Н.И. Сидняев

Статистический анализ и теория планирования эксперимента

Учебное пособие

BAUMANPRESS
Москва
ИЗДАТЕЛЬСТВО
МГТУ им. Н. Э. Баумана
2017

УДК 519.2
ББК 32.81я73
С34

Издание доступно в электронном виде на портале *ebooks.bmstu.ru*
по адресу: http://ebooks.bmstu.ru/catalog/109/book1687.html

Факультет «Фундаментальные науки»
Кафедра «Высшая математика»

Рекомендовано Редакционно-издательским советом
МГТУ им. Н.Э. Баумана в качестве учебного пособия

Рецензент
д-р физ.-мат. наук *А.А. Гурченков*

Сидняев, Н. И.

С34 Статистический анализ и теория планирования эксперимента : учебное пособие / Н. И. Сидняев. — Москва : Издательство МГТУ им. Н. Э. Баумана, 2017. — 195, [5] с. : ил.

ISBN 978-5-7038-4707-7

Изложены краткие теоретические сведения по курсу «Теория планирования эксперимента». Представлено введение в статистический анализ и теорию планирования эксперимента. Основные понятия иллюстрируются примерами практического содержания, рассмотренными с позиций регрессионного анализа. Издание носит справочный характер и поможет студентам старших курсов овладеть методами теории планирования эксперимента, которые широко используются при решении прикладных задач.

Для студентов 4–6-го курсов инженерных специальностей технических университетов.

УДК 519.2
ББК 32.81я73

ISBN 978-5-7038-4707-7

Предисловие

В учебном пособии, состоящем из четырех глав, изложены основы теории планирования эксперимента в соответствии с программой этого курса для технических и экономических специальностей технических вузов.

Материал учебного пособия представлен в строгой, но доступной пониманию форме, примеры имеют практическую направленность.

Содержание учебного пособия соответствует курсу лекций по дисциплине «Теория планирования эксперимента», который автор читает в МГТУ им. Н.Э. Баумана для студентов специальности «Прикладная математика». Учебное пособие предназначено для студентов старших курсов и может быть полезно преподавателям, аспирантам, а также специалистам в области статистического моделирования.

В первой главе рассмотрены статистические методы обработки экспериментальных данных, предназначенные для построения эмпирических зависимостей. Это дает возможность перейти к рассмотрению принципиально нового подхода к экспериментированию, который состоит в том, что в каждом опыте по специальному плану одновременно варьируют все независимые переменные (факторы). Последовательно объяснены основные понятия математической статистики, а именно: эмпирические функции распределения, плотность распределения, гистограмма, выборка, порядковая статистика. Приведены асимптотические свойства оценок, описаны способы доверительного оценивания, представлены основные распределения математической статистики (χ^2-распределение, распределение Стьюдента, распределение Фишера). Рассмотрено доверительное оценивание параметров биномиального, нормального и экспоненциального распределений. Введены понятия статистической гипотезы, критерия проверки гипотезы, статистики критерия. Дана общая логическая схема построения критерия проверки гипотезы.

Подробно изложены элементы регрессивного анализа (простая регрессия, парная линейная регрессия, ортогональная регрессия). Рассмотрены множественная линейная регрессия, статистические выводы о параметрах регрессии, оценка множественного коэффициента корреляции.

Во второй главе введены основные понятия теории планирования эксперимента, обсуждаются критерии планирования. Проанализированы этап предпланирования эксперимента, особенности выбора факторов, определения области проведения эксперимента, выбора базовой точки и интервалов варьирования. Приведены примеры полного факторного эксперимента типа 2^m, линейных моделей второго порядка.

Рассмотрена возможность сокращения числа опытов путем проведения дробного факторного эксперимента. В этом случае из всех опытов, необходимых при полном факторном эксперименте, некоторые сочетания исключаются и опыты при этих сочетаниях не проводятся. Отмечено, что опыты, в которых все факторы находятся на основном уровне (+1, 0, –1), исключать не следует. Приведено описание методов выделения существенных факторов, которые необходимо учитывать при построении математических моделей. Рассмотрены проблемы планирования эксперимента и расчета оценок параметров, входящих в модели нелинейно. Применение теории планирования эксперимента в таких задачах основано на линеаризации модели по неизвестным параметрам при некоторых начальных значениях этих параметров. В связи с тем что в дробных репликах часть взаимодействий заменена новыми факторами, найденные коэффициенты уравнения регрессии рассматриваются как совместные оценки линейных эффектов и эффектов взаимодействия.

В композиционных планах за основу принимают двухуровневый полный факторный эксперимент и к нему добавляют эксперименты, проводимые на других уровнях: как в центре, так и при различных значениях $\pm\alpha$ (расстояние от ребер гиперкуба, или длина «звездного плеча»). При большом числе факторов ($k > 3$) проведение полного факторного эксперимента связано с большим числом опытов, значительно превосходящим число коэффициентов линейной модели. Если при получении модели можно ограничиться линейным приближением, т. е. получить адекватную модель в виде полинома, то число опытов можно резко сократить в результате использования дробного факторного эксперимента.

В третьей главе представлены композиционные планы Бокса, ротатабельные, центральные композиционные планы, планы типа B_m, D-оптимальные планы, планы Кифера и Коно на отрезке $\Omega = (-1, 1)$. Центральные композиционные планы второго порядка рассматриваются в совокупности с ортогональными и ротатабельными планами второго порядка. Изложены основные методы получения математического описания на относительно линейных участках поверхности отклика и в области главного экстремума. Введено понятие информационного профиля плана.

Объяснено понятие разрешающей способности как числа несмешанных линейных эффектов в дробной реплике. Прямая оценка разрешающей способности дробной реплики затруднена, поэтому дробные реплики задают с помощью генерирующих соотношений. Генерирующее соотношение показывает, какое из взаимодействий принято незначимым и заменено новым фактором. Приведен пример, когда план типа 2^{3-1} может быть представлен двумя полурепликами, которые задаются генерирующими соотношениями. Описаны полные и дробные факторные планы, а также композиционные ортогональные и ротатабельные планы эксперимента для квадратичных моделей.

На примерах показано, что весьма сложной задачей является выбор некоторого универсального критерия оптимальности планирования, позволяющего оценивать планы со всех возможных точек зрения. Показано, что ротатабельные планы являются оптимальными и в более широком смысле: к ним приходится обращаться и тогда, когда надо минимизировать систематические ошибки, связанные с неадекватностью представления результатов исследования полиномами второго порядка. Построение ротатабельных планов второго порядка — сложная математическая задача.

Приведены данные, необходимые для построения матриц центрального композиционного ротатабельного планирования. Показано, как планы разбиваются на ортогональные блоки. Разобраны примеры применения ротатабельного планирования. Предложено несколько способов построения ротатабельных планов второго порядка. Отмечено, что ортогональные планы первого порядка являются одновременно ротатабельными планами. Дана геометрическая интерпретация ортогональных планов первого порядка. Рассмотрены примеры различных матриц D, задающих координаты экспериментальных точек в факторном пространстве для ортогонального плана первого порядка с минимальным числом экспериментов N.

Четвертая глава посвящена анализу почти стационарной области. Рассмотрено ортогональное и ротатабельное планирование второго порядка, а также ротатабельное планирование с учетом эффекта неадекватности модели.

В приложении приведены варианты домашнего задания.

В соответствии с целями конкретной общеобразовательной программы высшего профессонального образования и задачами профессиональной деятельности студент в дополнение к компетенциям, соответствующим квалификации магистра, должен обладать следующими компетенциями.

Общекультурные компетенции: умение находить, анализировать и обрабатывать информацию, в том числе относящуюся к новым областям знания, непосредственно не связанным с профессиональной деятельностью.

Профессиональные, в том числе *общенаучные*, компетенции: способность самостоятельно приобретать, осмысливать, структурировать и использовать в профессиональной деятельности новые знания, применять современные образовательные и информационные технологии, развивать свои инновационные способности, расширять и углублять свое научное мировоззрение; способность самостоятельно формулировать цели исследований, устанавливать последовательность решения задач с использованием теории планирования эксперимента, внедрять результаты научно-исследовательской работы в практику.

Инструментальные компетенции: владение современными средствами вычислительной техники, системным и прикладным программным обеспечением, позволяющее выбирать и творчески использовать их для решения научных и практических задач в профессиональной деятельности.

Профессионально-специализированные компетенции: владение методологией математического моделирования технических систем и технологических процессов, процессов управления ими, а также методами статистического анализа и теории массового обслуживания, математической теории надежности технических систем и технологических процессов на основе глубоких знаний фундаментальных математических и естественно-научных дисциплин, информационных технологий; способность самостоятельно разрабатывать математические модели и методы их качественного и количественного анализа с использованием теории планирования эксперимента применительно к техническим объектам, характерным для современного машино- и приборостроения.

После освоения дисциплины студент должен *знать*: методы теории планирования эксперимента, основы математического моделирования и его применения при исследовании физических, технических, биологических, экологических процессов; теоретический материал (исходные факты, определения, аксиомы, основные утверждения, теоремы, свойства математических объектов); методики решения задач теоретического и прикладного характера.

Студент должен *уметь*: анализировать и решать научные, научно-исследовательские и инженерно-экономические задачи в области теории планирования эксперимента и ее приложений; применять полученные знания при математическом моделировании явлений живой и неживой природы, экономических и социальных процессов общественной жизни, решении разнообразных задач; использовать информационные технологии в проектно-конструкторской, управленческой и финансовой деятельности.

Студент должен *иметь навыки*: постановки задач (описание физической задачи, выделение существенных факторов, использование физической модели и ограничение области ее применимости, математическое моделирование задачи); вывода расчетных формул (переход от уравнений математической модели к расчетным формулам; грамотное использование основных расчетных формул теории планирования эксперимента, если вывод расчетных формул является громоздким).

Введение

Практическая полезность научных исследований в значительной степени зависит от методов их проведения и формы представления результатов. Применение эффективной технологии исследований позволяет существенно сократить фазу внедрения, что приводит к экономии времени и средств [1–3].

Обеспечить с помощью традиционных методов исследования требуемые темпы развития ни в фундаментальных или прикладных исследованиях, ни при модернизации или проектировании производственных установок не удается [4–6]. В связи с этим в науке, технике, производстве при решении разнообразных задач во все бóльших масштабах применяют новые эффективные методы исследования. При этом особое внимание уделяют моделям процессов и способам их построения [7].

Для эффективного анализа механизмов явлений и управления производственными процессами необходимо выявить взаимосвязи между факторами, определяющими ход процесса, и представить их в количественной форме — в виде математической модели [7–9]. Математическая модель является математическим отображением наиболее существенных сторон процесса.

В зависимости от источника информации, используемого при построении математической модели, различают физико-химические модели (называемые также иногда аналитическими или теоретическими) и статистические (или эмпирические) модели [1, 7]. В первом случае за основу берут физико-химические закономерности моделируемых процессов, например, в виде уравнений баланса или кинетических уравнений для превращений вещества. Построение теоретических моделей сопряжено с проведением обширных и длительных исследований, поскольку при этом необходимо выяснить природу микропроцессов, протекающих в объекте, и описать их математически [10–12]. Как правило, модели процессов представляют в виде сложных систем уравнений (системы алгебраических, обыкновенных дифференциальных уравнений или уравнений в частных производных). Эти системы уравнений позволяют очень точно описать процессы, протекающие в объекте, и допускают экстраполяцию в точки факторного пространства, где непосредственное наблюдение процессов невозможно. Статистические модели получают в результате статистической обработки экспериментальных данных, собранных на исследуемом объекте. Структура статистической модели может быть выбрана относительно произвольно [13]. Соответствие модели объекту ограничивается лишь количественно.

Статистические модели имеют относительно простую структуру и очень часто представлены в виде полиномов. Область их применения отвечает лишь ближайшей окрестности рабочих точек, в которой проводятся эксперименты. Во многих случаях такие модели можно построить при сравнительно небольших затратах времени и средств [14–16].

В результате множества экспериментов должна быть построена подходящая модель, выявлены возможные ошибки. Экспериментатор, применяя свои технические знания, может обнаружить дефекты модели, скорректировать их и построить новую модель, которая также будет подвергнута испытаниям [7]. Функциональная форма, полученная на основе теоретического представления, обычно нелинейна в неизвестных параметрах. Эмпирический подход иногда применяют просто для того, чтобы избежать осложнений, связанных с нелинейными уравнениями. Решение нелинейных уравнений — область применения численных методов.

Эксперимент занимает главенствующее место среди способов получения информации о внутренних взаимосвязях явлений в природе и технике. Он является отправной точкой и критерием большинства знаний. Эксперименты и наблюдения стали основой для открытия многих известных законов природы и проверки теоретических гипотез [4]. По мере усложнения исследуемых процессов и явлений возрастают затраты на аппаратуру и проведение эксперимента.

В последние годы была создана последовательная и достаточно строгая теория регрессионного анализа, базирующаяся на современных теоретико-вероятностных представлениях [12]. Эта теория позволила значительно глубже понять и оценить результаты, получаемые методом наименьших квадратов.

Опыт показал, что классический регрессионный анализ, несмотря на хорошо разработанную теорию, не нашел широкого применения для решения экстремальных задач в физике, химии и металлургии [9]. При решении подобного рода задач приходится иметь дело с очень большим числом независимых переменных. В этом случае метод становится крайне громоздким и возникают практически непреодолимые трудности, связанные, с одной стороны, с необходимостью ставить очень большое число экспериментов, с другой стороны — с интерпретацией уравнения регрессии (все коэффициенты регрессии оказываются корреляционно связанными между собой).

Отметим следующие новые возможности, которые открывает перед исследователями теория планирования эксперимента:

1. Статистическое представление об эксперименте образует основу для исследования сложных объектов и систем [3]. Эти объекты характеризуются большим числом факторов, воздействующих на результаты эксперимента. При классическом подходе к экспериментам влияние совокупности факторов на результаты эксперимента исследовалось при условии, что изменяется только один из факторов и фиксируются «значения» всех остальных. В сложных системах, где большое число воздействий невозможно контролировать (или управлять ими), это условие не выполняется.

Статистическая концепция позволяет учесть влияние неконтролируемых факторов иным образом [16]. Воздействие этих факторов рассматривается как дополнительный стохастический шум, наложенный на истинные результаты экспериментов. Для того чтобы воздействие сделать случайным, применяют специальные методы. Благодаря этому удается надежно отделить факторы, интересующие экспериментатора, от шумового фона, обусловленного неконтролируемыми воздействиями.

2. Математическая статистика предоставляет в распоряжение экспериментатора методы анализа данных и принятия решений относительно исследуемого объекта на основании обработанных результатов эксперимента. Эти методы дают возможность учесть стохастический характер результатов и основаны на статистической проверке гипотез.

Планирование эксперимента — это новый подход к исследованию, при котором активная роль отводится математическим методам [7–9]. Основываясь на априорных сведениях об изучаемом процессе, исследователь выбирает некоторую оптимальную стратегию для управления экспериментом. Процесс исследования обычно разбивается на этапы. После каждого этапа исследователь получает новую информацию, позволяющую ему изменять стратегию исследования.

На математическом языке задача планирования эксперимента формулируется следующим образом: нужно выбрать в некотором смысле оптимальное расположение точек в факторном пространстве, для того чтобы получить какое-то представление о поверхности отклика. А выбор критерия оптимальности в значительной степени произволен.

Задачи, для решения которых может использоваться планирование эксперимента, чрезвычайно разнообразны. Поиск оптимальных условий, построение интерполяционных формул, выбор существенных факторов, оценка и уточнение констант теоретических моделей (например, кинетических), выбор из некоторого множества гипотез о механизме явлений наиболее приемлемых, исследование диаграмм состав–свойство — вот примеры задач, при решении которых применяется планирование эксперимента [10]. Можно сказать, что там, где есть эксперимент, присутствует и наука о его проведении — теория планирования эксперимента.

Поиск оптимальных условий — один из наиболее распространенных типов научно-технических задач. Эти задачи возникают в тот момент, когда установлена возможность проведения процесса и необходимо найти наилучшие (в некотором смысле оптимальные) условия его реализации.

Например, у химика возникла гипотеза о том, что при взаимодействии двух веществ должен получаться некоторый интересующий его продукт. Требуется подобрать концентрации реагирующих веществ, температуру, давление, время реакции и другие факторы таким образом, чтобы сделать выход продукта возможно более близким к 100 %. В этом примере находят условия проведения процесса, оптимальные в смысле максимизации выхода требуемого продукта. Но это далеко не единственно возможная постановка задачи. Найденные условия могли оказаться другими, если бы ставилась, например, задача минимизировать себестоимость продукта или количество вредных примесей. Следует подчеркнуть, что всегда необходимо четко формулировать, в каком смысле условия должны быть оптимальными. Этим определяется выбор цели исследования. Точная формулировка цели в значительной мере определяет успех исследования.

Задачи, сформулированные аналогичным образом, называются *задачами оптимизации*. Процесс их решения называется *процессом оптимизации* (или просто *оптимизацией*). Выбор оптимального состава многокомпонентных смесей или сплавов, повышение производительности действующих установок, повышение качества продукции, снижение затрат на ее получение — вот примеры задач оптимизации. Эксперимент, который ставится для решения задач оптимизации, называется *экстремальным*.

1. ЭЛЕМЕНТЫ МНОГОМЕРНОГО СТАТИСТИЧЕСКОГО АНАЛИЗА

1.1. Формы связи между случайными выборками

Результаты эксперимента обрабатывают, чтобы получить исходные данные для любых последующих расчетов, в том числе и вероятностных. Методы анализа результатов эксперимента и определения по ним вероятностей событий и характеристик случайных величин предоставляет математическая статистика — обширный раздел современной теории вероятностей [2–4].

Исходным материалом для применения статистических методов являются экспериментальные или статистические данные, под которыми понимают сведения о числе объектов, обладающих теми или иными признаками. Например, статистическими данными являются: $X_1, X_2, ..., X_n$ — отклонения размеров n однотипных деталей от номинального размера (X_i — отклонение размера i-й детали); $X_1, X_2, ..., X_i, ..., X_n$ — значения предельной прочности n образцов некоторого материала; $\tau_1, \tau_2, ..., \tau_i, ..., \tau_n$ — продолжительность функционирования до отказа (долговечность) n образцов прибора и т. д.

Перечисленные данные являются числовыми характеристиками массовых случайных явлений (сортности деталей, прочности материала, долговечности приборов и т. д.), поэтому предметом математической статистики служат случайные явления, а ее основная задача — количественный и качественный анализ этих явлений.

Напомним кратко основные исходные понятия математической статистики и те задачи, которые она позволяет решить. Совокупность всех возможных значений изучаемого признака статистических данных (другими словами, случайной величины X) называется *генеральной совокупностью*, конечной или бесконечной. Те значения признака X, которые зафиксированы в опыте на n образцах, случайным образом отобранных из генеральной совокупности, называют *выборкой* объемом n и обозначают $x_1, x_2, ..., x_n$ (строчными буквами) [1, 2].

При многократном повторении опыта (хотя бы мысленном!) исследователь будет получать различные конкретные выборки. В математической статистике постулируется, что множество всех конкретных выборок образует n-мерный случайный вектор $X_1, X_2, ..., X_n$, называемый *случайной выборкой* из генеральной совокупности, где, во-первых, все компоненты X_i независимы и, во-вторых, каждый компонент X_i есть случайная величина с тем же законом распределения (вообще говоря, неизвестным), которому подчиняется генеральная совокупность X.

2. ЭЛЕМЕНТЫ ТЕОРИИ ПЛАНИРОВАНИЯ ЭКСПЕРИМЕНТА

2.1. Наблюдение и эксперимент как основы математического моделирования

Из множества задач, которые экспериментатору приходится решать при исследовании интересующего его объекта или процесса, выделим следующие две задачи, встречающиеся на практике, пожалуй, наиболее часто:

1) построение математической модели объекта, представляющей собой аналитическую зависимость между выходной переменной (откликом) и набором входных переменных (факторов);

2) поиск оптимальных условий доведения объекта (протекания процесса) до рабочего состояния, т. е. поиск таких значений факторов, при которых отклик (или некоторый функционал от него) достигает экстремума.

В первой главе рассматривались вопросы построения модели по результатам пассивного эксперимента. При этом вторая задача не обсуждалась, так как возможность ее решения связана с целенаправленным поиском точек проведения эксперимента в пространстве факторов. Во второй главе обе отмеченные задачи рассмотрены в предположении, что экспериментатор имеет возможность целенаправленно влиять на условия проведения эксперимента (нужным образом устанавливать значения контролируемых переменных и число опытов), или, иными словами, может планировать эксперимент.

Планирование эксперимента — это процедура выбора числа и условий проведения опытов, необходимых и достаточных для решения поставленной задачи с требуемой точностью. При этом существенно следующее:

1) стремление к минимизации общего числа опытов;

2) одновременное варьирование всеми переменными, определяющими процесс, по специальным правилам — алгоритмам;

3) использование математического аппарата, формализующего многие действия экспериментатора;

4) выбор четкой стратегии, позволяющей принимать обоснованные решения после каждой серии экспериментов.

Цель планирования эксперимента — нахождение таких условий и правил проведения опытов, при которых удается получить надежную и достоверную информацию об объекте с наименьшей затратой труда, а также представить эту информацию в компактной и удобной форме с количественной оценкой точности.

Теория планирования эксперимента охватывает практически все встречающиеся на практике варианты исследования объектов. Исследуемый

3. АНАЛИЗ ФАКТОРНЫХ ЭКСПЕРИМЕНТОВ

3.1. Свойства факторных экспериментов

Пусть функция отклика имеет вид

$$\eta = \beta_0 + \sum_{1 \le i \le k} \beta_i x_i + \sum_{1 \le i < j \le k} \beta_{ij} x_i x_j + \sum_{1 \le i < j < l \le k} \beta_{ijl} x_i x_j x_l + \\ + \ldots + \beta_{12\ldots k} x_1 x_2 \ldots x_k. \tag{3.1}$$

Выразим ее в общем виде в предположении, что некоторые взаимодействия могут отсутствовать:

$$\eta = f(x_1, x_2, \ldots, x_k;\ \beta_0, \beta_1, \ldots, \beta_p). \tag{3.2}$$

Функция отклика линейна по параметрам, при этом $p \le 2^k - 1$, где 2^k — число неизвестных параметров в уравнении (3.1). Пусть $\bar{D} = (x_{iu})$, $i = 1, 2, \ldots, k$; $u = 1, 2, \ldots, N$, — матрица полного или дробного факторного плана, а $\bar{X} = (x_{ju})$, $j = 1, 2, \ldots, k$; $u = 1, 2, \ldots, N$, — соответствующая ей функция отклика (3.2) — матрица независимых переменных.

Сформулируем основные свойства факторных планов, когда $\operatorname{rank} X \le \le p + 1$. Рассмотрим два случая.

1. Ранг матрицы X равен $r = p + 1$, т. е. модель наблюдений является моделью полного ранга. Из условия $\operatorname{rank} X \le p + 1 \Rightarrow r \le N$ матрица независимых переменных обладает следующими свойствами:

$$\sum_{u=1}^{N} x_{ju} = 0,\ j = 1, 2, \ldots, p; \tag{3.3}$$

$$\sum_{u=1}^{N} x_{ju}^2 = N,\ j = 1, 2, \ldots, p; \tag{3.4}$$

$$\sum_{u=1}^{N} x_{ju} x_{iu} = 0,\ i,\ j = 1, 2, \ldots, p;\ i \ne j. \tag{3.5}$$

Поскольку планирование является ортогональными, матрица $(X'X)^{-1}$ диагональна и ее диагональные элементы $c_{jj} = 1/N$, $j = 1, 2, \ldots, p$, поэтому (см. (2.29), (2.30))

4. ОПИСАНИЕ КВАЗИСТАЦИОНАРНОЙ ОБЛАСТИ

4.1. Ортогональное планирование второго порядка

Движение по методу крутого восхождения заканчивается, когда исследователь попадает в почти стационарную область, которая не может быть описана с помощью линейного приближения. В этой части поверхности отклика доминирующими становятся коэффициенты регрессии, характеризующие эффекты взаимодействия.

Квазистационарную область обычно удается описать полиномами второго порядка. Для этого нужно иметь такую систему планирования, в которой каждая переменная будет принимать хотя бы три разных значения. В соответствии с общей идеей шагового эксперимента такое планирование может быть получено добавлением некоторого числа специальным образом расположенных точек к ядру, образованному планированием для линейного приближения. Такие планы называются *композиционными* или *последовательными* (*последовательно строящимися*) планами.

На рис. 4.1 показано центральное композиционное планирование, предложенное Боксом и Уилсоном [3] для случая, когда число факторов $k = 3$. Допустим, что вначале, в квазистационарной области, было поставлено пять опытов; четыре опыта, отмеченные на рисунке кружочками, образуют полуреплику типа 2^{3-1}, пятая точка в центре эксперимента служит для оценки кривизны поверхности. Если линейное приближение оказалось недостаточным, то: 1) добавляют четыре точки так, чтобы получить полный факторный эксперимент: восемь точек полного факторного эксперимента образуют вершины куба (в общем случае — гиперкуба); 2) добавляют шесть «звездных» точек с координатами ($\pm\alpha$, 0, 0), (0, $\pm\alpha$, 0), (0, 0, $\pm\alpha$), образующих октаэдр; 3) увеличивают, если нужно, число параллельных опытов n_0 в центре эксперимента. «Звездные» точки на рис. 4.1 отмечены крестиками. Общее число точек при k факторах будет $2^k + 2k + n_0$. Такое планирование требует значительно меньшего числа опытов, чем полный факторный эксперимент типа 3^k. Например, при $k_1 = 3$ и $k_2 = 4$ для полного факторного эксперимента потребуется число опытов $N_1 = 27$ и $N_2 = 81$, тогда как для центрального композиционного планирования можно будет ограни-

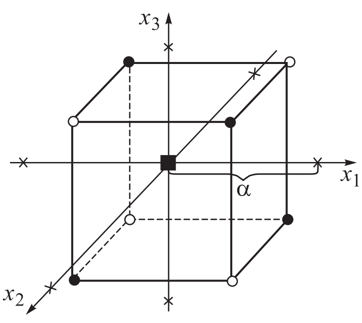

Рис. 4.1. Центральное композиционное планирование в трехмерном факторном пространстве

ТЕМЫ ДЛЯ ПОДГОТОВКИ К ЭКЗАМЕНАМ

1. Наблюдение и эксперимент как основы математического моделирования.
2. Обработка результатов измерений. Прямые равноточные измерения. Критерии оценки грубых погрешностей.
3. Ранговая корреляция при обработке результатов эксперимента.
4. Принципы моделирования систем на эмпирическом уровне.
5. Ошибки оценивания.
6. Проверка гипотезы адекватности модели.
7. Метод наименьших квадратов для одного фактора.
8. Обобщение метода наименьших квадратов на многофакторный линейный случай.
9. Статистический анализ.
10. Взвешенный метод наименьших квадратов и статистический анализ.
11. Обработка результатов дублированных опытов.
12. Использование регрессионных моделей при анализе результатов «разрозненного» эксперимента.
13. Планирование завершающих исследований.
14. Этап предпланирования эксперимента. Выбор факторов. Область проведения эксперимента. Пассивный и активный эксперименты. Сопоставление моделей.
15. Полные факторные эксперименты.
16. Интервал варьирования и кодированные переменные. Базовая точка.
17. Полный факторный эксперимент (ПФЭ) 2^2.
18. Полный факторный эксперимент 2^3.
19. Полный факторный эксперимент 2^k.
20. Дробный факторный эксперимент.
21. Построения полуреплик.
22. Построения четверть-реплик.
23. Выбор дробных реплик.
24. Определяющий контраст.
25. Обобщенный определяющий контраст.
26. Разрешающая способность и выбор дробных реплик.
27. Анализ факторных экспериментов.
28. Свойства факторных экспериментов.
29. Факторные эксперименты с повторными наблюдениями.
30. Проверка гипотезы адекватности.
31. Насыщенное и ненасыщенное планирование.
32. Линейные планы.
33. Построение линейных ортогональных планов.

34. Критерии оптимальности планов.
35. Типы планов эксперимента.
36. D-оптимальные планы (привести теоремы).
37. Непрерывные D-оптимальные планы на отрезке $\Omega = (-1, 1)$ Кифера и Коно.
38. Непрерывные D-оптимальные планы для квадратичной регрессии на гиперкубе.
39. Планы второго порядка.
40. Ортогональный центрально-композиционный план второго порядка.
41. Структура центральных композиционных планов.
42. Планы Бокса.
43. Планы Хартли.
44. Ортогональные центрально-композиционные планы второго порядка.
45. Произвольный симметричный центрально-композиционный план.
46. Многомерные ортогональные центрально-композиционные планы второго порядка.
47. Основные понятия ротатабельности центрально-композиционных планов.
48. Планы второго порядка с единичной областью планирования.
49. Многомерные модели ротатабельных центрально-композиционных планов.
50. Моменты ротатабельного плана.
51. Метод построения ротатабельных планов второго порядка в трех и более измерениях.
52. Проверка гипотезы адекватности модели при наличии повторных испытаний в центре плана.
53. Проверка гипотезы адекватности модели при наличии повторных испытаний в точках плана.
54. Линейная регрессия.
55. Проверка гипотез при использовании линейной регрессии.
56. Интервальные оценки при линейной регрессии.
57. Многофакторная линейная регрессия.
58. Проверка гипотез при использовании множественной линейной регрессии.
59. Другие модели линейной регрессии.
60. Исследование уравнения регрессии. Анализ остатков.
61. Проверка значимости коэффициентов многомерной модели.
62. Многофакторный дисперсионный анализ.
63. Группировка данных при однофакторном дисперсионном анализе.
64. Получение оценок дисперсий и выводов о степени влияния фактора.
65. Исследование поверхности отклика.
66. Метод крутого восхождения. Ротатабельные центральные композиционные планы. Планы типа B_m. Правило движения вдоль градиента. Последовательное планирование эксперимента.
67. Канонические модели второго порядка и их анализ.
68. Планы для подбора модели второго порядка и для изучения поверхности отклика.

ЛИТЕРАТУРА

1. *Сидняев Н.И., Мельникова Ю.С.* Оценки статистических параметров распределений: метод. рекомендации к домашнему заданию по математической статистике. URL: http://wwwcdl.bmstu.ru/fn1/OcenkiSPR.html (дата обращения 27.04.2012).

2. *Сидняев Н.И.* Теория вероятностей и математическая статистика: учеб. пособие. М.: Юрайт, 2011. 310 с.

3. *Сидняев Н.И., Вилисова Н.Т.* Введение в теорию планирования эксперимента: учеб. пособие. М.: Изд-во МГТУ им. Н.Э. Баумана, 2011. 463 с.

4. *Сидняев Н.И.* Теория планирования эксперимента и анализ статистических данных: учеб. пособие. М.: Юрайт, 2011. 399 с.

5. *Сидняев Н.И., Смирнов А.В.* Математическое моделирование оценки возможного ущерба с помощью ортогональных планов второго порядка // Современные естественно-научные и гуманитарные проблемы: сб. тр. М.: Логос, 2005. С. 376–384.

6. *Сидняев Н.И.* Теория планирования эксперимента и анализ статистических данных: учеб. пособие для магистров. 2-е изд., перераб. и доп. М.: Юрайт, 2014. 495 с.

7. *Сидняев Н.И., Садыхов Г.С., Савченко В.П.* Модели и методы оценки остаточного ресурса изделий радиоэлектроники. М.: Изд-во МГТУ им. Н.Э. Баумана, 2015. 382с.

8. *Малафеев С.И., Копейкин А.Н.* Надежность технических систем. Примеры и задачи: учеб. пособие. СПб.: Лань, 2012. 320 с.

9. *Гришин А.Ф., Кочерова Е.В.* Статистические модели: построение, оценка, анализ: учеб. пособие. М.: Финансы и статистика, 2005. 416 с.

10. *Сидняев Н.И., Мельникова Ю.С.* Метод нахождения показателей структурной надежности // Междунар. конф. «Теория вероятностей и ее приложения», посвящ. 100-летию со дня рождения Б.В. Гнеденко (Москва, 26–30 июня 2012 г.): тез. докл. / под ред. А.Н. Ширяева, А.В. Лебедева. М.: ЛЕНАНД, 2012. С. 260–261.

11. *Сидняев Н.И.* О теории восстановления для функционирования восполняемых объектов транспортировки углеводородов // Тр. междунар. симп. «Надежность и качество». В 2 т. / под ред. Н.К. Юркова. Т. 1. Пенза: Изд-во ПГУ, 2010, 532 с.

12. *Радченко С.Г.* Методология регрессионного анализа. Киев: Корнийчук, 2011. 376 с.

13. *Федоров В.А., Сергеев Н.П., Кондрашин А.А.* Контроль и испытания в проектировании и производстве радиоэлектронных средств. М.: Техносфера, 2005. 502 с.

14. *Sidnyaev N.I., Andreytseva K.S.* Independence of the Residual Quadratic Sums in the Dispersion Equation with Noncentral χ^2-Distribution // Applied Mathematics. Vol. 02. No. 10 (October 2011). Pp.1303–1308. DOI: 10.4236/am.2011.210181. URL: http://file.scirp.org/HTML/7855.html (дата обращения 06.04.2016).

15. *Сидняев Н.И., Говор С.А.* Деструкция металлических поверхностей гидравлических систем в морской воде // Тр. междунар. симп. «Надежность и качество». В 2 т. / под ред. Н.К. Юркова. Т. 2. Пенза: Изд-во ПГУ, 2014. С. 199–202.

16. *Дрейпер Н., Смит Г.* Прикладной регрессионный анализ. Множественная регрессия. 3-е изд. М.: Вильямс, 2007. 320 с.

ВАРИАНТЫ ДОМАШНЕГО ЗАДАНИЯ

Исходные данные для выполнения первой части задания приведены в табл. П1:

1) матрица плана $\bar{D}$;

2) функция отклика вида $\eta = \sum_{j=1}^{P} f_j\left(x_1, x_2, ..., x_k\right)\beta_j$, причем оценки $\beta_1, \beta_2, ..., \beta_j$ являются неизвестными;

3) наблюдения $y_1, y_2, ..., y_N$ в точках плана $\xi(N)$;

4) точка $(x_1, x_2, ..., x_k)$ факторного пространства R^k.

Требуется определить:

1) матрицу планирования X;

2) МНК-оценку $\hat{\beta} = \left(\hat{\beta}_1, \hat{\beta}_2, ..., \hat{\beta}_P\right)$ вектора неизвестных параметров $\beta = (\beta_1, \beta_2, ..., \beta_P)$;

3) МНК-оценку $\hat{\eta} = \hat{\eta}(x_1, x_2, ..., x_k)$ функции отклика $\eta = \eta(x_1, x_2, ..., x_k)$ в точке $(x_1, x_2, ..., x_k)$;

4) матрицу ковариаций $D\left(\hat{\beta}\right) = \sigma^2 (X^T X)^{-1}$.

Таблица П1

Номер варианта	D	η	$y_1, ..., y_k$	$x_1, ..., x_k$
1	$\begin{pmatrix} 2 & 1 \\ 2 & -1 \\ -2 & -1 \\ -2 & 1 \end{pmatrix}$	$\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$	1; 2; 3; 4	(2, 3)
2	$\begin{pmatrix} 1 & -1 \\ 1 & 1 \\ 0 & 0 \end{pmatrix}$	$\beta_0 + \beta_1 x_1 + \beta_2 x_2$	–2; 3; 1	(2, 4)
3	$\begin{pmatrix} 0 & 2 & 1 \\ 0 & -2 & 1 \\ 2 & 0 & -1 \\ -2 & 0 & -1 \end{pmatrix}$	$\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$	2; 3; 0; 1	(1, 1, 1)

Окончание табл. П1

Номер варианта	D	η	$y_1, \ldots, y_k$	$x_1, \ldots, x_k$
4	$\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$	$\beta_0 + \beta_1 x_1 + \beta_{11} x_1^2$	2; 0; 1	(2)
5	$\begin{pmatrix} -1 & 1 \\ 0 & 0 \\ 1 & 1 \end{pmatrix}$	$\beta_0 + \beta_1 x_1 + \beta_2 x_2$	3; 2; 5	(3, 1)

Исходные данные для выполнения второй части задания приведены в табл. П2:

1) план $\xi(N)$: $\bar{x}^1, \bar{x}^2, \ldots, \bar{x}^n$; $m_1, m_2, \ldots, m_n$, где $\sum_1^n m_i = N$; $\bar{x}^1, \bar{x}^2, \ldots, \bar{x}^n$ — спектр плана;

2) функция отклика вида $\eta = \sum_{j=1}^{P} f_j\left(x_1, x_2, \ldots, x_k\right)\beta_j$; при этом $\beta_1, \beta_2, \ldots, \beta_j$ являются неизвестными параметрами;

3) повторные наблюдения $\{y_{ls}\}$ $(l = \overline{1, n};\ \ s = \overline{1, m_l})$, т. е. заданы наблюдения $y_{l1}, y_{l2}, \ldots, y_{lm_l}$ в точке плана $\bar{x}^l = (x_{1l}^*, x_{2l}^*, \ldots, x_{kl}^*)$;

4) точка $(x_1, x_2, \ldots, x_k)$ факторного пространства R^k.

Требуется определить:

1) матрицу спектра плана $\tilde{D}$;

2) матрицу $\tilde{X}$;

3) МНК-оценку $\hat{\eta} = \hat{\eta}(x_1, x_2, \ldots, x_k)$ функции отклика $\eta = \eta(x_1, x_2, \ldots, x_k)$ в точке $(x_1, x_2, \ldots, x_k)$;

4) МНК-оценку $\hat{\beta} = \left(\hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta}_P\right)$ вектора неизвестных параметров $\beta = (\beta_1, \beta_2, \ldots, \beta_P)$;

5) матрицу ковариаций $D\left\{\hat{\beta}\right\} = \sigma^2 \left(\tilde{X}^T V \tilde{X}\right)^{-1}$.

Таблица П2

Номер варианта	$\xi(N)$	η	$\{y_{ls}\}$, $l = \overline{1, n}$; $s = \overline{1, m_l}$	$(x_1, x_2, \ldots, x_k)$
1	$\bar{x}^1 = (-1, -1, 1)$, $\bar{x}^2 = (1, -1, 1)$, $\bar{x}^3 = (1, 1, 1)$, $\bar{x}^4 = (-1, 1, -1)$, $m_1 = m_2 = m_3 =$ $= m_4 = 2$	$\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$	$y_{11} = 1, y_{12} = 4$, $y_{21} = 3, y_{22} = 2$, $y_{31} = 1, y_{32} = 5$, $y_{41} = 2, y_{42} = 2$, $n = 4$, $m = 2$	(–2, 1, 2)

Номер варианта	$\xi(N)$	η	$\{y_{ls}\}$, $l=\overline{1,n}$; $s=\overline{1,m_l}$	$(x_1, x_2, \ldots, x_k)$
2	$\bar{x}^1=(-1,\ 1)$, $\bar{x}^2=(-1,\ -1)$, $\bar{x}^3=(-1,\ 1)$, $\bar{x}^4=(1,\ -1)$, $m_1=m_2=m_3=$ $=m_4=2$	$\beta_0+\beta_1x_1+\beta_2x_2+\beta_{12}x_1x_2$	$y_{11}=-3$, $y_{12}=-1$, $y_{21}=-2$, $y_{22}=0$, $y_{31}=3$, $y_{32}=1$, $y_{41}=0$, $y_{42}=2$, $n=4$, $m=2$	$(-3,\ 1)$
3	$\bar{x}^1=(1)$, $\bar{x}^2=(2)$, $\bar{x}^3=(0)$, $\bar{x}^4=(-1)$, $m_1=m_2=m_3=$ $=m_4=2$	$\beta_0+\beta_1x_1$	$y_{11}=-1$, $y_{12}=3$, $y_{21}=0$, $y_{22}=4$, $y_{31}=2$, $y_{32}=0$, $y_{41}=-2$, $y_{42}=2$, $n=4$, $m=2$	(-2)
4	$\bar{x}^1=(0)$, $\bar{x}^2=(1)$, $\bar{x}^3=(-1)$, $m_1=m_2=m_3=2$	$\beta_0+\beta_1x_1+\beta_{11}x_1^2$	$y_{11}=6$, $y_{12}=4$, $y_{21}=3$, $y_{22}=5$, $y_{31}=1$, $y_{32}=3$, $n=4$, $m=2$	(4)
5	$\bar{x}^1=(-2,\ -1)$, $\bar{x}^2=(-2,\ 1)$, $\bar{x}^3=(2,\ 1)$, $\bar{x}^4=(2,\ -1)$, $m_1=m_2=m_3=$ $=m_4=2$	$\beta_0+\beta_1x_1+\beta_2x_2+$ $+\beta_{12}x_1x_2$	$y_{11}=1$, $y_{12}=3$, $y_{21}=2$, $y_{22}=4$, $y_{31}=4$, $y_{32}=0$, $y_{41}=1$, $y_{42}=-4$, $n=4$, $m=2$	$(-3,\ 3)$

ОГЛАВЛЕНИЕ

Учебное издание

Сидняев Николай Иванович

Статистический анализ
и теория планирования эксперимента

Редактор *Е.К. Кошелева*
Корректор *Л.В. Забродина*
Художник *Я.М. Асинкритова*
Компьютерная графика *В.А. Филатовой*
Компьютерная верстка *Т.В. Батраковой*

Оригинал-макет подготовлен
в Издательстве МГТУ им. Н.Э. Баумана.

В оформлении использованы шрифты
Студии Артемия Лебедева.

Подписано в печать 30.09.2017. Формат 70×100/16.
Усл. печ. л. 16,25. Тираж 100 экз. Изд. № 080-2016. Заказ

Издательство МГТУ им. Н.Э. Баумана.
105005, Москва, 2-я Бауманская ул., д. 5, стр. 1.
press@bmstu.ru
www.baumanpress.ru

Отпечатано в типографии МГТУ им. Н.Э. Баумана.
105005, Москва, 2-я Бауманская ул., д. 5, стр. 1.
baumanprint@gmail.com

Информационные технологии в образовательном процессе МГТУ им. Н.Э. Баумана

Соответствие современным тенденциям в высшем образовании:

- переход Университета на двухуровневую систему обучения
- введение блочно-модульной схемы учебного процесса и новых УМКД
- оперативный доступ к образовательным материалам через сеть Университета и Интернет

Учебники, учебные пособия и методические пособия, лабораторные практикумы, курсы лекций и другие материалы в электронной форме

В 2014 году Издательство МГТУ им. Н.Э. Баумана запустило в эксплуатацию web-портал (http://ebooks.bmstu.ru) для обеспечения оперативного доступа студентов к учебной литературе в электронной форме через сеть Университета и Интернет. На площадке web-портала в настоящее время размещены учебно-методические издания, вышедшие в Издательстве МГТУ им. Н.Э. Баумана за последние 10 лет.

Студентам:

- полнотекстовый поиск требуемых материалов как по базе данных контента, так и внутри самого документа
- рубрикаторы по факультетам/кафедрам МГТУ им. Н.Э. Баумана, областям знаний, кодам специальностей (ОКСО), ключевым словам, указателю авторов
- доступ к контенту изданий путем просмотра или загрузки на стационарный, мобильный или планшетный компьютер, смартфон, коммуникатор в различных форматах

Авторам:

- возможность оперативно вносить изменения и дополнения в текст учебных изданий
- использование в учебном процессе материалов web-портала вне зависимости от выхода их печатной версии
- публикация эксклюзивной малотиражной литературы и повышение индекса научного цитирования автора и рейтинга Университета

Москва
ИЗДАТЕЛЬСТВО
МГТУ им. Н. Э. Баумана

МГТУ им. Н.Э. Баумана на всех этапах своего развития считал издательскую деятельность одним из важнейших направлений в работе. На протяжении почти двух столетий здесь регулярно издавались научные труды, учебные пособия и другая литература, способствующая развитию системы технического образования в России. В 1876–1877 годах в условиях нехватки технической литературы, учебников для будущих инженеров ее редактированием и изданием занялись сами студенты. Это были лекции профессоров Н.Е. Жуковского, Ф.Е. Орлова, П.К. Худякова, А.П. Гавриленко, напечатанные на литографском камне!

Издательство в его нынешнем виде было организовано в 1989 году. Сейчас это мощная, великолепно технически оснащенная, имеющая высококвалифицированный персонал, динамично развивающаяся структура. Издательство обеспечивает литературой учебный процесс не только Университета, но и других технических вузов.

Мы выпускаем учебники и учебные пособия, методическую и научно-практическую литературу, монографии и многое другое.

С нами работают ведущие преподаватели МГТУ им. Н.Э. Баумана, известные ученые, инженеры-практики.

Издательство МГТУ им. Н.Э. Баумана по праву заслужило репутацию **лидера** вузовского книгоиздания России. Мы стремимся обеспечить высокое качество выпускаемой продукции, поэтому наши книги неоднократно удостаивались высших наград на общероссийских конкурсах.

По всем интересующим вас вопросам можно обращаться:

Алиев Азер Алиевич	8(499) 263-62-41
Петросян Марина Кимовна	8(499) 265-42-98
Козлов Сергей Владимирович	8(499) 265-42-98

press@bmstu.ru
www.baumanpress.ru