в составе MS Excel (Analysis ToolPak) до таких мощных программных комплексов, как STATISTICA (StatSoft, Inc. www.statsoft.com) и SPSS Statistics (SPSS Inc. www.spss.com). Применение подобных программных пакетов избавляет исследователя от рутинных вычислительных процедур, но требует знания теоретических и методологических основ применения тех или иных методов статистического анализа.

РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА

(приведена в хронологическом порядке)

Основная

1. Адлер, Ю.П. Планирование эксперимента при поиске оптимальных условий / Ю.П. Адлер, Е.В. Маркова, Ю.В. Грановский. – Москва : Наука, **1971. -** 279 с.
2. Саутин, С.Н. Планирование эксперимента в химии и химической технологии / С.Н. Саутин. - Л. : Химия, **1975. –** 48 с.
3. Бохан, Н.И. Основы научных исследований и обработки экспериментальных данных / Н.И. Бохан, В.К. Бензарь. – Горки : БСХА, **1980**. - 104 с.
4. Бохан, Н.И. Планирование экспериментов в исследованиях по механизации и автоматизации сельскохозяйственного производства / Н.И. Бохан, А.М. Дмитриев, И.С. Нагорский. – Горки : БСХА, **1986**. – 80 с.
5. Нагорский, И.С. Методические указания для выполнения курсовой работы по дисциплине «Основы научных исследований» / И.С. Нагорский. **–** Минск : БГАТУ**, 2004. –** 28 с.
6. Нагорский, И.С. Основы научных исследований: пособие по изучению дисциплины: в 4-х ч. Ч. 1, 2 / И.С. Нагорский. – Минск : БГАТУ, **2006. –** 132 с.
7. Нагорский, И.С. Основы научных исследований: пособие по изучению дисциплины: в 4-х ч. Ч. 3, 4 / И.С. Нагорский, Б.В. Ловкис, Ю.Т. Антонишин. – Минск : БГАТУ, **2008**. – 108 с.

Дополнительная

1. Налимов, В.В. Статистические методы планирования экстремальных экспериментов / В.В. Налимов, Н.А. Чернова. – Москва : Наука, **1965**. – 340 с.
2. Шторм, Р. Теория вероятностей. Математическая статистика. Статистический контроль качества / Р. Шторм; пер. с нем. - Москва : Мир, **1970**. - 368 с.
3. Кацев, П.Г. Статистические методы исследования режущего инструмента / П.Г. Кацев. – Москва : Машиностроение, **1974. –** 240 с.

4. Браунли К.А. Статистическая теория и методология в науке и технике / К.А. Браунли; пер. с англ. – Москва : Наука, **1977**. - 408 с.

8. Большев, Л.Н. Таблицы математической статистики / Л.Н. Большев, Н.В. Смирнов. – Москва : Наука, **1983**. - 416 с.

5. Основы научных исследований: учебник для технических вузов / под ред. В.И. Крутого, В.В. Попова. – Москва : Высшая школа, **1989**. – 400 с.

6. Хайлис, Г.А. Исследования сельскохозяйственной техники и обработка опытных данных / Г.А. Хайлис, М.М. Ковалев. – Москва : Колос, **1994**. – 164 с.

7. Протасевич, Г.Ф. Учебно-методическое пособие по дисциплине «Основы научных исследований» и «Организация эксперимента» / Г.Ф. Протасевич. – Минск : БНТУ, **2002**. - 68с.

8. Бабицкий, В.В. Планирование эксперимента: учебно-методическое пособие по проведению инженерных экспериментов и обработке. полученных результатов / В.В. Бабицкий, Я.Н. Ковалев, В.Д. Якимович. – Минск : БНТУ, **2003**. - 48с.

9. Леонов, А.Н. Пористые проницаемые материалы: теория проектирования изделий и технологий / А.Н. Леонов, М.М. Дечко, В.К. Шелег. – Минск : Тонпик, **2003**. - 218 с.

10. Беларусь: наука, технологии, инновации. – Минск : ГКНТ, **2005**. - 26 с.

11. О состоянии и перспективах развития науки в Республике Беларусь по итогам 2006 года: Аналитический доклад / А.Н. Коршунов [и др.] – Минск : БелИСА, **2007**. – 316 с.

12. Положение «О научно-исследовательской работе студентов БГАТУ». - Минск : БГАТУ, **2007**. - 14 с.