

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
ВОЛОГОДСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Е.А. Фокичева, М.И. Алексеев

ПЛАНИРОВАНИЕ ЭКСПЕРИМЕНТА И ОБРАБОТКА РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЙ

*Утверждено редакционно-издательским советом ВоГУ
в качестве учебного пособия*

ВОЛОГДА
2014

УДК 001.891.57
ББК 22.172
Ф 74

Рецензенты:

доктор технических наук, профессор Санкт-Петербургского
государственного университета путей сообщения *Н.А. Черников*;
доктор технических наук, профессор Череповецкого
государственного университета *А.И. Фоменко*

Фокичева, Е.А.

Ф 74 Планирование эксперимента и обработка результатов исследований:
учебное пособие / Е.А. Фокичева, М.И. Алексеев. – Вологда: ВоГУ,
2014. – 72 с.

Учебное пособие «Планирование эксперимента и обработка результатов исследований» составлено в соответствии с требованиями государственного стандарта и с программой «Планирование эксперимента, обработка результатов исследований». Приводятся основные классификационные признаки экспериментов, основные элементы научно-технического эксперимента, приемы выбора основных факторов эксперимента и технология построения факторных планов.

Учебное пособие предназначено для студентов направления 08.04.01 «Строительство» (МП1 «Водоснабжение городов и промышленных предприятий»), МП2 «Водоотведение и очистка сточных вод» и может быть использовано при подготовке к лабораторным и практическим занятиям и экзаменам.

УДК 001.891.57
ББК 22.172

© ВоГУ, 2014
© Фокичева Е.А., Алексеев М.И., 2014

ПРЕДИСЛОВИЕ

Планирование эксперимента и обработка результатов исследований все больше входят в круг вопросов, необходимых студентам старших курсов технических вузов, дипломникам и магистрам. Многогранность изучаемых явлений, сложность и высокая стоимость оборудования и сооружений, острая нехватка времени – все это вынуждает исследователя вести эксперимент на моделях, планировать исследования. Результаты эксперимента подвергаются обработке и анализу с применением методов математической статистики.

Планирование эксперимента предполагает проведение его в условиях неполной информации о механизме и действующих силах процесса. Методы теории планирования экспериментов направлены на разработку оптимальных планов проведения экспериментов с целью сокращения объема проводимых исследований при заданной точности и достоверности получения результатов, извлечения из полученных опытных данных максимума полезных сведений. Процедура построения математической модели во многом определяется целями исследований, свойствами изучаемого объекта, количеством и качеством фактического материала.

Задача настоящего учебного пособия – ознакомить студентов с кругом актуальных проблем, связанных с постановкой и проведением научных исследований. Целью пособия по дисциплине «Планирование эксперимента, обработка результатов исследований» является формирование у студентов знаний и практических навыков самостоятельного построения моделей объектов, исследования их на компьютере с использованием современных достижений информационных технологий и для решения практических задач.

В пособии рассматриваются вопросы подбора эмпирических формул и оценки их параметров, вопросы оценки истинных значений измеряемых величин и точности измерений, вопросы исследования корреляционных зависимостей, а также некоторые вопросы анализа – интегрирование, дифференцирование, интерполяция, изложены методики расчета погрешностей результатов измерений, построения доверительных интервалов для измеряемой величины, исследования линейной корреляции переменных величин, различные варианты расчета параметров линейных аппроксимирующих функций методом наименьших квадратов, а также содержатся краткие сведения об основных понятиях теории вероятностей.

Учебное пособие «Планирование эксперимента и обработка результатов исследований» содержит необходимые сведения о классификации экспериментов, погрешностях измерений, выборе необходимых факторов, особенностях построения регрессионной модели и т.д.

Учебное пособие «Планирование эксперимента и обработка результатов исследований» составлено в соответствии с требованиями государственного образовательного стандарта Российской Федерации высшего профессионального образования по направлению 270800.68 «Строительство» МП1 «Водоснабжение городов и промышленных предприятий», МП2 «Водоотведение и очистка сточных вод».

Пособие состоит из 5 разделов, тематика которых предполагает изучение: процессов моделирования; вероятностной оценки случайной погрешности и построения функциональных зависимостей при многократных измерениях; обработки результатов эксперимента статистическими методами; теории планирования эксперимента; а также включает проверочные тесты.

При составлении пособия авторы опирались на следующие фундаментальные работы: Адлер Ю.П., Маркова Е.В., Грановский Ю.В. «Планирование эксперимента при поиске оптимальных условий», Красовский Г.И., Филаретов Г.Ф. «Планирование эксперимента», Новицкий П.В., Зограф Э.Н. «Оценка погрешностей измерений», Г.И. Ивченко, Ю.И. Медведев, «Математическая статистика», Ю.В. Линник «Метод наименьших квадратов и основы теории обработки наблюдений» Н. Джонсон, Ф. Лион «Статистика и планирование эксперимента в технике и науке».

ВВЕДЕНИЕ

Экспериментальные исследования являются основным источником получения достоверных сведений об объектах реального мира. Они проводятся с целью выбора рациональных технологических режимов функционирования или оптимизации параметров систем, оценки степени выполнения заданных требований к создаваемым изделиям, выяснения закономерностей функционирования, анализа влияния факторов на показатели качества систем и так далее.

При исследовании технических систем могут использоваться теоретические и эмпирические методы познания. Каждое из этих направлений обладает относительной самостоятельностью, имеет свои достоинства и недостатки. В общем случае, теоретические методы в виде математических моделей позволяют описывать и объяснять взаимосвязи элементов изучаемой системы или объекта в относительно широких диапазонах изменения переменных величин. Однако при построении теоретических моделей неизбежно введение каких-либо ограничений, допущений, гипотез и т.п. Поэтому возникает задача оценки достоверности (адекватности) полученной модели реальному процессу или объекту. Для этого проводится экспериментальная проверка разработанных теоретических моделей. Практика является решающей основой научного познания. В ряде случаев именно результаты экспериментальных исследований дают толчок к теоретическому обобщению изучаемого явления. Экспериментальное исследование дает более точное соответствие между изучаемыми параметрами. Но не следует и преувеличивать результаты экспериментальных исследований, которые справедливы только в пределах условий проведенного эксперимента. Таким образом, теоретические и экспериментальные исследования дополняют друг друга и являются составными элементами процесса познания окружающего нас мира.

В настоящее время процедура обработки экспериментальных данных разработана достаточно хорошо и исследователю необходимо только иметь достаточную подготовку для того, чтобы правильно ее использовать. Основные вопросы, решаемые при обработке результатов эксперимента: вопросы подбора эмпирических формул и оценка их параметров, вопросы оценки истинных значений измеряемых величин и точности измерений, вопросы исследования корреляционных зависимостей и некоторые другие.

Экспериментальные данные формируются путем пассивного наблюдения либо с помощью активного эксперимента. При пассивном наблюдении информация получается путем регистрации необходимых сведений в условиях обычного функционирования объекта. В активном эксперименте производится целенаправленное воздействие на объект по заранее составленной схеме. Активный эксперимент позволяет расширить область исследования, точнее вскрыть закономерности функционирования, сократить потребности в ресурсах на проведение исследования. Но организация и проведение активно-

го эксперимента сложнее пассивного. Кроме того, следует учитывать и принципиальные ограничения в проведении активных экспериментов на действующих объектах, невозможность их осуществления для недоступных объектов.

Модель содержит физическое описание исследуемого объекта или процесса, которое позволяет составить его математическое описание. Неверно построенная модель, в которой не нашли отражения какие-то важные процессы или факторы, влияющие на результат измерений, также приводит к несоответствиям. Измеряемые в эксперименте величины, вычисляемые по полученным из модели рабочим формулам, содержат погрешности, которые носят название модельных погрешностей.

В настоящее время теория планирования эксперимента выступает как самостоятельное научное направление и находит практическое применение там, где проводятся сложные научные и технические экспериментальные исследования [1,2,3]. Теория использует аппарат математической статистики, линейной алгебры, комбинаторики и других разделов математики. Наличие формальной теории планирования эксперимента не исключает необходимости четкого представления физических основ процессов, протекающих в объекте исследования, факторов, воздействующих на объект. Эти сведения нужны на этапе составления плана эксперимента, при анализе и интерпретации результатов.

Методы теории планирования экспериментов направлены на разработку оптимальных планов проведения экспериментов с целью сокращения объема проводимых исследований при заданной точности и достоверности получения результатов, извлечения из полученных опытных данных максимума полезных сведений. Составной частью теории планирования эксперимента является исследование способов обработки результатов эксперимента, проведенного по выбранному плану, анализ свойств получаемых оценок показателей качества объекта. Экспериментальные данные, полученные с помощью теории планирования эксперимента, часто являются основой для применения других математических методов, например градиентных методов оптимизации.

1. НАБЛЮДЕНИЕ И ЭКСПЕРИМЕНТ. МОДЕЛИРОВАНИЕ

1.1. Наблюдение и эксперимент

Наблюдение – это систематическое, целенаправленное восприятие того или иного объекта или явления без воздействия на изучаемый объект или явление. Наблюдение позволяет получить первоначальную информацию по изучаемому объекту или явлению.

Эксперимент – метод изучения объекта, когда исследователь активно и целенаправленно воздействует на него путем создания искусственных условий или использует естественные условия, необходимые для выявления соответствующих свойств. Достоинствами эксперимента по сравнению с наблюдением реального явления или объекта является:

- 1) возможность изучения в “чистом виде”, без влияния побочных факторов, затемняющих основной процесс;
- 2) в экспериментальных условиях можно получить результат более быстро и точно;
- 3) при эксперименте можно проводить испытания столько раз, сколько это необходимо.

Результат эксперимента или измерения всегда содержит некоторую погрешность. Если погрешность мала, то ею можно пренебречь. Однако при этом неизбежно возникают два вопроса: во-первых, что понимать под малой погрешностью, и, во-вторых, как оценить величину погрешности, то есть и результаты эксперимента нуждаются в определенном теоретическом осмыслении.

Целью любого эксперимента является определение качественной и количественной связи между исследуемыми параметрами, либо оценка численного значения какого-либо параметра [1].

В некоторых случаях вид зависимости между переменными величинами известен по результатам теоретических исследований. Как правило, формулы, выражающие эти зависимости, содержат некоторые постоянные, значения которых и необходимо определить из опыта.

Другим типом задачи является определение неизвестной функциональной связи между переменными величинами на основе данных эксперимента. Такие зависимости называют эмпирическими.

Однозначно определить неизвестную функциональную зависимость между переменными невозможно даже в том случае, если бы результаты эксперимента не имели ошибок. Тем более не следует этого ожидать, имея результаты эксперимента, содержащие различные ошибки измерения.

Поэтому следует четко понимать, что целью математической обработки результатов эксперимента является не нахождение истинного характера зависимости между переменными или абсолютной величины какой-либо константы, а представление результатов наблюдений в виде наиболее простой формулы с оценкой возможной погрешности ее использования.

1.2. Моделирование

Многогранность изучаемых явлений, сложность и высокая стоимость оборудования и сооружений, острая нехватка времени – все это вынуждает исследователя вести эксперимент на моделях, планировать исследования. Процедура построения математической модели во многом определяется целями исследований, свойствами изучаемого объекта, количеством и качеством фактического материала.

Моделирование – это метод изучения объектов, при котором вместо оригинала (объект) эксперимент проводится на модели (другой объект), а результаты количественно распространяют на оригинал (натурное сооружение).

Таким образом, по результатам опытов с моделью мы должны количественно предсказать поведение оригинала (натурного сооружения) в рабочих условиях. Причем распространение на оригинал выводов, полученных в опытах с моделью, не обязательно должно означать простое равенство тех или иных параметров оригинала и модели. Достаточно получить правило расчета интересующих нас параметров оригинала. К процессу моделирования предъявляются два требования:

Эксперимент на модели должен быть проще, быстрее, экономичнее или безопаснее, чем эксперимент на оригинале (иначе не стоит создавать модель – лучше построить оригинал и эксперимент проводить на нем). Сказанное не означает, что модель должна быть дешевле оригинала. Часто бывает целесообразно построить дорогую, но универсальную модель, которую можно приспособить для моделирования многих оригиналов. Это может оказаться весьма экономичным.

Должно быть известно правило, по которому проводится расчет параметров оригинала на основе испытания модели.

Материальные и мысленные модели. Физические и математические модели

В современной науке термин модель применяют в основном в двух значениях. Первое следует из определения моделирования. Модель – это объект, который в процессе эксперимента заменяет оригинал. Для большей определенности в этом случае мы будем говорить о материальной модели.

Кроме этого, широко используется другое значение. Говоря о модели атома Резерфорда или молекулы Бутлерова, мы, разумеется, не считаем, что речь идет о каких-то опытных установках, на которых экспериментировали Резерфорд и Бутлеров. В данном случае мы имеем дело с мысленными моделями. Мысленная модель – это схема объекта (явления), отражающая его существенные стороны, которая возникает в сознании человека в процессе познания. В связи с этим предмет моделирования часто определяют как построение и практическое использование мысленных, либо материальных моделей. В принципе, вся наука есть, в известном смысле, не что иное как совокупность мысленных моделей.

В применении к технологии очистки воды под моделированием чаще всего понимают экспериментальный метод, связанный с проведением опытов на материальных моделях и распространением результатов на оригинал.

Любая материальная модель строится на основе мысленной. Поэтому при изложении вопросов моделирования мы постоянно будем обращаться к свойствам и особенностям не только материальных, но и мысленных моделей.

Целесообразно определить два важнейших варианта мысленных моделей. Обычно схема объекта первоначально возникает как совокупность наглядно-физических понятий, то есть мы представляем форму, размеры, конструкции. Это физическая мысленная модель. Затем на основе физических представлений мы строим количественное описание объекта. Оно представляет собой совокупность математических структур: уравнений, неравенств, таблиц, графиков и т. д. Это описание называют математической моделью или математическим описанием процесса.

Мысленная модель может развиваться не только от физической к математической, но и в обратном направлении.

Ни одна модель принципиально не способна отразить оригинал полностью и всесторонне. Это вытекает из общефилософских рассуждений и одинаково верно для материальных и мысленных моделей.

Более того, часто на практике целесообразнее пользоваться менее «совершенной» моделью, отражающей только отдельные черты оригинала.

В ряде случаев целесообразно один оригинал моделировать с помощью разных моделей, не похожих друг на друга. Рассмотрим этот вопрос на примере.

Необходимо смоделировать аэротенк. Как должна выглядеть модель аэротенка? Это зависит от того, какие вопросы желательно решать с помощью моделирования. Если это определение кинетических характеристик окисления органических загрязнений, то обычно используют лабораторные модели аэротенков, респирометры. Если необходимо исследовать аэрационное оборудование, то строят фрагмент установки. Если решается вопрос осаждения активного ила, то более тщательно прорабатывается модель вторичного отстойника. При исследовании коррозионной стойкости материалов аэротенка при воздействии сточных вод моделью будут служить просто кусочки материалов стен аэротенка, погруженные в соответствующие среды.

Важной особенностью мысленных моделей является то, что часто имеет смысл пользоваться упрощенной моделью, даже при наличии более совершенной. Как правило, чем проще модель, тем проще сделать на ее основе количественные выводы. Зачастую уточнение, получаемое при использовании более сложной модели, не оправдывает усложнения. Иногда этим уточнением можно вообще пренебречь. В таких случаях целесообразно пользоваться упрощенными моделями.

Теория подобия – аппарат моделирования

Принцип подобия позволяет из класса явлений, описываемых дифференциальными уравнениями, выделить при помощи приведения их к безразмерному виду группу подобных явлений.

Подобными называются явления, у которых все характеризующие их величины в любой точке пространства находятся в одинаковых отношениях между собой. В этом случае подобие называется полным. Если же подобие соблюдается лишь для нескольких величин, то оно называется частичным. Так как всякий процесс протекает в определенном геометрическом пространстве (контуре), то прежде всего устанавливают подобие геометрическое и затем – физическое, характеризующее процесс.

Простейшее представление о геометрическом подобии двух фигур известно из геометрии. Если рассматривать линейные размеры фигуры не только как скалярные величины, но и как имеющие определенное направление, то подобные фигуры должны быть так расположены в пространстве, чтобы их аналогичные «размеры» были параллельны друг другу.

Соотношение аналогичных размеров модели и натурального объекта называется масштабом модели. Взяв какой-либо линейный размер модели l_m и разделив его на соответствующий размер натуры l_n , получим линейный масштаб модели:

$$a_l = l_n / l_m. \quad (1)$$

Так, например, если рассматривать два аппарата с мешалками, диаметры которых соответственно равны D_1 и D_2 , высоты слоя перемешиваемой жидкости H_1 и H_2 , диаметры лопастей d_1 и d_2 , высоты лопастей h_1 и h_2 , то, чтобы указанные аппараты были подобны необходимо соблюдение равенства:

$$a_D = D_1/D_2; a_H = H_1/H_2; a_d = d_1/d_2; a_h = h_1/h_2. \quad (2)$$

Из условия геометрического подобия аппаратов:

$$a_D = a_H = a_d = a_h. \quad (3)$$

Однако геометрическое подобие аппаратов удобнее выражать, вводя определяющий геометрический размер. Так, если в рассматриваемой модели аппарата с мешалкой за определяющий размер принять диаметр лопасти мешалки d и с ним сравнивать в данной модели все основные размеры, $D_2/d_2 = i_D$; $H_1/d_1 = i_H$; $h_m/d_m = i_h$, то геометрическое подобие будет соблюдаться, если в любом другом аппарате при той же определяющей характеристике сохранится одинаковое значение безразмерных отношений:

$$D_2/d_2 = D_1/d_1 = i_D; H_2/d_2 = H_1/d_1 = i_H; h_2/d_2 = h_1/d_1 = i_h. \quad (4)$$

Безразмерное отношение i_D , i_H , i_h и т. д., сохраняющее одно и то же значение в модели и натуре, называется инвариантом геометрического подобия.

Таким образом, геометрическое подобие будет соблюдаться тогда, когда инварианты его в сравниваемых системах сохраняют одно и то же значение.

Между собой инварианты геометрического подобия могут быть численно и не равны. Безразмерность инвариантов подобия позволяет переносить условия геометрического подобия на аппараты любых размеров, важно лишь, чтобы отношение данного размера к определяющему следовало равенству

По аналогии с геометрическим физическое подобие соблюдается тогда, когда инварианты его в сравниваемых системах сохраняют одно и то же значение.

Инварианты физического подобия так же, как и геометрического, должны быть величинами безразмерными. Но поскольку физическое явление характеризуется рядом физических величин (скорость, плотность, вязкость и т. п.), то составление из этих величин безразмерных отношений представляет основную задачу метода подобия.

Решение этой задачи осуществляется двумя путями: при помощи подобного преобразования дифференциальных уравнений или при помощи анализа размерностей.

Так как дифференциальное уравнение представляет математическую модель описываемого физического явления, то его подобное преобразование означает подобие моделей явления. Границы соблюдения этого подобия устанавливаются опытным путем. В результате подобного преобразования дифференциальных уравнений последние заменяются так называемыми критериальными уравнениями. В этом случае инварианты физического подобия называются критериями подобия. Подобное преобразование сводится к замене физических величин под дифференциальными операторами конечными величинами.

Теоремы подобия

Практическое применение теории подобия к экспериментальному и теоретическому исследованию процессов основано на трех теоремах подобия.

Первая теорема подобия. Подобные явления характеризуются численно равными критериями подобия, или у подобных явлений критерии подобия равны между собой.

Вторая теорема подобия. Любая зависимость между физическими величинами, характеризующими явление или процесс, может быть представлена в виде взаимной зависимости между критериями подобия.

Таким образом, вторая теорема подобия содержит ответ на вопрос, как следует обрабатывать полученные экспериментальные данные, или, в какой форме может быть получено решение системы дифференциальных уравнений, описывающих процесс, с помощью методов теории подобия.

Третья теорема подобия (обратна первой). Подобны те явления или системы, которые описываются одинаковыми уравнениями связи и условия однозначности которых подобны.

Подобие условий однозначности обеспечивается равенством определяющих критериев подобия в случае, если явления или процессы качественно одинаковы. Качественно одинаковыми будут процессы, математическое описание которых одинаково.

Таким образом, третья теорема подобия формулирует необходимые и достаточные условия для подобия явлений или процессов. Методами теории подобия можно перенести результаты опытов, полученные на моделях, на группу или класс подобных систем.

Основные принципы анализа размерностей.

Метод анализа размерностей

В некоторых случаях из-за сложности явления или процесса не удастся составить его полное математическое описание в виде системы дифференциальных уравнений, а возможно лишь в самом общем виде представить зависимость между физическими величинами и геометрическими параметрами, характеризующими процесс. Вид такой зависимости можно найти на основе анализа размерностей физических величин, вошедших в уравнение. Этот метод базируется на π -теореме, являющейся частным случаем второй теоремы подобия.

Формулировка π -теоремы: всякое уравнение, связывающее между собой N физических величин (например, скорость, вязкость, плотность и т. п.), среди которых m величин обладают независимыми размерностями (например, масса, длина, время), может быть преобразовано к уравнению, связывающему $(N-m)$ безразмерных критериев (комплексов) и симплексов, составленных из этих величин.

Метод анализа размерностей позволяет представить общую функциональную зависимость для любого исследуемого процесса в виде уравнения связи между строго определенным числом безразмерных комплексов, состоящих из физических величин с определенной размерностью, выраженной с помощью основных единиц измерения. Этот метод базируется на двух допущениях: 1 – известно заранее (из практических данных), от каких именно параметров процесса и переменных зависит рассматриваемая величина; 2 – между всеми значимыми для исследуемого процесса физическими величинами существует связь в виде степенного многочлена.

Сущность физического моделирования заключается в решении двух задач: в нахождении постоянных в критериальном уравнении, которым описывается модель процесса, и построении модели. Если рассматриваемый процесс является сложным, зависящим от большого числа параметров, для которого трудно построить полную модель, прибегают к приближенной модели.

В качестве примера возьмем критериальное уравнение сопротивления при движении жидкости в аппаратах:

$$E_n = f(\text{Re}, l/d), \quad (5)$$

$$\frac{\Delta P}{\rho \omega^2} = f\left(\frac{\omega \cdot d \cdot \gamma}{\mu \cdot g}, \frac{l}{d}\right) \quad (6)$$

Изменяемые параметры: геометрические l , d и и т. п., физические γ , μ и т. п., зависимые переменные ω , ΔP и т. п.

Физическое моделирование выполняется в следующей последовательности:

1. Опыты ставятся в аппаратах различных размеров при соблюдении геометрического подобия и определяются зависимостями искомых переменных величин от физических параметров. Масштабы опытных установок выбираются с соблюдением геометрического подобия. Например, в биофильтрах в качестве определяющего геометрического размера принимается эквивалентный диаметр загрузки d_3 ; в аппаратах с мешалками – диаметр мешалки и т. д.

2. На каждой модели ставится несколько опытов, в которых меняются физические параметры для получения зависимости между ними. Взаимодействие параметров может быть непосредственно учтено при постановке направленного эксперимента или многофакторного планирования эксперимента.

3. Зависимость между безразмерными комплексами строится в логарифмических координатах, так как критериальные уравнения на основе принципа теории размерностей – степенные уравнения и, следовательно, в логарифмических координатах они представляются в виде безразмерных комплексов прямыми линиями. Для этого опытные точки, обработанные в виде безразмерных комплексов (например, безразмерный коэффициент трения $\lambda=2E\mu$ и число Рейнольдса Re), наносятся на логарифмический график. Тангенс угла наклона есть показатель степени критерия Рейнольдса.

Появление на графике, выражающем связь между безразмерными комплексами, горизонтального или вертикального хода прямой указывает на возникновение так называемого автомодельного режима. Автомодельный режим появляется, например, в шероховатых трубах при больших числах Рейнольдса, при этом коэффициент сопротивления становится постоянной величиной, не зависящей от λ .

Автомодельный режим может возникать в различных процессах. Автомодельность может характеризоваться независимостью процесса от любого параметра. Например, в биофильтрах, насадочных колоннах распределение жидкости по сечению колонны не зависит от диаметра колонны.

Наличие автомодельных условий, т. е. исключение влияния одного или нескольких параметров на процесс, значительно упрощает задачу моделирования процесса.

Если число критериев подобия, входящих в уравнение, достаточно велико, построение модели, подобной оригиналу, — задача совершенно нереальная.

Действительно, каждая реакция (в реакторе часто одновременно протекают несколько реакций) характеризуется своим типом зависимости от опре-

деленных факторов – концентрации и температуры. Кроме того, реакции могут сопровождаться массообменом и практически всегда— выделением (поглощением) тепла и теплообменом. На работу реактора влияет структура потока жидкости в нем. В результате число факторов, а следовательно и критериев подобия, становится очень большим.

Правда, в ряде случаев условия работы реакторов позволяют выделить лимитирующую стадию процесса. Тогда влиянием большинства факторов можно пренебречь, условия подобия упрощаются, создание модели на основе теории подобия становится возможным.

При моделировании процесса технологии очистки воды особое значение имеет проблема увеличения аппаратуры. Как показывает промышленная практика, с увеличением размеров аппарата обычно снижается его эффективность по сравнению с эффективностью лабораторной модели. Основная причина заключается в нарушении равномерности распределения потоков, то есть в изменении гидродинамических условий в результате продольного перемешивания.

Упорядочить гидродинамику промышленного аппарата в некоторых случаях удастся на гидравлической модели, диаметр которой соответствует диаметру аппарата в натуральную величину, а высота значительно уменьшена.

В сложных условиях для расчета аппарата промышленного масштаба необходимо учитывать пять групп переменных величин: геометрические, гидродинамические, тепловые, массообменные, химические.

Геометрический масштаб может быть определен из условия пропорциональности между определенными линейными размерами модели и промышленного аппарата:

$$K = D_a / D_m, \quad (7)$$

где D_a , D_m – диаметры аппарата и модели.

Гидродинамическое подобие при условии использования в аппарате и модели вещества с одинаковыми свойствами (плотность, вязкость, теплопроводность и др.) упростится и выразится равенством критериев Рейнольдса:

$$D_a \cdot v_a = D_m \cdot v_m, \quad (8)$$

где v_a , v_m — скорость в аппарате и модели.

При отсутствии химических реакций в системе тепловое подобие также обеспечивается равенством критериев Рейнольдса:

$$\omega_a \cdot D_a = \omega_m \cdot D_m. \quad (9)$$

Химическое подобие требует равенства критериев Дамклера:

$$D_a = \frac{k \cdot D_a}{\omega_a} = \frac{k \cdot D_m}{\omega_m}, \quad (10)$$

где k – константа скорости реакции.

При увеличении масштаба аппаратов, в которых осуществляются процессы, осложненные химической реакцией, можно основываться только на приближенном моделировании.

Масштабный переход может быть успешно решен путем создания таких аппаратных конструкций, в которых поверялись бы отдельные хорошо отработанные элементы. Такой многоэлементный или многосекционный метод масштабирования нашел применение в практике моделирования.

Аналогия

Обычно в теории подобия подразумевают, что в оригинале и модели протекают процессы, одинаковые по своей физической сущности, разница между которыми состоит только в значении параметров. Если обобщить понятие масштаб, включив в него соотношение между любыми одноименными параметрами оригинала и модели (например, между размерами, скоростями реакции и т. д.), то модель будет отличаться от оригинала только масштабами. По разным параметрам масштабы выбирают не произвольно, а исходя из равенства критериев подобия.

Для практики чрезвычайно важно, чтобы результаты опытов на модели были количественно приложимы к оригиналу, в противном случае моделирование бессмысленно; важно, чтобы модель была проще оригинала – иначе моделирование нецелесообразно: легче экспериментировать прямо на оригинале. Но вот физическая идентичность модели и оригинала не обусловлена никакими практическими соображениями. Стало быть, без нее можно обойтись.

В таком случае, чем же руководствоваться при построении модели? Очевидно, теми требованиями, о которых шла речь выше, прежде всего – возможностью количественного переноса данных с модели на оригинал. Если такой перенос возможен, то безразлично, как устроена модель – лишь бы данные были надежными.

Что же тогда, если не общность физической сущности процесса, обеспечивает общность количественных закономерностей в оригинале и модели? Ответ на этот вопрос вытекает из замечательного свойства многих природных явлений: самые различные по физическому механизму явления очень часто описываются одинаковыми уравнениями, т. е. количественные закономерности могут совпадать для физически совершенно разных явлений.

Аналогичными называют объекты и процессы, описываемые аналогичными уравнениями. Уравнения аналогичны, если они содержат разные физические величины, но все операторы (знаки алгебраического действия, символы функций, символы дифференцирования и интегрирования и т. д.) в них совпадают и следуют в одном и том же порядке. Величины, которые в аналогичных уравнениях стоят на одинаковых местах, называют аналогами.

В математике мы, по существу, никогда не имеем дела с физическими величинами. Математика оперирует не величинами, а их численными значениями. Поэтому аналогичность уравнений, описывающих различные физиче-

ские объекты, приводит к одинаковому распределению численных значений в пространстве и времени, хотя эти значения характеризуют разные физические величины. При наличии аналогии можно построить модель, подобную оригиналу, хотя физическая сущность явлений в модели и оригинале различна.

Примером является аналогия, основанная на уравнениях, описывающих фильтрацию жидкости и прохождение электрического тока в трехмерном проводнике (например, электролите) – так называемая электрогидродинамическая аналогия.

Процесс фильтрации подчиняется закону Дарси-Вейсбаха, прохождение тока – закону Ома.

Закон Дарси-Вейсбаха:

$$\vec{w} = -k \cdot \text{grad}P , \quad (11)$$

закон Ома:

$$\vec{j} = -\chi \cdot \text{grad}U , \quad (12)$$

где $\vec{w}$ – вектор скорости жидкости; k – коэффициент фильтрации; P – давление; $\vec{j}$ – вектор плотности тока; χ – удельная проводимость среды; U – электрический потенциал; grad – градиент, векторный оператор.

Градиент скалярной величины (например, давления, потенциала и т. д.) – это вектор, направление которого соответствует возрастанию данного скаляра с наибольшей скоростью. Значение градиента равно производной от скалярной величины по данному направлению.

Оба уравнения совпадают по форме. Следовательно, существует аналогия между потоком жидкости и электрическим полем. Проводимость служит аналогом коэффициента фильтрации, зависящего от пористости среды, вязкости жидкости и формы частиц. Аналог давления – электрический потенциал.

Предположим, что надо смоделировать движение жидкости через фильтрующую загрузку. Аппарат прямоугольного сечения, вода на входе в слой загрузки равномерно распределяется по всему живому сечению, следовательно, все сечения, параллельные данному, находятся в одинаковых условиях. Тогда трехмерную задачу мы заменяем двухмерной, так как моделировать движение жидкости можно в одном сечении.

Для получения аналоговой модели строят горизонтальную ванну, заливают слоем раствора электролита (он моделирует загрузку фильтра). Модель делают геометрически подобной сечению оригинала. К поперечным сечениям А и В подводят разность потенциалов, моделирующую перепад давления в слое. После этого измеряют направление и силу тока в разных точках ванны – эти измерения определяют направление и значение скорости в соответствующих точках оригинала. Опыты на модели позволяют найти наиболее целесообразную форму оригинала.

Последовательность проведения аналитических исследований при моделировании

Проведение аналитических исследований можно грубо разделить на семь последовательных стадий:

1. Постановка задачи. Это, возможно, наиболее важная стадия, поскольку не существует общих правил, достаточно полезных во всех случаях. Технические проблемы настолько разнообразны, что для успеха анализа должна быть ясна природа задачи. Постановка задачи не только определяет цель анализа, но и пути решения задачи.

2. Изучение теоретических основ процесса. Определение фундаментальных законов, которым подчиняется механика явления, лежащих в основе проблемы. Теоретические основы процессов изучаются обычно по различным источникам (опубликованным и неопубликованным). Если не удастся подобрать удовлетворительную теорию, прибегают к постулатам. Справедливость последних проверяется сравнением результатов решений математической модели, построенной на основе принятых постулатов, с экспериментальными данными. Таким же образом выясняется, какая из возможных теорий правильнее отражает сущность изучаемых явлений. Быстрота подобного выбора является одним из преимуществ аналитического подхода к исследованию процессов.

3. Составление уравнений процессов. На основе выбранной физической модели применительно к решаемой задаче записывается система соответствующих уравнений. Этот шаг ведет анализирующего к ясному, недвусмысленному пониманию и определению проблемы. На данном этапе не требуется производить никаких действий кроме, если это возможно, упрощения уравнений путем пренебрежения незначительными членами. Отбрасывая их, надо быть осторожным и убедиться, что невключенный член уравнения действительно незначим на протяжении всего хода решения задачи.

4. Построение модели. Когда уравнения составлены, определяется метод решения их совместной системы.

5. Выбор одного из нескольких возможных способов решения в зависимости от уровня проведенного исследования процесса и сложности уравнений модели.

6. Анализ модели. Фактически можно выделить три основных уровня анализа модели. Логический – если надо решить несложный вопрос, и уравнения достаточно просты, ответ получают путем просмотра модели, не решая входящие в нее уравнения. Ясно, что этот метод не может быть распространен на более сложные случаи без значительного увеличения количества неапробированных предложений и допущений.

Аналитический уровень – очень ограниченный задачами умеренной сложности, связан с решением уравнений аналитическими методами.

Численный уровень с применением ЭВМ – наиболее результативный, единственно целесообразный для решения задач высокой сложности.

7. Изучение и подтверждение результатов, полученных при решении математической модели. Любому непредполагаемому заранее решению должно быть дано рациональное объяснение, чтобы гарантировать себя от ошибок, которые могут появиться в результате вычислений.

В достаточно простых случаях расчет несомненно предпочтительнее моделирования: дешевле сосчитать, чем строить модели и экспериментировать. Однако по мере усложнения процессов усложняются и математические модели. Все чаще возникают ситуации, когда математическая модель есть, но решить эти сложные уравнения человеку не под силу.

Необходимо иметь в виду, что нельзя противопоставлять физическое и математическое моделирование, и совсем бессмысленно считать, что одно — лучше, а другое — хуже. Важно решить задачу, и на каждом этапе следует принять тот метод, который является более эффективным.

Во многих случаях физический эксперимент оказывается дешевле сложного расчета, хотя возможно и то и другое

И наконец, не следует забывать, что, как правило, теория дает общий вид уравнений математического описания а численные коэффициенты этих уравнений, значения которых различаются в каждом частном случае, должны быть найдены из опыта.

2. ИЗМЕРЕНИЯ. ОЦЕНКА ПОГРЕШНОСТЕЙ ИЗМЕРЕНИЙ

2.1. Виды измерений и причины ошибок

Под измерением понимают сравнение измеряемой величины с другой величиной, принятой за единицу измерения. Различают два типа измерений: прямые и косвенные. При прямом измерении измеряемая величина сравнивается непосредственно со своей единицей меры. Значение измеряемой величины отсчитывается при этом по соответствующей шкале прибора.

При косвенном измерении измеряемая величина определяется (вычисляется) по результатам измерений других величин, которые связаны с измеряемой величиной определенной функциональной зависимостью.

При измерении любой физической величины производят проверку и установку соответствующего прибора, наблюдение их показаний и отсчет. При этом никогда истинного значения измеряемой величины не получить. Это объясняется тем, что измерительные средства основаны на определенном методе измерения, точность которого конечна. При изготовлении прибора задается класс точности [2]. Его погрешность определяется точностью делений шкалы прибора. Кроме приборной погрешности на результат измерения влияет еще ряд объективных и субъективных причин, обуславливающих появление ошибки измерения – разности между результатом измерения и истинным значением измеряемой величины. Ошибка измерения обычно неизвестна, как неизвестно и истинное значение измеряемой величины. Исключение составляют измерения известных величин при определении точности измерительных приборов или их тарировке. Поэтому одной из важнейших задач математической обработки результатов эксперимента и является оценка истинного значения измеряемой величины по данным эксперимента с возможно меньшей ошибкой.

2.2. Классификация погрешностей измерений. Вероятностная оценка случайной погрешности и построение функциональных зависимостей при многократных измерениях

Классификация погрешностей измерений. Основные типы погрешностей измерений и их краткая характеристика приведены в таблице 1.

Основные типы погрешностей измерений

I. Погрешности средств измерения и результатов измерений			
Погрешности средств измерения		Погрешности результатов измерения	
отклонения метрологических свойств или параметров средств измерений от номинальных, влияющие на погрешности результатов измерений (создающие так называемые инструментальные ошибки измерений).		отклонение результата измерения $x_{изм}$ от действительного (истинного) значения измеряемой величины x_0 , определяемая по формуле $\Delta x = x_{изм} - x_0$ – погрешность измерения.	
инструментальная	методическая	статическая	динамическая
обусловлена несовершенством применяемых средств измерений. Причины возникновения - неточности, допущенные при изготовлении и регулировке приборов, изменение параметров элементов конструкции и т.д.	обусловлена несовершенством метода измерений или упрощениями, допущенными при измерениях. Причиной методической погрешности может быть недостаточно учтенное взаимное влияние объекта измерений и измерительных приборов	погрешность, свойственная условиям статического измерения, то есть при измерении постоянных величин после завершения переходных процессов в элементах приборов и преобразователей.	погрешность, свойственная условиям динамического измерения. Динамическая погрешность появляется при измерении переменных величин и обусловлена инерционными свойствами средств измерений.
II. Систематические и случайные погрешности			
систематическая		случайная	
составляющая погрешности измерения, остающаяся постоянной или закономерно изменяющаяся при повторных измерениях одной и той же физической величины. Систематические погрешности являются в общем случае функцией измеряемой величины, влияющих величин (температуры, влажности, напряжения питания и пр.) и времени. В функции измеряемой величины систематические погрешности входят при поверке и аттестации образцовых приборов.		называют составляющие погрешности измерений, изменяющиеся случайным образом при повторных измерениях одной и той же величины. Случайные погрешности определяются совместным действием ряда причин.	
III. Погрешности адекватности и градуировки			
адекватности		градуировки	
погрешность при выборе функциональной зависимости. Если зависимость параметра состояния от уровней входного фактора задана при моделировании объекта достаточно точно, то погрешность адекватности оказывается минимальной.		погрешность действительного значения величины, приписанного той или иной отметке шкалы средства измерений в результате градуировки.	

IV. Аддитивная и мультипликативная		
аддитивная		мультипликативная
погрешность, постоянная в каждой точке шкалы. Ярким примером аддитивной погрешности является погрешность квантования или оцифровки		погрешность, линейно возрастающая или убывающая с ростом измеряемой величины
IV. Абсолютная, относительная и приведенная погрешности		
абсолютная	относительная	приведенная
алгебраическая разность между номинальным и действительным значениями измеряемой величины. $\Delta x(x), \Delta y(y)$ – абсолютные погрешности	выраженное в процентах отношение абсолютной погрешности к действительному значению измеряемой или воспроизводимой данным средством измерений величины. $\frac{\Delta x}{x}, \frac{\Delta y}{y}$ – относительные погрешности.	равна отношению абсолютной погрешности измерительного прибора к некоторому нормирующему значению. В качестве нормирующего значения принимается значение, характерное для данного вида измерительного прибора.

2.3. Вероятностная оценка случайной погрешности

Характеристики статистических распределений

Формула для расчета вероятности, что случайная величина x принимает значения в некотором интервале $x_1 \leq x \leq x_2$ имеет вид:

$$P(x_1 \leq x \leq x_2) = \int_{x_1}^{x_2} \rho(x) dx, \quad (13)$$

где величина $\rho(x)$ – плотность распределения вероятности случайной величины x (или статистическое распределение) [3]. Поскольку величина x находится в интервале $-\infty < x < \infty$ с вероятностью равной единице, функция $\rho(x)$ удовлетворяет условию нормированности:

$$\int_{-\infty}^{\infty} \rho(x) dx = 1. \quad (14)$$

Среднее значение этой функции можно вычислить по следующей формуле:

$$\bar{x} = \int_{-\infty}^{\infty} x \cdot \rho(x) dx. \quad (15)$$

Расчет среднего значения по формуле [14] возможен только при $\rho(x) = \text{const}$.

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i. \quad (16)$$

Таким образом, даже для оценки точности вычисления средней величины $\bar{x}$ необходимо учитывать форму статистического распределения исходных данных.

Для оценки статистических распределений принято используют значения их моментов. Моменты случайных величин, найденные без исключения систематических составляющих, называются *начальными*, а моменты для центрированных распределений – *центральными*.

Центральный момент k -го порядка для непрерывной случайной величины рассчитывается по формуле:

$$M_k = \int_{-\infty}^{+\infty} (x - \bar{x})^k \rho(x) dx, \quad (17)$$

при этом M_1 – математическое ожидание; $M_2 = \sigma^2$ – дисперсия (для конечной выборки – среднеквадратичное отклонение); M_3 характеризует асимметрию распределения, а безразмерный коэффициент асимметрии $A = \frac{M_3}{\sigma^3}$ есть третий центральный момент, поделенный на среднеквадратичное отклонение; M_4 характеризует протяженность распределения, отношение $E = \frac{M_4}{\sigma^4}$ – эксцесс, характеризующее остроту вершины распределения.

Наиболее широкое распространение при обработке экспериментальных данных получил центральный момент второго порядка M_2 который повсеместно используют для оценки погрешностей измерений. Для конечной выборки (конкретного числа отсчетов n) среднеквадратичное отклонение принято рассчитывать по следующей формуле:

$$D(x) = M_2 = \sigma^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2. \quad (18)$$

Эта формула для вычисления $\bar{x}$, не учитывает форму распределения и не является строгой.

Оценку асимметрии и эксцесса при конечной выборке n осуществляют по следующим формулам

$$A = \frac{\sum_{i=1}^n (x_i - \bar{x})^3}{n\sigma^3}, \quad (19)$$

$$E = \frac{\sum_{i=1}^n (x_i - \bar{x})^4}{n\sigma^4} - 3. \quad (20)$$

Число 3 в формуле (18) определяет эксцесс нормального распределения. Если эксцесс распределения отрицательный, то вершина функции распределения острее, чем у нормального распределения.

Определение этих характеристик распределений (моментов) называется *точечными оценками*, которые характеризуют распределение достаточно грубо.

Пусть в процессе экспериментальных измерений регистрируется сразу несколько случайных величин $x_1, x_2, \dots$, каждая из которых имеет свое среднее значение $\bar{x}_1, \bar{x}_2, \dots$ и дисперсию $\sigma_1, \sigma_2, \dots$. Многомерная случайная величина $X = (x_1, x_2, \dots)$ будет иметь многомерное распределение вероятностей $\rho(x_1, x_2, \dots)$ с условием нормировки

$$\int_{-\infty}^{\infty} \rho(x_1, x_2, \dots) dx_1 dx_2 \dots = 1 \quad (21)$$

Величины $x_1, x_2, \dots$ считаются статистически независимыми, если $\rho(x_1, x_2, \dots) = \rho(x_1)\rho(x_2)\dots$. Но эти величины могут быть статистически связаны, и для численной оценки этой связи двух случайных величин принято использовать *смешанный момент второго порядка* который называют *корреляционным моментом* или *ковариацией*.

Для расчета корреляционного момента K_{x_1, x_2} используется следующая формула:

$$K_{x_1, x_2} = \overline{(x_1 - \bar{x}_1)(x_2 - \bar{x}_2)} = \overline{(x_1 x_2)} - \bar{x}_1 \bar{x}_2, \quad (22)$$

где черта сверху означает статистическое усреднение соответствующего выражения. Корреляционный момент есть смешанная дисперсия двух величин, поэтому для расчета *коэффициента корреляции* R используется нормировка на дисперсию каждой из случайных величин:

$$R = \frac{1}{\sigma_1 \sigma_2} K_{x_1, x_2}. \quad (23)$$

Коэффициент корреляции меняется в пределах от -1 до +1 и определяет степень связи случайных величин.

Правила сложения случайных погрешностей. При сложении погрешностей в сложных измерительных системах необходимо учитывать, насколько эти погрешности статистически независимы или коррелированы. Как правило, при сложении погрешностей их делят на две группы: коррелированные и некоррелированные. Разделение на большее число групп (некоррелированные, слабо коррелированные, сильно коррелированные) слишком трудоемко. Поэтому на практике коррелированными погрешностями считаются те, у которых $|R| \geq 0,7$, остальные – некоррелированные.

Из приведенных выше формул может быть получено следующее правило сложения погрешностей двух случайных величин:

$$D(x_1 + x_2) = D(x_1) + D(x_2) + 2K_{x_1, x_2}, \quad (24)$$

а дисперсия их суммы:

$$\sigma = \sqrt{\sigma_1^2 + 2R\sigma_1\sigma_2 + \sigma_2^2}. \quad (25)$$

Таким образом, складывать среднеквадратические отклонения необходимо с учетом того, насколько случайные погрешности коррелированы.

Если:

- 1) $R = 1$, то $\sigma = |\sigma_1 + \sigma_2|$;
- 2) $R = -1$, то $\sigma = |\sigma_1 - \sigma_2|$.

3. ОБРАБОТКА РЕЗУЛЬТАТОВ ЭКСПЕРИМЕНТА СТАТИСТИЧЕСКИМИ МЕТОДАМИ. ПОСТРОЕНИЕ ФУНКЦИОНАЛЬНЫХ ЗАВИСИМОСТЕЙ

3.1. Основные законы распределения случайных величин, применяемые при обработке результатов эксперимента статистическими методами

Нормальное распределение

Нормальная модель распределения вероятностей играет исключительно важную роль в теории вероятностей и математической статистике. Главная особенность нормального распределения состоит в том, что оно является предельным, к которому приближаются другие распределения при соблюдении некоторых условий.

Нормальные распределения часто встречаются на практике в самых различных областях. Принято считать, что все ошибки измерений (измерение массы, размера) и многие другие случайные величины имеют нормальное распределение.

Нормальное распределение задается функцией плотности вероятности:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}, \quad (26)$$

где a – математическое ожидание случайной величины X , т.е. $M(X) = a$;

σ – среднее квадратичное отклонение случайной величины X , т.е. $\sqrt{D(X)} = \sigma$ ($D(X)$ – дисперсия случайной величины).

Из формулы (15) видно, что нормальная модель зависит от двух параметров a и σ , поэтому ее называют двухпараметрической моделью распределения.

Если случайная величина X имеет нормальное распределение с параметрами $M(X) = a$ и $\sigma = \sqrt{D(X)}$, то этот факт кратко записывают с помощью символической записи: случайная величина $X \in N(a, \sigma)$.

График функции плотности вероятности называют нормальной кривой или кривой Гаусса. Эта кривая изображена на рисунке 1.

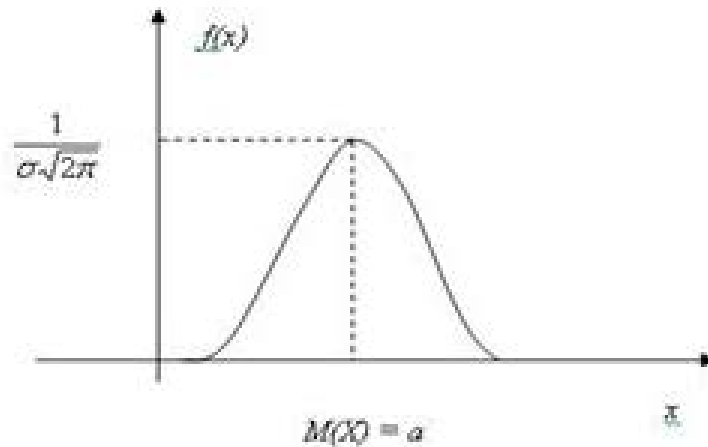


Рис.1. Кривая Гаусса

При этом необходимо учесть, что:

- $f(x)$ определена при всех $x \in \mathbb{R}$.
- кривая нормального распределения симметрична относительно прямой $x = a$.
- кривая Гаусса имеет максимум в точке $x = a$: $f(a) = \frac{1}{\sigma\sqrt{2\pi}}$.
- кривая Гаусса имеет две точки перегиба: $x_1 = a - \sigma$ и $x_2 = a + \sigma$.
- площадь, заключенная между кривой Гаусса и осью абсцисс, равна 1; между осью абсцисс, кривой Гаусса и прямыми $a \pm 2\sigma$ равна $\approx 0,95$.

При увеличении (уменьшении) параметра σ максимальная ордината уменьшается (увеличивается). Другими словами, параметр σ характеризует форму кривой, при неизменном положении центра кривой; так как площадь под кривой Гаусса всегда равна 1 ($\int_{-\infty}^{+\infty} f(x)dx = 1$), то, если σ увеличивается, то кривая становится плоско – вершинной, σ уменьшается – кривая Гаусса вытягивается вверх. Параметр σ иногда называют параметром масштаба.

Если изменять математическое ожидание a при неизменном σ , то кривая Гаусса будет смещаться вдоль оси абсцисс, т.е. параметр $a = M(X)$ характеризует положение кривой при неизменной форме. Иногда параметр a называют параметром сдвига.

Если случайная величина $X \in N(a, \sigma)$, то случайная величина $U = \frac{x-a}{\sigma}$ имеет нормальное распределение с параметром $M(U) = 0$ и $\sigma(U) = 1$, т.е. $U \in N(0,1)$. Поэтому случайную величину $U = \frac{x-a}{\sigma}$ называют нормированной или стандартизованной нормальной величиной. Плотность распределения вероятностей нормированной случайной величины U имеет вид:

$$f(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}. \quad (27)$$

Функция распределения СВ $X \in N(a, \sigma)$ имеет следующий вид:

$$F(x) = \int_{-\infty}^x f(x) dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(x-a)^2}{2\sigma^2}} dx. \quad (28)$$

Функция распределения нормализованной случайной величины

$$F(u) = P(U \leq u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^u e^{-t^2/2} dt. \quad (29)$$

Для облегчения вычисления вероятности попадания случайной величины $X \in N(a, \sigma)$ в интервал $[\alpha, \beta]$ вводится нормированная функция Лапласа:

$$\Phi(x) = \frac{2}{\sqrt{2\pi}} \int_0^x e^{-\frac{u^2}{2}} du. \quad (30)$$

$$\text{Тогда } P(\alpha < x < \beta) = F(\beta) - F(\alpha) \Leftrightarrow P(\alpha < x < \beta) = \frac{1}{2} \left(\Phi\left(\frac{\beta-a}{\sigma}\right) - \Phi\left(\frac{\alpha-a}{\sigma}\right) \right).$$

Используя нормированную функцию Лапласа, можно записать функцию распределения случайной величины $X \in N(a, \sigma)$ в виде:

$$F(x) = \frac{1}{2} + \frac{1}{2} \Phi\left(\frac{x-a}{\sigma}\right). \quad (31)$$

Распределение χ^2 (хи – квадрат)

Рассмотрим случайную величину Y , распределенную по нормальному закону $Y \in N(a, \sigma)$. Тогда случайная величина $U = \frac{Y-a}{\sigma} = \chi^2$ распределена по нормальному закону с параметрами $M(U) = 0$ и $\sigma(U) = 1$, т.е. $U \in N(0, 1)$.

Квадрат такой стандартизованной случайной величины:

$$U^2 = \left(\frac{Y-a}{\sigma}\right)^2 = \chi^2 \quad (32)$$

называется случайной величиной χ^2 (хи – квадрат) с одной степенью свободы.

Рассмотрим n независимых случайных величин $Y_1, Y_2, \dots, Y_n$, распределенных по нормальному закону с $M(Y_i) = a_i$ и средними квадратическими отклонениями $\sigma_i, i = \overline{1, n}$.

Образует для каждой из этих случайных величин стандартизованную случайную величину:

$$U_i = \frac{Y_i - a_i}{\sigma_i}, \quad i = \overline{1, n} \quad (33)$$

Сумма квадратов стандартизованных переменных:

$$\chi^2 = U_1^2 + U_2^2 + \dots + U_n^2 = \left(\frac{Y_1 - a_1}{\sigma_1}\right)^2 + \left(\frac{Y_2 - a_2}{\sigma_2}\right)^2 + \dots + \left(\frac{Y_n - a_n}{\sigma_n}\right)^2 \quad (34)$$

называется случайной величиной χ^2 с $\nu = n$ степенями свободы.

Плотность распределения случайной величины χ^2 имеет вид:

$$f(\chi^2) = \begin{cases} 0, \\ \frac{1}{2^{\frac{\nu}{2}} \Gamma(\frac{\nu}{2})} (\chi^2)^{\frac{\nu}{2}-1} e^{-\frac{\chi^2}{2}}, & \text{если } \chi^2 \geq 0 \end{cases} \quad (35)$$

Итак, распределение χ^2 зависит от одного параметра ν – числа степеней свободы.

Функция распределения χ^2 имеет вид:

$$F(\chi^2) = P(\chi^2 < \chi_0^2) = \begin{cases} 0, \\ \frac{1}{2^{\frac{\nu}{2}} \Gamma(\frac{\nu}{2})} \int_0^{\chi^2} (\chi^2)^{\frac{\nu}{2}-1} e^{-\frac{\chi^2}{2}} d(\chi^2), & \text{если } \chi^2 \geq 0 \end{cases} \quad (36)$$

На рис. 2 изображен график плотности вероятности и функции χ^2 – распределения для разных степеней свободы.

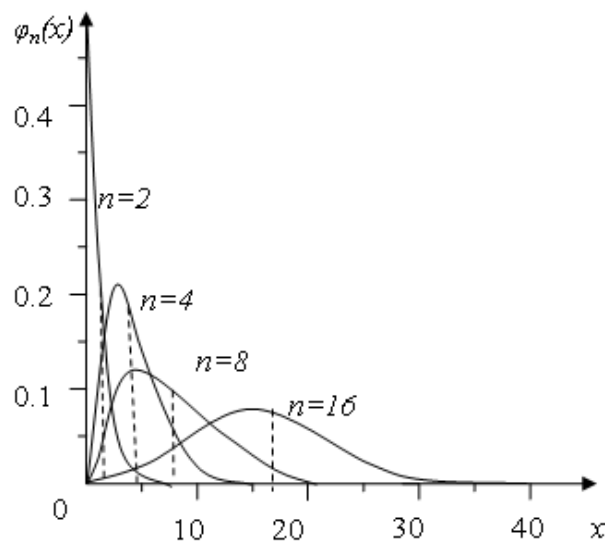


Рис.2. График плотности вероятности и функции χ^2 – распределения

На практике, как правило, используются не $f(\chi^2)$ и $F(\chi^2)$, а квантили χ^2 – распределения $\chi^2_{\alpha, \nu}$. Квантилем $\chi^2_{\alpha, \nu}$, отвечающим заданному уровню вероятности α , называется такое значение $\chi^2 = \chi^2_{\alpha, \nu}$, при котором

$$P(\chi^2 > \chi^2_{\alpha, \nu}) = \int_{\chi^2_{\alpha, \nu}}^{\infty} f(\chi^2) d(\chi^2) = \alpha \quad (37)$$

Распределение Стьюдента

Распределение Стьюдента (t – распределение) имеет важное значение при статистических вычислениях, связанных с нормальным законом, а именно тогда, когда среднее квадратическое отклонение σ неизвестно и подлежит определению по опытным данным.

Пусть $Y, Y_1, Y_2, \dots, Y_n$ – независимые случайные величины, имеющие нормальное распределение с параметрами $M(Y) = M(Y_i) = 0$ и $\sigma_Y = \sigma_{Y_i} = 1$, $i = \overline{1, n}$.

Случайная величина

$$t = \frac{Y}{\sqrt{\frac{1}{n} \sum_{i=1}^n Y_i^2}} = \frac{Y}{\sqrt{\frac{1}{n} \chi_n^2}}, \quad (38)$$

являющаяся функцией нормально распределенных случайных величин, называется безразмерной дробью Стьюдента.

Плотность распределения случайной величины t имеет вид:

$$f(t) = S(t, \nu) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\pi\nu} \Gamma(\frac{\nu}{2})} \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}}, \quad -\infty < t < +\infty \quad (39),$$

где ν – число слагаемых в подкоренном выражении дроби Стьюдента, т.е. $\nu = n$. Такое обозначение числа степеней свободы общепринято в математической статистике.

Из формулы (29) видно, что распределение СВ t зависит только от одного параметра – числа степеней свободы ν , равного числу слагаемых в подкоренном выражении дроби Стьюдента (28).

Известно, что математическое ожидание и дисперсия СВ t соответственно равны $M(t) = 0$; $D(t) = \frac{\nu}{\nu-2}$; ($\nu > 2$).

При увеличении числа степеней свободы ν график плотности распределения Стьюдента приближается к кривой Гаусса.

В статистических расчетах используются квантили t – распределения $t_{\frac{\alpha}{2}; \nu}$. Значения квантилей находятся из решения уравнения:

$$P(|t| > t_{\frac{\alpha}{2}; \nu}) = 2 \int_{t_{\frac{\alpha}{2}; \nu}}^{\infty} f(t) dt = \alpha \quad (40)$$

3.2. Построение функциональных зависимостей при многократных измерениях

Построение статистических распределений и критерии оценивания Пусть необходимо построить зависимость $y=f(x)$ в том случае, если каждому значению x соответствует набор значений y_i . В практических задачах в результате эксперимента имеется набор точек с координатами x_i, y_i как это показано на рис.3.

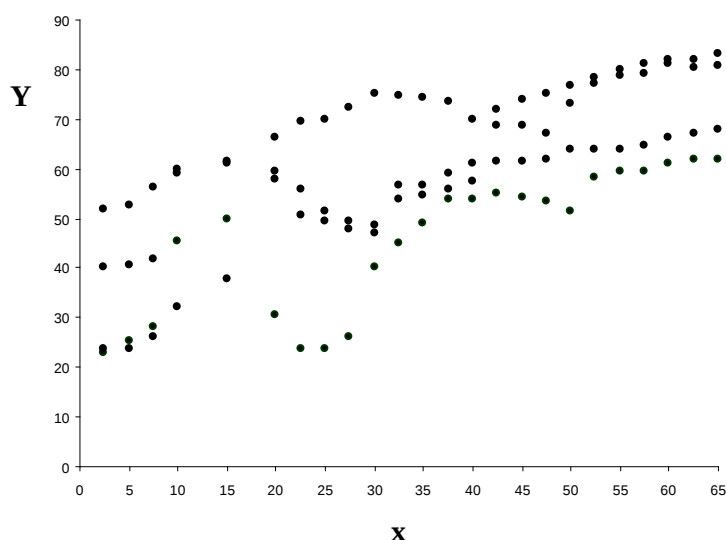


Рис.3. Результаты эксперимента

При построении математических моделей состояния водной среды необходимо оценивать качество воды с помощью одной или нескольких числовых величин, т.е. провести формализацию параметра "качество воды". Качество воды представляет собой совокупность физических, химических и биологических показателей, определяющих степень ее пригодности для конкретных видов водопользования и отвечающих требованиям охраны окружающей среды. Оно характеризуется составом и количеством растворенных и взвешенных в воде веществ, содержанием биомассы и микроорганизмов, температурой и другими физико-химическими характеристиками. Действительное число компонент экосистемы может быть бесконечно велико, можно выделить несколько десятков и даже сотен различных параметров. Поэтому

индивидуальный учет в модели каждой из них невозможен и бесполезен. В зависимости от характера задачи и на основе имеющихся результатов исследования экосистемы, большей частью компонент пренебрегают. Как правило, реальные методы используют несколько простейших параметров: растворенный кислород, биологическое потребление кислорода, концентрация примеси и т.п.

Для того, чтобы построить вероятностную модель эксперимента, необходимо знать, по какому закону распределены экспериментальные данные (см. пункт 3.1). Для оценки закона распределения служат различные критерии, позволяющие с определенной вероятностью по конечной выборке оценить распределение генеральной совокупности данных.

В общем случае для оценки статистического распределения предлагается действовать в следующем порядке:

1. Построить гистограмму с последующим нахождением центра и симметрированием и сглаживанием.
2. Провести аппроксимацию полученной гистограммы к соответствующему распределению.
3. Проверить построенное распределение по критериям согласия.

При построении гистограммы (рис.5) важно выбрать необходимое число интервалов m , в которые попадает случайная величина x (обычно $6 < m < 20$). Затем по вертикальной оси гистограммы откладывается количество

отсчетов f_i или нормированная частота $F_i = \frac{f_i}{n}$, попадающих в i -й интервал (n - общий объем случайной выборки).

Если при малой выборке n выбрано слишком большое число столбцов m , то гистограмма может оказаться очень неравномерной (см. рис. 5.а) и судить о форме статистического распределения невозможно. Если исходить из того, что генеральная совокупность имеет гладкую форму распределения, то выбросы и провалы в гистограмме можно считать случайным шумом вызванным попаданием m в тот или иной столбец. Тогда задачу выбора числа m можно считать задачей оптимальной фильтрации. m должно быть таким, чтобы максимально сглаживать но при этом минимально искажать случайную зависимость, описывающую распределение.

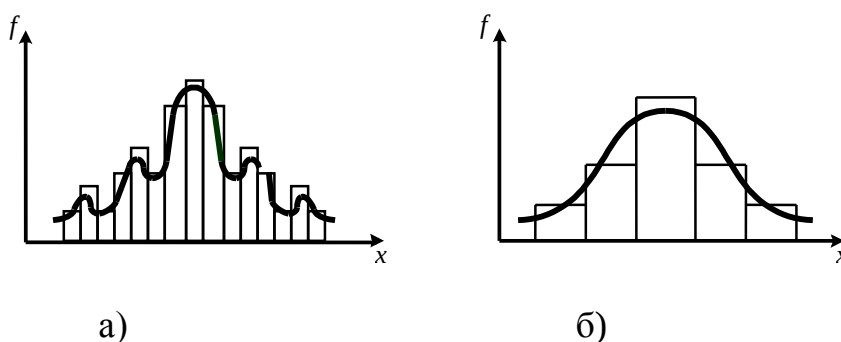


Рис.4. Гистограмма распределения в общем виде

Для выбора числа m в литературе предлагается большое количество формул, мы приведем лишь три из них:

- 1) формула Старджеса: $m = \log_2 n + 1 = 3,3 \lg n + 1$;
- 2) $m = 5 \lg n$;
- 3) $m = \sqrt{n}$.

Для того чтобы определить однозначный центр распределения (предполагая симметричность распределения) число столбцов должно быть нечетным. Для его нахождения рассчитывают медиану или 50% – квантиль (половина выборки – больше, а половина – меньше этого числа). Медиана может не совпадать со средним значением $\bar{X}$.

Симметрирование гистограммы заключается в переносе некоторого числа отсчетов f_i из данного столбца в симметричный с ним столбец для их выравнивания.

Для идентификации распределения $\rho(x)$... используются критерии согласия. Наиболее распространен критерий χ^2 (критерий Пирсона). Он используется при проверке гипотез о принадлежности выборки к определенной генеральной совокупности.

Для этого выбирается χ^2 по формуле:

$$\chi^2 = \sum_{i=1}^m \frac{(f_i - M_i)^2}{M_i} , \quad (41)$$

где f_i – число значений случайной величины, попавшей в i -й интервал (высота столбца в гистограмме на рис.5); M_i – значение частоты в выбранной модели распределения.

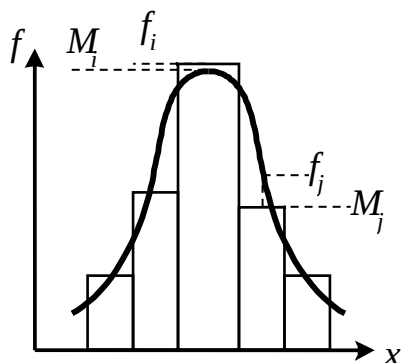


Рис.5. Гистограмма для идентификации распределения

Критерий Пирсона дает $\chi^2 = 0$, если в центрах всех столбцов гистограммы выполняется равенство $f_i - M_i = 0$.

Критерий χ^2 , кроме того, позволяет произвести сравнение двух моделей распределений в том случае, если для них используется разное число столбцов.

В общем случае χ^2 существенно возрастает с увеличением числа столбцов, но для квантилей распределения χ^2 имеются специальные таблицы, в которых используется число степеней свободы

$$\nu = m - 1 - r \quad (42)$$

где r – число параметров, необходимых для описания распределения.

Для нормального распределения достаточно определить первые два момента, поэтому $r = 2$. В отличие от распределения Стюдента, рассчитанного для нормальной генеральной совокупности, критерий χ^2 может использоваться для любых распределений.

Второй пример широко используемого критерия – *критерий Колмогорова-Смирнова*. Он позволяет сравнить две независимые выборки и ответить на вопрос, относятся ли они к одной генеральной совокупности. Его удобство состоит в том, что для использования нет необходимости строить гистограмму.

В качестве статистики служит наибольшая по модулю разница между нормированными частотами в двух выборках

$$D = \max \left(\frac{f_1}{n_1} - \frac{f_2}{n_2} \right). \quad (43)$$

$$\text{При } n_1 = n_2 = n \text{ получим } D = \max \left(\frac{f_1 - f_2}{n} \right).$$

Величина D табулирована и для $(n_1 + n_2) > 40$ можно задать граничное значение

$$D_{\text{кр}} = k \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}, \quad (44)$$

где k – постоянная, зависящая от уровня значимости $q = 1 - p$ (вероятность ошибки при идентификации распределения). Для этой величины имеется приближенная аналитическая формула $k = \sqrt{-0,5 \ln (q/2)}$. Если выборка объемом n сопоставляется с аналитической моделью распределения $\rho(x)$

($n_2 \rightarrow \infty$), то $D_{\text{кр}} = \frac{k}{\sqrt{n}} = \sqrt{-0,5 \ln (q/2)/n}$. Отсюда может быть оценена вероятность ошибки идентификации распределения по выборке объемом n :

$$p \geq 2e^{-2nD_{cp}^2} . \quad (45)$$

Для оценки разницы между дисперсиями двух конечных выборок из нормальной совокупности используется *критерий Фишера*. В качестве значения критерия берут отношение большей дисперсии к меньшей

$$F = \frac{S_1^2}{S_2^2} , \quad (46)$$

где $S_{1,2}$ — среднеквадратичные отклонения.

Вычисленное таким образом значение критерия сравнивается с табулированным значением в соответствии с уровнем значимости q для степеней свободы m и n распределений (m и n — величины выборок).

4. ПЛАНИРОВАНИЕ ЭКСПЕРИМЕНТА. ОБЩИЕ ПОЛОЖЕНИЯ

4.1. История планирования эксперимента

Планирование эксперимента возникло в 50-х годах 20 века из потребности устранить или хотя бы уменьшить систематические ошибки в сельскохозяйственных исследованиях путем рандомизации условий проведения эксперимента. Процедура планирования оказалась направленной не только на уменьшение дисперсии оцениваемых параметров, но также и на рандомизацию относительно сопутствующих, спонтанно изменяющихся и неконтролируемых переменных [4]. В результате удалось избавиться от смещения в оценках.

С 1918 г. Р. Фишер начал свою известную серию работ на Рочемстедской агробиологической станции в Англии. В 1935 появилась его монография «Design of Experiments», давшая название всему направлению. В 1942 А. Кишен рассмотрел планирование эксперимента по латинским кубам, которое явилось дальнейшим развитием теории латинских квадратов. Затем Р. Фишер независимо опубликовал сведения об ортогональных гипергреко-латинских кубах и гиперкубах. Вскоре после этого в 1946 г. Р. Рао рассмотрел их комбинаторные свойства. Дальнейшему развитию теории латинских квадратов посвящены работы Х. Манна (1947–1950 гг.).

Первое глубокое математическое исследование блок-схемы выполнено Р. Боузом в 1939 г. Вначале была разработана теория сбалансированных неполноблочных планов (BIB-схемы). Затем Р. Боуз, К. Нер и Р. Рао обобщили эти планы и разработали теорию частично сбалансированных неполноблочных планов (PBIB-схемы). С тех пор изучению блок-схем уделяется большое внимание как со стороны специалистов по планированию эксперимента

(Ф. Йетс, Г. Кокс, В. Кохрен, В. Федерер, К. Гульден, О. Кемптгорн и другие), так и со стороны специалистов по комбинаторному анализу (Боуз, Ф. Шимамото, В. Клатсворси, С. Шрикханде, А. Гофман и др.).

Исследования Р. Фишера знаменуют начало первого этапа развития методов планирования эксперимента. Фишер разработал метод факторного планирования. Йетс предложил для этого метода простую вычислительную схему. Факторное планирование получило широкое распространение. Особенностью факторного эксперимента является необходимость ставить сразу большое число опытов.

В 1945 г. Д. Финни ввел дробные реплики от факторного эксперимента. Это позволило сократить число опытов и открыло дорогу техническим приложениям планирования. Другая возможность сокращения необходимого числа опытов была показана в 1946 г. Р. Плакеттом и Д. Берманом, которые ввели насыщенные факторные планы.

Г. Хотеллинг в 1941 г. предложил находить экстремум по экспериментальным данным с использованием степенных разложений и градиента. Следующим важным этапом было введение принципа последовательного шагового экспериментирования. Этот принцип, высказанный в 1947 г. М. Фридманом и Л. Сэвиджэм, позволил распространить на экспериментальное определение экстремума – итерацию.

Чтобы построить современную теорию планирования эксперимента, не хватало одного звена – формализации объекта исследования. Это звено появилось в 1947 г. после создания Н. Винером теории кибернетики. Кибернетическое понятие «черный ящик» играет в планировании важную роль.

В 1951 г. работой американских ученых Дж. Бокса и К. Уилсона начался новый этап развития планирования эксперимента. В ней сформулирована и доведена до практических рекомендаций идея последовательного экспериментального определения оптимальных условий проведения процессов с использованием оценки коэффициентов степенных разложений методом наименьших квадратов, движение по градиенту и отыскание интерполяционного полинома в области экстремума функции отклика (почти стационарной области).

В 1954—1955 гг. Дж. Бокс, а затем П. Юл. показали, что планирование эксперимента можно использовать при исследовании физико-химических процессов, если априори высказаны одна или несколько возможных гипотез. Направление получило развитие в работах Н. П. Клепикова, С. Н. Соколова и В. В. Федорова в ядерной физике.

Третий этап развития теории планирования эксперимента начался в 1957 г., когда Бокс применил свой метод в промышленности. Этот метод стал называться «эволюционным планированием». В 1958 г. Г. Шиффе предложил новый метод планирования эксперимента для изучения физико-химических диаграмм состав – свойство под названием «симплексная решетка».

Развитие теории планирования эксперимента в СССР отражено в работах В. В. Налимова, Ю. П. Адлера, Ю. В. Грановского, Е. В. Марковой, В. Б. Тихомирова.

4.2. Основные понятия и определения теории планирования эксперимента

Теория планирования эксперимента (ТПЭ) охватывает практически все встречающиеся на практике варианты исследования объектов. Основными типовыми задачами экспериментального исследования являются:

- 1) поиск значений параметров системы, обеспечивающих достижение оптимального значения показателя качества исследуемого объекта при известных ограничениях на значения этих параметров.

- 2) приближенное аналитическое описание функциональной связи показателей качества с параметрами системы по результатам проведенного экспе-

римента, при этом аналитическое описание объекта можно использовать для предварительного исследования вариантов построения системы или в интересах построения модели старшей системы, включающей данный объект на правах элемента;

3) оценка дифференциального влияния уровней параметров системы на показатель качества (если параметры системы являются по своей природе качественными или когда количественные параметры могут принимать не большое число различных значений).

Исследователь на этапе планирования эксперимента должен владеть следующей информацией :

- класс, к которому можно отнести моделируемую систему (статическая или динамическая, детерминированная или стохастическая и др.);
- режим работы (стационарный или нестационарный);
- промежуток времени, в течение которого следует наблюдать за функционированием системы;
- объём испытаний, позволяющий обеспечить требуемую точность оценок (в статистическом смысле) исследуемых характеристик системы.

Таким образом, основные цели планирования модельных экспериментов заключаются в:

- сокращении общего объёма испытаний при соблюдении требований к достоверности и точности их результатов;
- повышении информативности каждого из экспериментов в отдельности.

Поиск плана эксперимента проводится в факторном пространстве. Факторное пространство (ФП) – это множество внешних и внутренних параметров модели, значения которых исследователь может контролировать в ходе подготовки и проведения эксперимента.

Значения факторов обычно называются уровнями.

Если при проведении эксперимента исследователь может изменять уровни факторов, эксперимент называется *активным*, в противном случае – *пассивным*.

Каждый из факторов имеет верхний и нижний уровни, расположенные симметрично относительно некоторого нулевого уровня. Точка в ФП, соответствующая нулевым уровням всех факторов, называется *центром плана*.

Интервалом варьирования фактора называется некоторое число, прибавление которого к нулевому уровню даёт верхний уровень, а вычитание – нижний. Как правило, план эксперимента строится относительно одного (основного) выходного скалярного параметра Y , который называется *наблюдаемой переменной*.

Если моделирование используется как инструмент принятия решения, то в роли наблюдаемой переменной выступает *показатель эффективности*. При этом предполагается, что значение наблюдаемой переменной, полученное в ходе эксперимента, складывается из двух составляющих:

$$Y = f(x) + e(x), \quad (47)$$

где $f(x)$ – функция отклика (неслучайная функция факторов),
 $e(x)$ – ошибка эксперимента (случайная величина).

Дисперсия Dy наблюдаемой переменной, которая характеризует точность измерений, равна дисперсии ошибки опыта, т.е. $Dy = De$.

Dy называют *дисперсией воспроизводимости эксперимента*. Она характеризует качество эксперимента. Эксперимент называется *идеальным* при $Dy = 0$.

Существует два основных варианта постановки задачи планирования имитационного эксперимента:

1. Из всех допустимых требуется выбрать такой план, который позволил бы получить наиболее достоверное значение функции отклика $f(x)$ при фиксированном числе опытов;

2. Из всех допустимых требуется выбрать такой план, при котором статистическая оценка функции отклика может быть получена с заданной точностью при минимальном объёме испытаний.

Решение задачи планирования в первой постановке называется *стратегическим планированием*, во второй – *тактическим планированием*.

При проведении опытных исследований различают *пассивный и активный эксперимент*.

Методология *пассивного экспериментирования* предполагает проведение большой серии опытных исследований с поочередным варьированием значений входных переменных x и анализом результатов измерений выходной переменной y (лабораторный эксперимент или эксперимент на пилотной установке).

К *пассивному эксперименту* принято относить также и сбор опытных данных в режиме эксплуатации промышленной установки (*промышленный эксперимент*).

В данном случае вид уравнения регрессии необходимо определять по характеру изменения переменных на графике эмпирической линии регрессии, полученной по выборке экспериментальных данных.

Для решения этой задачи при одной входной переменной x существуют эффективные методы, в которых предусматривается преобразование системы координат как для входной (x), так и для выходной переменной (y). При большем числе входных переменных ($x_1, \dots, x_m$) надёжных методов определения вида уравнения регрессии (вида эмпирической модели) в настоящее время не существует.

Активный эксперимент проводится по заранее составленному плану, в соответствии с которым ставится задача не только определения оптимальных условий проведения эксперимента, но и оптимизации процесса (оптимальное планирование эксперимента).

При определении оптимальных условий проведения процесса с использованием эмпирических моделей (например, методом Бокса-Вильсона)

выходная переменная y является критерием оптимальности или целевой функцией.

Активный эксперимент планируется таким образом, чтобы упростить обработку его результатов методами регрессионного и корреляционного анализа.

К достоинствам активного экспериментирования относятся:

- возможность предсказания количества опытов, которые следуют провести;
- определение точек факторного пространства, где следует проводить опыты;
- отсутствие проблем, связанных с выбором вида уравнения регрессии;
- возможность определения оптимальных параметров процесса экспериментально-статистическим методом;
- сокращение объема опытных исследований.

4.3. Основы планирования многофакторного эксперимента

В общем случае объект исследования представляется в виде структурной схемы, основанной на принципе «черного ящика» (рис.6).

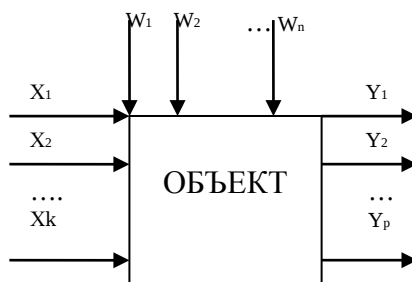


Рис.6. Структурная схема объекта исследования

При этом имеются следующие группы параметров:

- 1) управляющие (входные) x , которые называются *факторами*;
- 2) выходные параметры y , которые называются *параметрами состояния*;
- 3) w – возмущающие воздействия.

Предполагается, что возмущающие воздействия не поддаются контролю и либо являются случайными, либо меняются во времени. Каждый фактор x_i имеет область определения, которая должна быть установлена до проведения эксперимента. Комбинацию факторов можно представить как точку в многомерном пространстве, характеризующую состояние системы.

На практике целью многофакторного эксперимента является установление зависимости

$$Y = f(x_1, x_2, \dots, x_k) \quad (48)$$

описывающей поведение объекта.

Целью эксперимента может быть, например, построение зависимости (48) при минимальном количестве измерений значений управляющих параметров x_i' .

На первом этапе планирования эксперимента необходимо выбрать область определения факторов x_i . Выбор этой области производится исходя из априорной информации. Значения x_i называются уровнями управляющего параметра.

Основные требования, предъявляемые к факторам:

- управляемость,
- однозначность,
- совместимость,
- независимость

Рассмотрим эти требования более подробно. При планировании эксперимента факторы должны быть *управляемыми*. Это значит, что экспериментатор, выбрав нужное значение фактора, может его поддерживать постоянным в течение всего опыта, т. е. может управлять фактором. В этом состоит особенность «активного» эксперимента. Планировать эксперимент можно только в том случае, если уровни факторов подчиняются воле экспериментатора.

Точность замера факторов должна быть по возможности более высокой. Степень точности определяется диапазоном изменения факторов. Факторы должны быть непосредственными воздействиями на объект. Факторы должны быть *однозначны*. Сложно управлять фактором, который является функцией других факторов. Но в планировании могут участвовать сложные факторы, такие как соотношения между компонентами, их логарифмы и т. п. При планировании эксперимента обычно одновременно изменяется несколько факторов. Поэтому очень важно сформулировать требования, которые предъявляются к совокупности факторов.

Совместимость факторов означает, что все их комбинации осуществимы и безопасны. Это очень важное требование. Несовместимость факторов может наблюдаться на границах областей их определения. Избавиться от нее можно сокращением областей. Положение усложняется, если несовместимость проявляется внутри областей определения. Одно из возможных решений – разбиение на подобласти и решение двух отдельных задач.

При планировании эксперимента важна *независимость* факторов, т. е. возможность установления фактора на любом уровне вне зависимости от уровней других факторов. Если это условно невыполнимо, то невозможно планировать эксперимент. Таким образом, необходимо выполнение еще одного требования – *отсутствие корреляции между факторами*. Требование некоррелированности не означает, что между значениями факторов нет никакой связи. Достаточно, чтобы связь не была линейной.

Необходимость введения сложных факторов возникает при желании представить динамические особенности объекта в статической форме. Факто-

ры в эксперименте подразделяются на качественные и количественные. Качественные факторы можно квантифицировать или приписать им числовые обозначения, тем самым перейти к количественным значениям. В дальнейшем будем считать, что все факторы являются количественными и представлены непрерывными величинами (если другое не оговорено особо).

Область планирования задается интервалами возможного изменения факторов $v_{i,min} < v_i < v_{i,max}$ для $i = 1, 2, \dots, k$, где k – количество факторов. В теории ПЭ часто используют нормализацию факторов, т.е. преобразование натуральных значений факторов в безразмерные (кодированные) величины. Переход к безразмерным значениям x_i задается преобразованием

$$x_i = (v_i - v_{i0}) / \Delta v_i, \quad (49)$$

где v_i – натуральное значение фактора, v_{i0} – натуральное значение основного уровня фактора, соответствующее нулю в безразмерной шкале, Δv_i – интервал варьирования.

Совокупность основных уровней всех факторов представляет собой точку в пространстве параметров, называемую центральной точкой плана или центром эксперимента. С геометрической точки зрения нормализация факторов равноценна линейному преобразованию пространства факторов, при котором проводятся две операции: перенос начала координат в точку, соответствующую значениям основных уровней факторов; сжатие – растяжение пространства в направлении координатных осей.

Активный эксперимент включает: систему воздействий, при которых воспроизводится функционирование объекта; регистрацию отклика объекта. *План эксперимента* задает совокупность данных, определяющих количество, условия и порядок реализации опытов. *Опыт* составляет элементарную часть эксперимента и предусматривает воспроизведение исследуемого явления в конкретных условиях с последующей регистрацией результата. В условиях случайности в одних и тех же условиях проводятся параллельные (повторные) опыты в интересах получения статистически устойчивых результатов. Опыт u предполагает задание конкретных значений факторам $\mathbf{v}_u = v_{1u}, v_{2u}, \dots, v_{ku}$, а совокупность значений факторов во всех N точках плана эксперимента образует матрицу плана

$$\begin{array}{c} v_{11}, v_{21}, \dots, v_{k1} \\ v_{12}, v_{22}, \dots, v_{k2} \\ \dots \dots \dots \\ v_{1N}, v_{2N}, \dots, v_{kN} \end{array} \quad (50)$$

Строки матрицы соответствуют опытам, столбцы – факторам, элемент матрицы v_{iz} задает значение z -го фактора в i -м опыте.

Вектор u называется *откликом*. В ТПЭ обычно изучается ситуация, в которой вектор отклика u состоит из одного элемента u . При наличии нескольких составляющих вектора u , каждую из них можно исследовать отдельно. Зависимость отклика от факторов носит название *функции отклика*, а

геометрическое представление функции отклика – *поверхности отклика*. Функция отклика рассматривается как показатель качества или эффективности объекта. Этот показатель является функцией от параметров – факторов. На практике широкое распространение получили простые функции вида $M\{y'\} = bf(v)$, где $b=(b_0, b_1, \dots, b_h)$ – вектор неизвестных параметров модели размерности $h+1$, $f(v)=(f_0(v), f_1(v), \dots, f_h(v))$ – вектор заданных базисных функций, $M\{y'\}$ – математическое ожидание функции отклика. Такое представление функции отклика соответствует линейной по параметрам модели регрессионного анализа, т.е. функция отклика есть линейная комбинация базисных функций от факторов.

Вследствие влияния на результаты экспериментов случайных воздействий истинные значения коэффициентов можно определить только приближенно. Оценку вектора неизвестных параметров b находят по результатам экспериментов, в ходе которых получают значения y_u при заданных значениях факторов v_u . Эти оценки обычно рассчитываются с помощью метода наименьших квадратов (МНК) на основе выборок значений факторов и откликов системы на воздействия [8]. В качестве оценки вектора b выбирается такое значение, которое минимизирует

$$\frac{1}{N} \sum_{u=1}^N (y'_u - y_u)^2, \quad (51)$$

где y'_u – вычисленное на модели значение функции отклика в u -й точке факторного пространства. Приравнявая нулю частные производные от данной квадратичной формы, взятые по переменным $b_0, b_1, \dots, b_h$, можно получить систему уравнений вида

$$\frac{1}{N} \sum_{u=1}^N (y'_u - y_u) f(v_i) = 0 \quad (52)$$

где $i=0, 1, 2, \dots, h$. Значение b находят путем решения этой системы уравнений. Решение системы возможно при линейной независимости базисных функций.

Если не принимать специальных мер, то оценки коэффициентов станут взаимозависимыми, и полученное выражение для функции отклика можно рассматривать только как интерполяционную формулу, что затрудняет ее физическую интерпретацию и последующие расчеты. Однако, формируя специальным образом матрицу плана, можно получить независимые значения. И эти величины будут характеризовать вклад каждого фактора в значение функции отклика.

В большинстве случаев вид функции отклика y' , получаемый из теоретических соображений, является сложным для практического применения, а при неполном знании объекта вообще неизвестен. По данным причинам функцию целесообразно представить в универсальном, удобном для практи-

ческого применения виде, чему соответствует представление в виде полинома. Тогда системой базисных функций является совокупность степенных функций с целыми неотрицательными значениями показателей степени. Полиномиальная форма представления функции отклика примет вид

$$y=b_0+b_1x_1+ b_2x_2 \quad (53)$$

или

$$y=b_0+b_1x_1+ b_2x_2+b_{11}x_1^2+ b_{22}x_2^2 +b_{12}x_1 x_2 \quad (54)$$

Такая функция отклика линейна относительно неизвестных коэффициентов и будет полностью определена, если заданы степень полинома и коэффициенты. Степень полинома обычно задается исследователем априорно и уточняется в ходе исследования. На практике наибольшее распространение получили полиномы первого и второго порядка, соответственно линейные и квадратичные модели. Коэффициенты полинома принято называть *эффектами факторов*.

Иногда функцию отклика целесообразно представить в другом виде, например, в виде степенной функции, так как достижение заданной точности требует применения полинома высокого порядка. Однако использование функций, нелинейных относительно неизвестных параметров, усложняет вычисления, затрудняет оценку их свойств. В некоторых случаях задачу можно упростить путем искусственного преобразования нелинейной функции в линейную. При этом требуется соответствующее преобразование и результатов экспериментов.

Применение ТПЭ основано на ряде допущений, а именно [8,9]:

1. Функция отклика содержит в своем составе неслучайную и случайную составляющую. В процессе проведения эксперимента некоторые показатели носят случайный характер, что требует многократного повторения опытов в одних и тех же условиях для получения статистически устойчивых результатов, а получаемые оценки показателей должны обладать свойствами состоятельности, эффективности, несмещенности и достаточности. Оценки типовых показателей формируются путем усреднения результатов наблюдений. Поэтому при достаточно большом количестве наблюдений можно считать, что случайная составляющая ϵ распределена по нормальному закону с нулевым математическим ожиданием, что позволяет получить несмещенную оценку математического ожидания функции отклика в конкретной точке плана. Результаты, полученные путем усреднения повторных опытов в каждой точке плана, представляют собой независимые, нормально распределенные случайные величины;

2. Факторы $v_1, v_2, \dots, v_k$ измеряются с пренебрежимо малой ошибкой по сравнению с ошибкой в определении величины y (учет помех в задании факторов приводит к трудно разрешимым проблемам в оценке коэффициентов функции отклика). Ошибка в определении значения функции отклика объяс-

няется не столько погрешностью измерений, сколько влиянием на результат работы системы неучтенных или случайных факторов, например различиями в формируемой последовательности случайных чисел при статистическом моделировании;

3. Дисперсии среднего значения функции отклика в различных точках равны друг другу (выборочные оценки дисперсии однородны). Это означает, что при многократных повторных наблюдениях над величиной y_u при некотором наборе значений $v_{1u}, v_{2u}, \dots, v_{ku}$, получаемая оценка дисперсии среднего значения не будет отличаться от оценки дисперсии, полученной при многократных наблюдениях для любого другого набора значений независимых переменных $v_{1s}, v_{2s}, \dots, v_{ks}$.

Указанные допущения позволяют использовать для расчетов коэффициентов полинома МНК, который дает эффективные и несмещенные оценки коэффициентов и обеспечивает простоту проведения самих расчетов. Применение МНК, вообще говоря, не требует соблюдения нормального распределения результатов наблюдения. Этот метод в любом случае дает решение, минимизирующее сумму квадратов отклонений результатов наблюдения от значений функции отклика. Допущение о нормальном распределении используется при проведении различного рода проверок, например, при проверке адекватности функции отклика и экспериментальных данных. Естественно, что точность оценок коэффициентов функции отклика повышается с увеличением числа опытов, по которым вычисляются коэффициенты.

Планируя эксперимент, стремятся получить линейную модель, однако в выбранных интервалах варьирования априори не известно, что линейная модель адекватно описывает поведение системы. Нелинейность связана со смешанным взаимодействием.

С ростом числа факторов число возможных взаимодействий возрастает. Например, для факторного эксперимента $N = 2^3$ кроме x_0, x_1, x_2, x_3 в матрице планирования появляются столбцы $x_1x_2, x_1x_3, x_2x_3, x_1x_2x_3$. Всего в матрице планирования оказывается восемь столбцов, следовательно, необходимо определять восемь коэффициентов. Все восемь коэффициентов необходимо определять в том случае, если учитывать смешанное взаимодействие. Если же модель задается в виде гиперплоскости (линейная модель), то достаточно определить четыре коэффициента: x_0, x_1, x_2, x_3 . Полный факторный эксперимент оказывается избыточным и возникает выбор:

1. Построить гиперплоскость по четырем экспериментам, а остальные четыре опыта использовать для оценки погрешности.

2. Провести эксперимент, состоящий из 4-х опытов, то есть реализовать экономный план эксперимента.

Таким образом, в отличие от модели гиперболоида, которая требует определение 2^k неизвестных коэффициентов, модель гиперплоскости, содер-

жит $k+1$ коэффициент и требует соответствующего числа опытов, то есть полный факторный план (ПФП) для модели гиперплоскости сильно избыточен.

Для построения гиперплоскости, достаточно использовать лишь некоторую часть из ПФП. Эту часть в теории планирования эксперимента называют *дробной* репликой или дробным факторным планом (ДФП). Если дробление ПФП производится последовательным делением числа опытов на 2, то реплику называют *регулярной*. Число p последовательного деления называют *дробностью* реплики.

Число опытов регулярного ДФП равняется $n = 2^{k-p}$. При $p = 1$ ДФП называют полурепликой (или 1/2 реплика), при $p = 2$ – 1/4 реплика и т.д.

Соответствующее число опытов и параметров планирования приведены в таблице 2.

Таблица 2

Число опытов и параметров планирования

Число факторов, k	Число коэфф. модели, $k+1$	Число опытов ПФП	Вид плана	Число опытов плана	Избыточность
2	3	4	ПФП	4	1
3	4	8	Полуреплика	4	0
4	5	16	Полуреплика	8	3
5	6	32	Четвертьреплика	8	2
6	7	64	1/8 реплика	8	1
7	8	128	1/16 реплика	8	0
8	9	256	1/16 реплика	16	7

Для составления планов-таблиц регулярных дробных реплик часто используют так называемое правило двоичного кода. Оно гласит, что для модели в виде гиперболоида знаки “+” и “–” в столбцах плана должны чередоваться по правилу чередования двоичных чисел в разряде двоичного кода, то есть в столбце x_1 – через 1, в столбце x_2 – через 2, в столбце x_3 – через 4, в столбце x_k – через 2^{k-1} .

4.4. Критерии оптимальности и типы планов

В настоящее время используется свыше двадцати различных критериев оптимальности планов, которые подразделяются на две основные группы.

К первой группе относят критерии, связанные с ошибками оценок коэффициентов, а ко второй – с ошибкой оценки поверхности отклика [2, 3, 6].

Критерии *первой группы* представляют интерес для задач оптимизации, выделения наиболее значимых параметров на начальных этапах решения оптимизационных задач или для выявления несущественных параметров в задачах восстановления закономерности функционирования объекта. Геометрическое истолкование свойств ошибок коэффициентов связано со свойствами эллипсоида их рассеяния, определяемого математическим ожиданием и дисперсией значений ошибок. Пространственное расположение, форма, и размер эллипсоида полностью зависят от плана эксперимента.

Критерию *D*-оптимальности соответствует минимальный объем эллипсоида рассеяния ошибок (минимум произведения всех дисперсий коэффициентов полинома). В соответствующем плане эффекты факторов максимально независимы друг от друга. Этот план минимизируют ожидаемую ошибку предсказания функции отклика. Критерию *A*-оптимальности соответствует план с минимальной суммарной дисперсией всех коэффициентов. Критерию *E*-оптимальности – план, в котором максимальная дисперсия коэффициентов будет минимальна.

Выбор критерия зависит от задачи исследования, так при изучении влияния отдельных факторов на поведение объекта применяют критерий *E* - оптимальности, а при поиске оптимума функции отклика – *D*-оптимальности. Если построение *D*-оптимального плана вызывает затруднения, то можно перейти к *A*-оптимальному плану, построение которого осуществляется проще.

Критерии *второй группы* используются при решении задач описания поверхности отклика, определения ограничений на значения параметров. Основным здесь является критерий *G*-оптимальности, который позволяет построить план с минимальным значением наибольшей ошибки в описании функции отклика. Применение *G*-оптимального плана дает уверенность в том, что в области планирования нет точек с чрезмерно большой ошибкой описания функции.

Среди всех классов планов основное внимание в практической работе уделяется ортогональным и ротатабельным планам.

Ортогональным называется план, для которого выполняется условие парной ортогональности столбцов матрицы планирования, в частности, для независимых переменных, где N – количество точек плана эксперимента, k – количество независимых факторов. При ортогональном планировании коэффициенты полинома определяются независимо друг от друга – вычеркивание или добавление слагаемых в функции отклика не изменяет значения остальных коэффициентов полинома. Для ортогональных планов эллипсоид рассеяния ориентирован в пространстве так, что направления его осей совпадают с направлениями координат пространства параметров.

Использование *ротатабельных* планов обеспечивает для любого направления от центра эксперимента равнозначность точности оценки функ-

ции отклика (постоянство дисперсии предсказания) на равных расстояниях от центра эксперимента. Это особенно важно при решении задач поиска оптимальных значений параметров на основе градиентного метода, так как исследователь до начала экспериментов не знает направление градиента и поэтому стремится принять план, точность которого одинакова во всех направлениях. В ряде случаев при исследовании поверхности отклика требуется униморфность модели, а именно, соблюдение постоянства значений дисперсии ошибки в некоторой области вокруг центра эксперимента. Выполнение такого требования целесообразно в тех случаях, когда исследователь не знает точно расположение области поверхности отклика с оптимальными значениями параметров. Указанная область будет определена на основе упрощенной модели, полученной по результатам экспериментов.

По соотношению между количеством оцениваемых неизвестных параметров модели и количеством точек плана эксперимента все планы подразделяются на три класса: *ненасыщенные* – количество параметров меньше числа точек плана; *насыщенные* – обе величины одинаковы; *сверхнасыщенные* – количество параметров больше числа точек плана. Метод наименьших квадратов применяют только при ненасыщенном и насыщенном планировании, и он не применим для сверхнасыщенного планирования.

Для некоторых планов важную роль играет свойство *композиционности*. Так, композиционные планы для построения полиномов второго порядка получают добавлением некоторых точек к планам формирования линейных функций. Это дает возможность в задачах исследования сначала попытаться построить линейную модель, а затем при необходимости, добавив наблюдения, перейти к моделям второго порядка, используя ранее полученные результаты и сохраняя при этом некоторое заданное свойство плана, например его ортогональность. Между критериями оптимальности и методами построения оптимальных планов экспериментов существует жесткая связь. Построение планов производится или с использованием каталогов планов или с использованием непосредственно методов планирования экспериментов, что является непростой задачей и требует достаточно высокой квалификации исследователя в области теории планирования эксперимента.

ТЕСТЫ

Приведенные тесты предназначены как для проверки степени усвоения материала на аудиторных занятиях, так и для самостоятельной работы студентов.

Тест 1

1. Как собирается информация при пассивном наблюдении?

А) получается путем регистрации необходимых сведений в условиях обычного функционирования объекта;

Б) путем целенаправленного воздействия на объект по заранее составленной схеме

В) путем воздействия на объект непредусмотренных факторов.

2. Как собирается информация при активном эксперименте?

А) путем воздействия на объект непредусмотренных факторов;

Б) производится целенаправленное воздействие на объект по заранее составленной схеме;

В) получается путем регистрации необходимых сведений в условиях обычного функционирования объекта.

3. На что направлены методы теории планирования экспериментов?

А) на разработку оптимальных планов проведения экспериментов с целью сокращения объема проводимых исследований при заданной точности и достоверности получения результатов, извлечения из полученных опытных данных максимума полезных сведений;

Б) на расчет входных параметров эксперимента;

В) на составление схемы планирования эксперимента.

4. Автором монография «Design of Experiments», давшей название всему направлению теории планирования эксперимента, является:

А) Р. Боуз; Б) Ф. Йетс; В) Р. Фишер.

5. Понятие дробные реплики от факторного эксперимента были введены?

А) Д. Финни в 1945 г.;

Б) Г. Хотеллинг в 1941 г.;

В) Дж. Боксом в 1951 г.

6. Формализация объекта исследования появилась:

А) в 1951 г. в работах американских ученых Дж. Бокса и К. Уилсона;

Б) в 1947 г. после создания Н. Винером теории кибернетики;

В) в работах Н. П. Клепикова, С. Н. Соколова и В. В. Федорова по ядерной физике.

7. Этап «эволюционного планирования» появился:

А) в 1957 г.; Б) в 1947 г.; В) в 1951 г.

8. Какие из следующих типовых задач экспериментального исследования решает теория планирования эксперимента: 1) поиск значений параметров системы, обеспечивающих достижение оптимального значения показателя качества исследуемого объекта при известных ограничениях на значения этих параметров; 2) приближенное аналитическое описание функциональной связи показателей качества с параметрами системы по результатам проведенного эксперимента; 3) оценка дифференциального влияния уровней параметров системы на показатель качества?

А) только 1; Б) 1, 2, 3; В) 2 и 3.

9. Что из вышеперечисленного должен знать исследователь на этапе планирования эксперимента: 1) к какому классу относится моделируемая система? 2) какой режим работы интересует (стационарный (установившийся) или нестационарный)? 3) в течение какого промежутка времени следует наблюдать за поведением (функционированием) системы? 4) какой объем испытаний

(то есть повторных экспериментов) сможет обеспечить требуемую точность оценок (в статистическом смысле) исследуемых характеристик системы?

А) 1, 2, 3 и 4; Б) 2 и 3; В) 3 и 4.

10. Факторное пространство это:

А) все внешние параметры модели;

Б) множество внешних и внутренних параметров модели, значения которых исследователь может контролировать в ходе подготовки и проведения эксперимента;

В) Все внутренние параметры модели.

11. Уровнями в теории планирования эксперимента называют:

А) значения факторов;

Б) значения выходных параметров;

В) интервалы варьирования.

12. При проведении активного эксперимента исследователь может изменять уровни факторов, верно ли это?

А) верно;

Б) неверно.

13. Центром плана называется:

А) Точка в факторном пространстве (ФП), соответствующая минимальным уровням всех факторов;

Б) Точка в ФП, соответствующая нулевым уровням всех факторов;

В) Точка в ФП, соответствующая максимальным уровням всех факторов.

14. Из каких составляющих складывается значение наблюдаемой переменной, полученное в ходе эксперимента:

А) из функции отклика и ее дисперсии;

Б) из ошибки эксперимента и дисперсии;

В) из функции отклика (неслучайная функция факторов) и ошибки эксперимента (случайная величина).

15. Эксперимент называется идеальным при дисперсии воспроизводимости эксперимента D_y :

А) равной 0; Б) равной 0,5; В) равной 0,01.

16. Стратегическое планирование это:

А) решение задачи планирования во второй постановке;

Б) решение задачи планирования в первой постановке;

В) полное построение плана эксперимента.

17. К какому типу эксперимента относится промышленный эксперимент?

А) к пассивному; Б) к активному; В) к пассивному и активному вместе.

18. Что означает требование некоррелированности факторов?

А) между значениями факторов линейная зависимость;

Б) между значениями факторов нет никакой связи;

В) чтобы связь между факторами не была линейной.

19. Факторы в эксперименте бывают:

А) только качественными;

Б) качественными и количественными;

В) только количественными.

20. Для того, чтобы перейти к количественным значениям, качественные факторы:

А) квантифицируют;

Б) унифицируют;

В) проводят рандомизацию.

21. Центральной точкой плана называется:

А) точка в пространстве параметров, представляющую совокупность основных уровней всех откликов;

Б) точка в пространстве параметров, представляющую совокупность основных уровней всех факторов;

В) точка в пространстве параметров, представляющую совокупность основных уровней всех откликов и всех факторов.

22. Строки матрицы плана эксперимента соответствуют:

А) опытам; Б) факторам; В) откликам.

23. Функция отклика можно рассматривать как :

А) вектор заданных базисных функций;

Б) вектор заданных базисных функций;

В) показатель качества или эффективности объекта.

24. В каком виде целесообразно представить функцию отклика:

А) в виде логарифма; Б) в виде дроби; В) в виде полинома.

25. Эффектами факторов называют:

А) коэффициенты полинома;

Б) функцию отклика;

В) вектор заданных базисных функций.

Тест 2

1. Принципиальные ограничения для границы областей определения факторов принимаются для:

А) значений факторов, связанных с технико-экономическими соображениями;

Б) значений факторов, которые не могут быть нарушены ни при каких обстоятельствах;

В) для значений факторов, определяемых конкретными условиями проведения процесса.

2. Априорной называется информация, полученная:

А) до начала эксперимента, содержащаяся в результатах предыдущих исследований;

Б) в ходе проведения эксперимента;

В) после проведения эксперимента.

3. Интервал варьирования – это:

А) координаты основного уровня;

Б) координаты нижнего уровня;

В) расстояние на координатной оси между основным и верхним (или нижним) уровнем.

4. Каким ограничениям должен отвечать интервал варьирования:

А) его применение должно приводить к выходу фактора за пределы области допустимых значений;

Б) его применение не должно приводить к выходу фактора за пределы области допустимых значений, и он должен быть больше погрешности задания значений фактора (в противном случае уровни фактора станут не различимыми);

В) он должен быть не больше погрешности задания значений фактора.

5. Сколько значений принимает каждый фактор при вычислении оценок коэффициентов полиномов второго порядка или линейных полиномов:

А) четыре

Б) три

В) два

6. Количество точек в k -мерном пространстве в полном факторном плане $N =$:

А) $2k$; Б) $2k-1$; В) $4k$.

7. Для чего в табличной записи матрицы планирования эксперимента вводится фиктивная переменная $x_0 = 1$?

А) для единообразия записи расчетных формул коэффициентов полинома;

Б) для ограничения величин факторов;

В) для характеристики вектора результатов.

8. В матрице планирования эксперимента столбцы с произведениями переменных вычисляются:

А) путем сложения значений элементов в соответствующих столбцах простых переменных;

Б) путем умножения значений элементов в соответствующих столбцах простых переменных;

В) путем деления значений элементов в соответствующих столбцах простых переменных.

9. На что указывают коэффициенты при независимых переменных при полном факторном эксперименте?

А) на силу влияния факторов;

Б) на погрешность исследования;

В) на ошибку в расчете отклика.

10. С ростом числа факторов число возможных взаимодействий:

А) быстро растет; Б) не меняется; В) уменьшается.

11. Генерирующее соотношение – это:

А) соотношение, показывающее, с каким из эффектов смешан данный эффект ;

Б) соотношение коэффициентов при независимых переменных;

В) соотношение числа факторов.

12. С ростом числа факторов дробность реплик:

А) увеличивается; Б) уменьшается; В) не меняется.

13. С ростом числа факторов система смешивания:

А) не меняется; Б) усложняется; В) упрощается.

14. Полуреплика – это:

А) реплика, включающая только половину экспериментов ПФЭ;

Б) реплика, включающая только четверть экспериментов ПФЭ;

В) реплика, включающая все эксперименты ПФЭ.

15. Для построения дробных реплик следует выбирать:

А) незначимые взаимодействия факторов, которые выбираются по физическим соображениям на основе априорных сведений

Б) значимые взаимодействия факторов, которые выбираются после проведения эксперимента

В) значимые взаимодействия факторов, которые выбираются по физическим соображениям на основе априорных сведений

16. ДФЭ позволяет получить несмещенную оценку градиента функции отклика, если:

А) ее обобщающий определяющий контраст не больше трех;

Б) ее обобщающий определяющий контраст больше трех;

В) ее обобщающий определяющий контраст не больше двух.

17. Дисперсией называется:

А) минимальное значение квадрата отклонений величины от ее минимального значения;

Б) максимальное значение квадрата отклонений величины от ее максимального значения;

В) среднее значение квадрата отклонений величины от ее среднего значения.

18. Средним квадратическим отклонением называется:

А) корень квадратный из дисперсии, взятый с положительным знаком;

Б) квадрат дисперсии, взятый с положительным знаком;

В) среднее значение квадрата отклонений величины от ее среднего значения.

19. Систематические ошибки порождаются причинами:

А) действующими регулярно, в определенном направлении;

Б) действующими нерегулярно, в определенном направлении;

В) действующими регулярно, в случайном направлении;

20. Можно ли усреднять неоднородные дисперсии?

А) можно;

Б) можно в исключительных случаях;

В) нельзя.

21. Критерий Фишера (F-критерий) представляет собою:

А) отношение большей дисперсии к большей;

Б) отношение меньшей дисперсии к меньшей;

В) отношение средней дисперсии к меньшей.

22. Гипотеза об однородности оценок дисперсии воспроизводимости в различных точках плана принимается, если выполняется условие:

А) $F \geq F_{кр}$; Б) $F \leq F_{кр}$; В) $F > F_{кр}$.

23. Существенным недостатком критерия Фишера является:

А) учет всех оценок дисперсии воспроизводимости;

Б) учет только максимальных оценок дисперсии воспроизводимости;

В) игнорирование всех оценок дисперсии воспроизводимости, кроме максимального и минимального значения.

24. Критерий Кохрена :

А) это отношение максимальной дисперсии к сумме всех дисперсий;

Б) это отношение минимальной дисперсии к сумме всех дисперсий;

В) это отношение максимальной дисперсии к минимальной.

25. Числом степеней свободы в статистике называется:

А) сумма числа опытов и числа коэффициентов (констант);

Б) разность между числом опытов и числом коэффициентов (констант), которые уже вычислены по результатам этих опытов независимо друг от друга;

В) разность между числом опытов и дисперсиями.

Тест 3

1. При построении плана полного факторного эксперимента (ПФЭ) типа $3k$ для трех переменных количество точек плана составит:

А) 27; Б) 9; В) 3.

2. На каком этапе исследования используются композиционные планы?

А) на начальном этапе;

Б) на заключительном этапе;

В) в середине исследования.

3. Центральный композиционный план второго порядка называют планом Бокса:

А) если его ядром является ПФЭ $3k$, а также два любых парных взаимодействия по модулю равны друг другу;

Б) если его ядром является ПФЭ $2k$ или регулярная реплика типа $2k - p$, для которой парные взаимодействия не равны по модулю линейным факторам, а также два любых парных взаимодействия по модулю не равны друг другу;

В) если его ядром является ПФЭ $2k$, а также два любых парных взаимодействия по модулю равны друг другу.

4. Центральный композиционный план второго порядка называют планом Хартли, если:

А) его ядром является ПФЭ $2k$, а также два любых парных взаимодействия по модулю равны друг другу;

Б) если его ядром является ПФЭ $3k$, а также два любых парных взаимодействия по модулю равны друг другу;

В) его ядром является регулярная реплика типа $2k - p$, в которой некоторые парные взаимодействия равны по модулю линейным факторам.

5. Планы Хартли отличаются от планов Бокса тем, что:

А) обычно более экономны по числу опытов, но уступают по точности оценивания коэффициентов;

Б) являются ортогональными и ротатабельными;

В) обеспечивают более высокую точность оценивания коэффициентов.

6. В центральном композиционном плане (ЦКП) Бокса каждый фактор варьируется на:

А) трех уровнях; Б) пяти уровнях; В) шести уровнях.

7. Точки ротатабельного ЦКП Бокса второго порядка располагают на:

А) концентрических гиперсферах, количество которых не более двух;

Б) одной концентрической гиперсфере;

В) концентрических гиперсферах, количество которых не менее двух.

8. Результаты опытов в нулевой точке служат для:

А) проверки гипотезы об адекватности модели экспериментальным данным;

Б) изменения числа уровней варьирования факторов;

В) обеспечения более высокой точности оценивания коэффициентов.

9. Если многомерное пространство допустимых значений факторов представляет собой куб, то допустимые области значений факторов должны удовлетворять условиям:

А) $-1 \geq x_i \leq 1, i = 1, 2, \dots, k;$

Б) $x_{12} + x_{22} + \dots + x_{k2} \leq 1;$

В) $x_i \geq 1, i = 1, 2, \dots, k.$

10. Если многомерное пространство допустимых значений факторов представляет собой шар, то допустимые области значений факторов должны удовлетворять условиям:

А) $x_i \geq 1, i = 1, 2, \dots, k;$

Б) $x_{12} + x_{22} + \dots + x_{k2} \leq 1;$

В) $x_{12} + x_{22} + \dots + x_{k2} \geq 1.$

11. Латинский квадрат позволяет исследовать влияние:

А) не более чем трех факторов;

Б) более трех факторов;

В) неограниченного числа факторов.

12. Греко-латинские квадраты образуются путем:

А) суммирования столбцов;

Б) суммирования строк;

В) наложения друг на друга латинских квадратов.

13. План гипер-греко-латинского квадрата получается:

А) при наложении друг на друга трех различных вариантов латинских квадратов;

Б) при наложении друг на друга пяти различных вариантов латинских квадратов;

В) при наложении друг на друга десяти различных вариантов латинских квадратов.

14. Первый постулат регрессионного анализа:

А) проводится учет всех оценок дисперсии воспроизводимости;

Б) проводится учет только максимальных оценок дисперсии воспроизводимости;

В) параметр оптимизации y есть случайная величина с нормальным законом распределения.

15. Второй постулат регрессионного анализа:

А) Дисперсия y не зависит от абсолютной величины y .

Б) Дисперсия y зависит от абсолютной величины y .

В) Дисперсия y частично зависит от абсолютной величины y .

16. Из третьего постулата регрессионного анализа следует, что:

А) значения факторов – случайные величины;

Б) ошибка воспроизводимости существенно точнее, чем установление каждого фактора на заданный уровень и его поддержание на нем;

В) установление каждого фактора на заданный уровень и его поддержание существенно точнее, чем ошибка воспроизводимости.

17. Если величина коэффициента регрессии положительна, то:

А) с увеличением значения фактора величина параметра оптимизации убывает;

Б) с увеличением значения фактора растет величина параметра оптимизации;

В) с увеличением значения фактора величина параметра оптимизации не меняется.

18. При поиске максимума функции отклика:

А) увеличение значений всех факторов, коэффициенты которых имеют знак плюс, благоприятно, а имеющих знак минус – неблагоприятно;

Б) уменьшение значений всех факторов, коэффициенты которых имеют знак плюс, благоприятно, а имеющих знак минус – благоприятно;

В) уменьшение значений всех факторов, коэффициенты которых имеют знак плюс или минус, одинаково благоприятно.

19. Если в некоторой задаче взаимодействие двух факторов значимо и имеет положительный знак, то:

А) это свидетельствует о том, что одновременное увеличение, как и одновременное уменьшение, значений двух факторов приводит к уменьшению параметра оптимизации (без учета линейных эффектов);

Б) это свидетельствует о том, что одновременное увеличение, как и одновременное уменьшение, значений двух факторов приводит к увеличению параметра оптимизации (без учета линейных эффектов);

В) это свидетельствует о том, что одновременное увеличение, как и одновременное уменьшение, значений двух факторов не приводит к изменению параметра оптимизации (без учета линейных эффектов).

20. Если в линейной модели часть коэффициентов регрессии значима, часть незначима, то возможно принятие следующего решения:

А) и движение по градиенту;

Б) переход к планам второго порядка;

В) увеличение числа параллельных опытов и достройка плана.

21. Метод «перевала» при достройке плана заключается в том, что:

А) у исходной реплики изменяют знаки на обратные. В этом случае основные эффекты оказываются не смешанными с парными эффектами;

Б) осуществляется переход к полному факторному эксперименту;

В) осуществляется переход к реплике меньшей подробности.

22. В случае, если линейная модель адекватна, а все коэффициенты регрессии незначимы (кроме b_0), необходимо:

А) перейти к плану второго порядка;

Б) увеличение точности эксперимента и расширение интервалов варьирования;

В) перейти к полному факторному эксперименту.

23. Если линейная модель неадекватна, то для получения адекватной модели принимают следующие решения:

А) изменение интервалов варьирования факторов, перенос центра плана, достройка плана.

Б) осуществление переходов к реплике меньшей дробности;

В) осуществление переходов к плану второго порядка.

24. Симметричной относительно коэффициентов называется:

А) функция у которой величины коэффициентов которой существенно различаются;

Б) функция, величины коэффициентов которой различаются не существенно;

В) функция, у которой величины коэффициентов которой равны.

Тест 4

1. Метод Монте-Карло часто называют методом статистических испытаний, так как он:

А) требует проведения большого числа испытаний;

Б) не требует проведения большого числа испытаний;

В) требует проведения большого количества расчетов.

2. Что такое $er f$?

А) число испытаний;

Б) функция ошибок;

В) среднее квадратическое отклонение.

3. Применение метода Монте-Карло становится оправданным при кратности интеграла:

А) меньше 3; Б) равной 3; В) больше трех.

4. Несмещенной оценкой математического ожидания случайной величины является:

А) ее среднее арифметическое;

Б) ее максимальное значение;

В) ее минимальное значение.

5. Точность расчета математического ожидания и определенного интеграла будет расти:

А) с уменьшением числа испытаний;

Б) с ростом среднего квадратического отклонения;

В) с ростом числа испытаний.

6. Получение случайных чисел с требуемым законом распределения обычно выполняется:

А) в два этапа; Б) в три этапа; В) в четыре этапа.

7. Генераторы случайных чисел могут быть:

А) физические и программные;

Б) программные;

В) физические.

8. Корреляционная функция называется автокорреляционной, если:

А) производится статистический анализ нескольких случайных процессов;
Б) производится статистический анализ одного случайного процесса (или одной выборки случайных чисел);

В) производится статистический анализ нескольких выборок случайных чисел.

9. Нормированной корреляционной функцией называется:

А) отношение центрированной корреляционной функции к дисперсии случайного процесса;

Б) отношение дисперсии случайного процесса к центрированной корреляционной функции;

В) отношение центрированной корреляционной функции к функции ошибок.

10. Для нахождения значений нормированной корреляционной функции необходимо:

А) перемножить значения случайного процесса или значения отсчетов случайной величины;

Б) приравнять значения случайного процесса или значения отсчетов случайной величины;

В) предварительно центрировать значения случайного процесса или значения отсчетов случайной величины.

11. Метод интервальных оценок применяется в случае:

А) значительного числа наблюдений, по которым необходимо произвести оценку параметра;

Б) небольшого числа наблюдений, по которым необходимо произвести оценку параметра;

В) небольших значений центрированной корреляционной функции.

12. Отрицательное биномиальное распределение относится к:

А) дискретным распределениям;

Б) случайным распределениям;

В) степенным распределениям.

13. Поиск экстремума функции отклика происходит:

А) посредством измерения факторов;

Б) посредством измерения поверхности отклика в различных точках факторного пространства;

В) посредством измерения коэффициентов.

14. Вычисление оценки градиента точке по результатам эксперимента – это одна из составляющих метода:

А) Бокса и Уильсона; Б) Фишера; В) Паскаля.

15. После определения максимума оценки функции отклика производится:

А) переход к кодированной системе координат;

Б) пересчет факторов в логарифмическую функцию;

В) переход к системе координат натуральных факторов (переход от кодированных к натуральным).

16. Интеграл вероятностей, это:

А) функция Лапласа;

Б) функция ошибок;

В) среднее квадратическое отклонение.

17. Вычисление объема тел с помощью трехкратного интеграла сводится к:

А) взятию интеграла по области при подынтегральной функции, не равной единице;

Б) взятию интеграла по области при подынтегральной функции, тождественно равной единице;

В) взятию интеграла по области при подынтегральной функции.

18. Генератор случайных чисел, применявшийся в начале 50-х годов XX века, использовал метод:

А) Фибоначчи; Б) Паскаля; В) Лагранжа.

19. Для определения корреляционной функции по результатам опыта выбирается:

А) маленький объем выборки;

Б) средний объем выборки;

В) достаточно большой объем выборки, чтобы можно было в широком диапазоне формировать разницу между двумя соседними значениями случайных чисел.

20. Величина τ обозначает:

А) разницу для непрерывного времени;

Б) корреляционную функцию;

В) объем выборки.

21. $R(\tau)$ – это:

А) область суммирования;

Б) корреляционная функция;

В) объем выборки.

22. Диапазон вычисления корреляционной функции определяется как:

А) $N + \tau$; Б) $N : \tau$; В) $N - \tau$.

23. Область суммирования принимает значения:

А) от 1 до $N + \tau$;

Б) от 1 (первое случайное число выборки) до $N - \tau$;

В) от 1 до $N : \tau$.

24. Нормированная корреляционная функция отличается от централизованной в s раз, где s :

А) область суммирования;

Б) дисперсия данной выборки случайных чисел;

В) объем выборки.

25. ГСЧ считается хорошим, если при τ , не равным нулю, модуль нормированной корреляционной функции:

А) меньше 0,1; Б) больше 0,1; В) равен 1.

ГЛОССАРИЙ

Активный эксперимент – эксперимент, в котором уровни факторов в каждом опыте задаются исследователем.

Априорное ранжирование факторов – метод выбора наиболее важных факторов, основанный на экспертной оценке.

Временный дрейф – случайное или неслучайное изменение функции отклика во времени.

Дисперсия оценки функции отклика – дисперсия оценки математического ожидания отклика в некоторой данной точке факторного пространства.

Интервал варьирования фактора – половина размаха варьирования фактора.

Нормализация факторов – преобразование натуральных значений факторов в безразмерные значения.

Область оптимума – область факторного пространства в окрестности точки, в которой функция отклика достигает экстремального значения.

Область экспериментирования (область планирования) – область факторного пространства, где могут размещаться точки, отвечающие условиям проведения опытов.

Основной уровень фактора – натуральное значение фактора, соответствующее нулю в безразмерной шкале.

Опыт – воспроизведение исследуемого явления в определенных условиях проведения эксперимента при возможности регистрации его результатов.

Отклик – наблюдаемая случайная переменная, по предположению, зависящая от факторов.

Оценка функции отклика – зависимость, получаемая при подстановке в функцию отклика оценок значений ее параметров.

Параллельные опыты – рандомизированные во времени опыты, в которых уровни всех факторов сохраняются неизменными.

Пассивный эксперимент – эксперимент, при котором уровни факторов в каждом опыте регистрируются исследователем, но не задаются.

План эксперимента – совокупность данных, определяющих число, условия и порядок реализации опытов.

Планирование эксперимента – выбор плана эксперимента, удовлетворяющего заданным требованиям.

Поверхность отклика (*Поверхность регрессии*) – геометрическое представление функции отклика.

Поверхность уровня функции отклика – геометрическое место точек в факторном пространстве, которому соответствует некоторое фиксированное значение функции отклика.

Последовательный эксперимент (шаговый эксперимент) – эксперимент, реализуемый в виде серий, в котором условия проведения каждой последующей серии определяются результатами предыдущих.

Размах варьирования фактора – разность между максимальным и минимальным натуральными значениями фактора в данном плане.

Рандомизация плана – один из приемов планирования эксперимента, имеющий целью свести эффект некоторого неслучайного фактора к случайной ошибке.

Уровень фактора – фиксированное значение фактора относительно начала отсчета.

Фактор – переменная величина, по предположению влияющая на результаты эксперимента.

Факторное пространство – пространство, координатные оси которого соответствуют значениям факторов.

Функция отклика – зависимость математического ожидания отклика от факторов.

Эксперимент – система операций, воздействий и (или) наблюдений, направленных на получение информации об объекте при исследовательских испытаниях.

Эффект взаимодействия факторов – показатель зависимости изменения эффекта одного фактора от уровней других факторов.

ПРИЛОЖЕНИЕ 1

1. Нормальное распределение. Плотность вероятностей нормированного нормального распределения:

$$u \rightarrow N(0,1) \quad f(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}.$$

U	0	1	2	3	4	5	6	7	8	9
1	2	3	4	5	6	7	8	9	10	11
0,0	0,3989	3989	3989	3988	3986	3984	3982	3980	3977	3973
0,1	3970	3965	3961	3956	3951	3945	3939	3932	3925	3918
0,2	3910	3902	3894	3885	3876	3867	3857	3847	3836	3825
0,3	3814	3802	3790	3778	3765	3752	3739	3726	3712	3697
0,4	3683	3668	3653	3637	3621	3605	3589	3572	3555	3538
0,5	3521	3503	3485	3467	3448	3429	3410	3391	3372	3352
0,6	3332	3312	3292	3271	3251	3230	3209	3187	3166	3144
0,7	3123	3101	3079	3056	3034	3011	2989	2966	2943	2920
0,8	2897	2874	2850	2827	2803	2780	2756	2732	2709	2685
0,9	2661	2637	2613	2589	2565	2541	2516	2492	2468	2444
1,0	0,2420	2396	2371	2347	2323	2299	2275	2251	2227	2203
1,1	2179	2155	2131	2107	2083	2059	2036	2012	1989	1965
1,2	1942	1919	1895	1872	1849	1826	1804	1781	1758	1736
1,3	1714	1691	1669	1647	1626	1604	1582	1561	1539	1518
1,4	1497	1476	1456	1435	1415	1394	1374	1354	1334	1315
1,5	1295	1276	1257	1238	1219	1200	1182	1163	1145	1127
1,6	1109	1092	1074	1057	1040	1023	1006	0989	0973	0957
1,7	0940	0925	0909	0893	0878	0863	0848	0833	0818	0804
1,8	0790	0775	0761	0748	0734	0721	0707	0694	0681	0669
1,9	0656	0644	0632	0620	0608	0596	0584	0573	0562	0551
2,0	0,0540	0529	0519	0508	0498	0488	0478	0468	0459	0449
2,1	0440	0431	0422	0413	0404	0396	0387	0379	0371	0363
2,2	0355	0347	0339	0332	0325	0317	0310	0303	0297	0290
2,3	0283	0277	0270	0264	0258	0252	0246	0241	0235	0229
2,4	0224	0219	0213	0203	0203	0198	0194	0189	0184	0180
2,5	0175	0181	0167	0158	0158	0154	0151	0147	0143	0139
2,6	0136	0132	0129	0122	0122	0119	0116	0113	0110	0107
2,7	0104	0101	0099	0096	0093	0091	0088	0086	0084	0081
2,8	0079	0077	0075	0071	0071	0069	0067	0065	0063	0061
2,9	0060	0058	0056	0053	0053	0051	0050	0048	0047	0046
3,0	0,0044	0043	0042	0040	0039	0038	0037	0036	0035	0034
3,1	0033	0032	0031	0030	0029	0028	0027	0026	0025	0025
3,2	0024	0023	0022	0022	0021	0020	0020	0019	0018	0018
3,3	0017	0017	0016	0016	0015	0015	0014	0014	0013	0013

3,4	0012	0012	0012	0011	0011	0010	0010	0010	0009	0009
3,5	0009	0008	0008	0008	0008	0007	0007	0007	0007	0006
3,6	0006	0006	0006	0005	0005	0005	0005	0005	0005	0004
3,7	0004	0004	0004	0004	0004	0004	0003	0003	0003	0003
3,8	0003	0003	0003	0003	0003	0002	0002	0002	0002	0002
3,9	0002	0002	0002	0002	0002	0002	0002	0002	0001	0001

2. Нормальное распределение. Значение функции:

$$\Phi(u_i) = \frac{2}{\sqrt{2\pi}} \int_0^{u_i} e^{-\frac{x^2}{2}} dx = P(|u_i| < u_i).$$

Целые и десятичные доли u_i	Сотые доли u_i									
	0	1	2	3	4	5	6	7	8	9
0,0	0,0000	0,0080	0,0160	0,0239	0,0319	0,3999	0,0478	0,0558	0,0638	0,0717
0,1	0797	0876	0955	1034	1113	1192	1271	1350	1428	1507
0,2	1585	1663	1741	1819	1897	1974	2051	2128	2205	2282
0,3	2358	2434	2510	2586	2661	2737	2812	2886	2960	3035
0,4	3108	3182	3255	3328	3401	3473	3545	3616	3688	3759
0,5	3829	3899	3969	4039	4108	4177	4245	4313	4381	4448
0,6	4515	4581	4647	4713	4778	4843	4907	4971	5035	5098
0,7	5161	5223	5285	5346	5407	5467	5527	5587	5646	5705
0,8	5763	5821	5878	5935	5991	6047	6102	6157	6211	6265
0,9	6319	6372	6424	6476	6528	6579	6629	6679	6729	6778
1,0	0,6827	0,6875	0,6923	0,6970	0,7017	0,7063	0,7109	0,7154	0,7199	0,7243
1,1	7287	7330	7373	7415	7457	7499	7540	7580	7620	7660
1,2	7699	7737	7775	7813	7850	7887	7923	7959	7994	8029
1,3	8064	8098	8132	8165	8198	8230	8262	8293	8324	8355
1,4	8385	8415	8444	8473	8501	8529	8557	8584	8611	8638
1,5	8664	8690	8715	8740	8764	8789	8812	8836	8859	8882
1,6	8904	8926	8948	8969	8990	9011	9031	9051	9070	9090
1,7	9109	9127	9146	9164	9181	9199	9216	9233	9249	9265
1,8	9281	9297	9312	9327	9342	9357	9371	9385	9399	9412
1,9	9426	9439	9451	9464	9476	9488	9500	9512	9523	9534
2,0	0,9545	0,9556	0,9566	0,9576	0,9586	0,9596	0,9606	0,9616	0,9625	0,9634
2,1	9643	9651	9660	9668	9676	9684	9692	9700	9707	9715
2,2	9722	9729	9736	9743	9749	9756	9762	9768	9774	9780
2,3	9786	9791	9797	9802	9807	9812	9817	9822	9827	9832
2,4	9836	9841	9845	9849	9853	9857	9861	9865	9869	9872
2,5	9876	9879	9883	9886	9889	9892	9895	9898	9901	9904
2,6	9907	9910	9912	9915	9917	9920	9922	9924	9926	9928
2,7	9931	9933	9935	9937	9939	9940	9942	9944	9946	9947
2,8	9949	9951	9952	9953	9955	9956	9958	9959	9960	9961
2,9	9963	9964	9965	9966	9967	9968	9969	9970	9971	9972
3,0	0,9973	0,9974	0,9975	0,9976	0,9976	0,9977	0,9979	0,9979	0,9979	0,9980
3,1	9981	9981	9984	9983	9983	9984	9984	9985	9985	9986
3,2	9986	9987	9987	9988	9988	9989	9989	9989	9990	9990
3,3	9990	9991	9991	9991	9992	9992	9992	9992	9993	9993
3,4	9993	9994	9994	9994	9994	9994	9995	9995	9995	9995
3,5	9995	9996	9996	9996	9996	9996	9996	9996	9997	9997

3,6	9997	9997	9997	9997	9997	9997	9997	9998	9998	9998	9998
3,7	9998	9998	9998	9998	9998	9998	9998	9998	9998	9998	9998
3,8	9999	9999	9999	9999	9999	9999	9999	9999	9999	9999	9999
3,9	9999	9999	9999	9999	9999	9999	9999	9999	9999	9999	9999

3. χ^2 - распределение

В таблице представлены значения квантилей $\chi^2_{\alpha, v}$ в зависимости от числа степеней свободы v и вероятности α .

$\alpha \backslash v$	0,20	0,10	0,05	0,02	0,01	0,001
1	1,642	2,706	3,841	5,412	6,635	10,827
2	3,219	4,605	5,991	7,824	9,210	13,815
3	4,642	6,251	7,815	9,837	11,345	16,266
4	5,989	7,779	9,488	11,668	13,277	18,467
5	7,289	9,236	11,070	13,388	15,086	20,515
6	8,558	10,645	12,592	15,033	16,812	22,457
7	9,803	12,017	14,067	16,622	18,475	24,322
8	11,030	13,362	15,507	18,168	20,090	26,125
9	12,242	14,684	16,919	19,679	21,666	27,877
10	13,442	15,987	18,307	21,161	23,209	29,588
11	14,631	17,275	19,675	22,618	24,725	31,264
12	15,812	18,549	21,026	24,054	26,217	32,909
13	16,985	19,812	22,362	25,472	27,688	34,528
14	18,151	21,064	23,685	26,783	29,141	36,123
15	19,311	22,307	24,996	28,259	30,578	37,697
16	20,465	23,542	26,296	29,633	32,000	39,252
17	21,615	24,769	27,587	30,995	33,409	40,790
18	22,760	25,989	28,869	32,346	34,805	42,312
19	23,900	27,204	30,144	33,687	36,191	43,820
$\alpha \backslash v$	0,20	0,10	0,05	0,02	0,01	0,001
20	25,038	28,412	31,410	35,020	37,556	45,315
21	26,171	29,615	32,671	36,343	38,932	46,797
22	27,301	30,813	33,924	37,659	40,289	48,268
23	28,429	32,007	35,172	38,968	41,638	49,728
24	29,559	33,196	36,415	40,270	42,980	51,179
25	30,675	34,382	37,652	41,556	44,314	52,620
26	31,795	35,563	38,885	42,856	45,642	54,052
27	32,912	36,741	40,113	44,140	46,963	55,476
28	34,027	37,916	41,337	45,419	48,278	56,893
29	35,139	39,087	42,557	46,693	49,588	58,302
30	36,250	40,256	43,773	47,962	50,892	59,703

ПРИЛОЖЕНИЕ 2

Пример построения математической модели методом полного факторного эксперимента

В качестве примера построения математической модели методом полного факторного эксперимента предлагается пример построения математической модели процесса дефосфотации (удаления соединений фосфора) из сточных вод [9,10].

Эффективность удаления соединений фосфора Ξ [%] зависит от многих факторов, среди которых: исходная концентрация соединений фосфора в воде $C_{\text{исх.}}(\text{P})$ [мг/л], доза реагента $C_{\text{реаг.}}$ [мг/л], остаточная концентрация соединений фосфора $C_{\text{ост.}}(\text{P})$ [мг/л], водородный показатель pH , окислительно-восстановительный потенциал Eh [мВ], температура воды t [$^{\circ}\text{C}$], время отстаивания τ [мин], вид сточных вод, который обозначается буквой B и характеристики, определяющие физические свойства сточной воды: плотность сточной воды $\rho_{\text{ст.в.}}$, коэффициенты динамической и кинематической вязкости жидкости η и ν , плотность взвешенных веществ $\rho_{\text{взв.}}$, а также гидравлическая крупность u_0 образующихся хлопьев. Таким образом, характеристическое уравнение процесса дефосфотации сточных вод в общем виде можно скомпоновать следующим образом:

$$Y(\Xi, C_{\text{исх.}}, C_{\text{реаг.}}, C_{\text{ост.}}, \text{pH}, Eh, t, \tau, B, \rho_{\text{ст.в.}}, \nu, \eta, \rho_{\text{взв.}}, u_0) = 0$$

В данном случае часть факторов представляет собой взаимосвязанные величины, так, например: остаточная концентрация $C_{\text{ост.}}(\text{P})$ в очищенной сточной воде входит в показатель эффективности, поэтому ее можно исключить из уравнения; вид исследуемых сточных вод – высококонцентрированные сточные воды животноводческих комплексов, поэтому исключаем B из исходного уравнения, величины $\rho_{\text{ст.в.}}$, ν , η , $\rho_{\text{взв.}}$ входят в показатель u_0 , поэтому оставляем в уравнении u_0 . Таким образом, после исключения из уравнения вышеназванных факторов, получаем:

$$Y(\Xi, C_{\text{исх.}}, C_{\text{реаг.}}, \text{pH}, Eh, t, \tau, u_0) = 0$$

Ограничимся при проведении экспериментов рассмотрением исключительно физико-химических процессов коагуляции фосфатов, исключив из рассмотрения показатели физико-механических процессов осаждения скоагулированной взвеси, так как эти показатели определяют размеры сооружений механической очистки и не оказывают прямого влияния на формальную кинетику химических взаимодействий фосфора с реагентом. Величина pH высококонцентрированных сточных вод варьируется в пределах от 6,5 до 8,5, а в

исследуемой сточной воде рН лежит в пределах от 7,2 до 7,5. В то же время окислительно-восстановительный потенциал (ОВП) любого раствора неразрывно связан с его рН.

Величина E_h связана с активностями ионных форм элементов с переменной валентностью в растворе следующим уравнением:

$$E_h = E_0 + RT / nF \ln (a_{ox} / a_{red}),$$

где E^0 – постоянная для каждой окислительно-восстановительной системы величина нормального потенциала; $R=8,31$ Дж/мольК; T – температура, К; n – число электронов; a_{ox} , a_{red} – активности окисленных и восстановленных форм элементов, образующих обратимую окислительно-восстановительную систему.

В конечном итоге характеристическое уравнение процесса дефосфотации сточных вод принимает вид:

$$Y(\Theta, C_{исх.}, C_{реак.}, pH, E_h, t) = 0$$

Для дальнейшего использования целесообразно объединить содержание фосфатов в исходной воде (в пересчете на фосфор) и дозу реагента в безразмерный комплекс – соотношение $C_{реак. \text{ Me}} : C_{исх.}(P)$, заменив им в уравнении $C_{исх.}$ и $C_{реак.}$. В итоге уравнение примет вид:

$$Y(\Theta, C_{реак. \text{ Me}} : C_{исх.}(P), pH, E_h, t) = 0$$

$$\Theta = f(C_{реак. \text{ Me}} : C_{исх.}(P), pH, E_h, t).$$

Таким образом, для моделирования процессов дефосфотации необходимо изучить зависимость функции отклика – эффективность дефосфотации от определяющих факторов: отношения $C_{реак. \text{ Me}} : C_{исх.}(P)$, E_h и температуры воды t .

При более подробном рассмотрении определяющие факторы имеют следующий вид:

1. Соотношение $D=C_{коаг \text{ по Me}} : C_{исх.}(P)$, которое будет находиться в пределах от 15 до 1 обозначаем за x_1 , причем величина $x_1 = +1$ соответствует 15, а величина $x_1 = -1$ соответствует 1;

2. Окислительно-восстановительный потенциал E_h (находится в пределах от -100 до 100 мВ) обозначаем за x_2 , причем $x_2 = +1$ соответствует максимальному значению $E_h = 100$ мВ, а $x_2 = -1$ – минимальному $E_h = -100$ мВ;

3. Температуру сточной воды t от 20 °С до 30 °С обозначаем за x_3 , причем $x_3 = +1$ соответствует верхней границе температуры в 30 °С, а $x_3 = -1$ – нижней в 20 °С.

План эксперимента, его пятикратная реализация с учетом рандомизации и первичная обработка результатов представлены в таблицах 1 и 2. Рандомизацию проводим с помощью таблицы равномерно распределенных случайных чисел. Определяем последовательность реализации матрицы планирования в каждой из m серий опытов. Для этого в качестве начала выбирается любое число из табл.П.6[2] и записывается в столбец k_1 для номера строки 1. Остальные места этого столбца заполняют числа от 1 до N , следующие по порядку из табл.П.6[2] за выбранным начальным. Аналогично рандомизируются

испытания в каждой из оставшихся серий экспериментов; порядок реализации записывается в столбцах $k_2, k_3, \dots, k_m$.

Таблица 1

План эксперимента с пятикратной реализацией

N стр.	Циклы					z ₀	z ₁	z ₂	z ₃	z ₄	z ₅	z ₆	z ₇	Результаты, Э { % }				
	k ₁	k ₂	k ₃	k ₄	k ₅	x ₀	x ₁	x ₂	x ₃	x ₁ x ₂	x ₁ x ₃	x ₂ x ₃	x ₁ x ₂ x ₃	Y _{g1}	Y _{g2}	Y _{g3}	Y _{g4}	Y _{g5}
1	4	2	3	6	8	+	-	-	-	+	+	+	-	39,8	42,4	41,5	41,8	40,6
2	3	3	6	2	5	+	+	-	-	-	-	+	+	54,1	55,6	54,8	55,2	54,7
3	8	6	2	4	1	+	-	+	-	-	+	-	+	62,0	61,3	61,9	61,7	62,4
4	6	1	7	1	6	+	+	+	-	+	-	-	-	64,6	64,3	64,1	63,9	64,1
5	5	8	1	3	4	+	-	-	+	+	-	-	+	48,6	48,3	47,8	47,4	48,0
6	2	5	5	7	2	+	+	-	+	-	+	-	-	51,4	51,7	50,90	50,80	51,30
7	1	7	4	8	7	+	-	+	+	-	-	+	-	60,2	59,1	59,3	60,4	60,3
8	7	4	8	5	3	+	+	+	+	+	+	+	+	86,3	85,4	86,8	85,0	86,0

Таблица 2

Первичная обработка результатов эксперимента

Обработка		Адекватность	
$\bar{Y}$	S^2_g	$\bar{Y}_g$	$(\frac{\bar{Y}_g - \bar{Y}}{\bar{Y}})^2$
41,22	1,052	41,32	0,01
54,88	0,317	54,98	0,01
61,86	0,163	61,74	0,0144
64,2	0,07	64,08	0,0144
48,02	0,212	48,12	0,01
51,22	0,137	51,34	0,0144
59,86	0,373	59,74	0,0144
85,9	0,51	85,8	0,01
Σ 467,16	2,834		0,0976

Результаты эксперимента в каждой из серий испытаний записываются в столбцах $Y_1, Y_2, \dots, Y_m$. Усредненное значение выходной величины $\bar{Y}$:

$$\bar{Y}_g = (1/m) \cdot \sum_{l=1}^m Y_{lg}.$$

Оценки дисперсий:

$$S^2_g = (1/m-1) \cdot \sum_{l=1}^m (Y_{lg} - \bar{Y}_g)^2.$$

Так как все дисперсии получены по выборкам одинакового объема m , то число степеней свободы для всех дисперсий одинаково и равно

$$V_1 = m-1.$$

Для проверки гипотезы об однородности оценок дисперсий используется критерий Кохрена, который основан на законе распределения отношения максимальной эмпирической дисперсии к сумме всех дисперсий, т. е

$$G = \max\{S_g^2\} / \sum_{g=1}^n S_g^2 < G_{кр}(5\%; \nu_1=4; \nu_2=8)$$

$$G = 1,052/2,834 = 0,371 < G_{кр}(5\%; \nu_1=4; \nu_2=8) = 0,3910.$$

Таким образом, подтверждена воспроизводимость опытов (отсутствие в данных грубых промахов), что позволяет, в свою очередь, найти среднюю дисперсию строчных выборок (дисперсию опытов)

$$S^2\{Y\} = (1/N) \cdot \sum_{g=1}^N S_g^2 = 2,834/8 = 0,354$$

$$\text{с } \nu_3 = 8 \cdot (5-1) = 32 \text{ степенями свободы.}$$

Определяем оценки коэффициентов уравнения регрессии:

$$b_i = (1/N) \cdot \sum_{g=1}^N z_{ij} \cdot \bar{Y}_g$$

Отсюда $b_0=58,39$; $b_1=5,65$; $b_2=9,44$; $b_3=2,85$; $b_{12} = 1,44$; $b_{13} = 1,65$; $b_{23} = 2,07$; $b_{123} = 4,27$.

Проверка гипотезы проводится с помощью t-критерия Стьюдента, который при проверке нуль-гипотезы формируется в виде:

$$t_i = |b_i| / \sqrt{S^2\{b_i\}},$$

где $S^2\{b_i\}$ – дисперсия ошибки определения коэффициента b_i . При полном и дробном факторном планировании

$$S^2\{b_i\} = S^2\{v\} / N \cdot m.$$

Если вычисленная величина параметра t_i превышает табличное значение $t_{кр}$, найденное для q-процентного уровня значимости и $\nu_3 = N(m-1)$ числа степеней свободы, то нуль-гипотеза отвергается и коэффициент считается незначимым и его следует отбросить, не включая в искомую модель.

Проверяем значимость оценок коэффициентов по критерию Стьюдента предварительно найдя их дисперсию:

$$S^2\{b_i\} = S^2\{v\} / N \cdot m = 0,00885; \nu S^2\{b_i\} = 0,094.$$

Отсюда $t_1=60,16$; $t_2=100,48$; $t_3=30,37$; $t_{12}=15,32$; $t_{13}=17,6$; $t_{23}=22,02$; $t_{123}=45,42$. В общем виде искомое уравнение регрессии имеет вид:

$$\tilde{Y} = b_0 + b_1 x_1 + \dots + b_n x_n + b_{12} x_1 x_2 + b_{1\dots n} x_1 \dots x_n.$$

Табличное значение критерия t_i [4] $t_{кр}(5\%; \nu_3=32) = 2,046$, поэтому все найденные оценки коэффициентов признаются значимыми и должны войти в модель:

$$Y = 58,39 + 5,65x_1 + 9,44x_2 + 2,85x_3 + 1,44x_1x_2 + 1,65x_1x_3 + 2,07x_2x_3 + 4,27x_1x_2x_3.$$

Для определения дисперсии адекватности по формуле необходимо сначала найти числовые значения модели для каждой g-ой строки матрицы планирования, а затем подсчитать сумму квадратов разностей между модельным значением и средним арифметическим той же строки

$$S_{ад}^2 = 0,0976/8-7 = 0,0976.$$

Определяем критерий Фишера:

$$F = S_{ад}^2 / S^2\{Y\} \leq F_{кр.}(q; \nu_{ад}; \nu_3).$$

Если вычисленное по формуле значение критерия F меньше табличного $F_{кр}$, найденного для q -процентного уровня значимости, $v_{числ} = v_{ад} = v_4 = N-d$ числа степеней свободы числителя и $v_{зн} = v_3 = N(m-1)$ числа степеней свободы знаменателя, то нуль-гипотеза принимается. В противном случае она отвергается и описание (модель) признается неадекватным объекту.

$$F=0,0976/0,354=0,275 < F_{кр}=1,$$

что доказывает адекватность найденной модели.

Таким образом, формула для расчета требуемого эффекта очистки высококонцентрированных сточных вод имеет вид:

$$\begin{aligned} \Delta = & 0,001 \cdot D \cdot E_h \cdot t - 0,028 \cdot D \cdot E_h + 0,0462 \cdot D \cdot t - 0,006 \cdot E_h \cdot t - \\ & 0,364 \cdot D + 0,222 \cdot E_h + 0,194 \cdot t + 47,1 \end{aligned}$$

Определим с помощью полученной выше математической модели оптимальную дозу реагента и сравним ее с экспериментальным значением. Исследуемая сточная вода имеет следующие характеристики: окислительно-восстановительный потенциал $E_h=100$ Мв (соответствует $x_2=+1$), температура сточной воды 25°C (соответствует $x_3=0$), исходная концентрация фосфатов $29,4$ мг/л. Согласно требованиям ПДК по сбросу в горколлектор содержание фосфатов в очищенной сточной воде должно быть не более $5,29$ мг/л, что соответствует 80% эффекту дефосфотации в точке перед азротенками первой ступени. Таким образом, получаем следующее выражение:

$$80 = 58,39 + 5,65x_1 + 9,44 + 1,44x_1; 7,09 x_1 = 12,17; x_1 = 1,71.$$

Пересчитываем относительную величину в фактическую, исходя из соотношения $x_1=0,14D-1,14$. Отсюда получаем дозу реагента $D=16,5:1$.

По результатам проведенных экспериментально-опытных исследований наибольший эффект очистки достигался при введении реагента концентрацией 800 мг/л (по Fe^{2+} – 161 мг/л) при этом доза реагента составляет $16,8:1$. Расхождение между экспериментально подобранной дозой коагулянта и рассчитанной по формуле, полученной методом ПФЭ, составляет $1,78\%$.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Алексеев, М.И. Основы постановки научных исследований по очистке сточных вод: учебное пособие для слушателей повышения квалификации преподавателей / М.И. Алексеев, Б.Г.Мишуков, Е.М. Протасовский, В.П. Голосун. – Ленинград: ЛИСИ, 1987. – 52 с
2. Славутский, Л.А. Основы регистрации данных и планирования эксперимента: учебное пособие / Л.А. Славутский. – Чебоксары: Изд-во ЧГУ, 2006. – 200 с.
3. Ходасевич, Г.Б. Планирование эксперимента: учебник для вузов[электронный ресурс]/ Г.Б. Ходасевич. – Санкт-Петербург: СПб. государственный университет телекоммуникаций, 2004. – Режим доступа (<http://pds.sut.ru/pe/>)
4. Новицкий, П.В. Оценка погрешностей результатов измерений / П.В. Новицкий, И.А. Зограф. – 2-е изд., перераб. и доп. – Ленинград : Энергоатомиздат, 1991. – 304 с.
5. Волосухин, В.А. Планирование научного эксперимента: учебник для вузов по направлению подготовки 280100 "Природообустройство и водопользование" / В.А. Волосухин, А.И. Тищенко. – 2-е изд. – Москва: РИОР, 2014 . – 174, [1] с.
6. Сидняев, Н.И. Теория планирования эксперимента и анализ статистических данных: учеб. пособие для вузов / Н.И. Сидняев. – Москва: Юрайт, 2011. – 399 с.
7. Горелова, Г.В. Теория вероятностей и математическая статистика в примерах и задачах с применением Excel / Г.В. Горелова. – Москва: Феникс, 2005. – 476 с.
8. Лагутин, М.Б. Наглядная математическая статистика: учебное пособие – 2-е изд., испр. / М.Б. Лагутин. – Москва: БИНОМ. Лаборатория знаний, 2009. – 472 с.
9. Фокичева, Е.А. Особенности построения математической модели процесса реагентной дефосфотации высококонцентрированных сточных вод: статья // Вестник гражданских инженеров. – 2012. – № 6 (35). – С. 113–120.
10. Алексеев, М.И. Фокичева, Е.А. Удаление биогенных элементов из сточных вод животноводческих комплексов в целях предотвращения эвтрофирования водных объектов // Вестник гражданских инженеров. – 2013. – № 1 (36). – С. 117–123.

ОГЛАВЛЕНИЕ

ПРЕДИСЛОВИЕ.....	3
ВВЕДЕНИЕ	5
1. НАБЛЮДЕНИЕ И ЭКСПЕРИМЕНТ. МОДЕЛИРОВАНИЕ.....	7
1.1. Наблюдение и эксперимент	7
1.2. Моделирование.....	8
2. ИЗМЕРЕНИЯ. ОЦЕНКА ПОГРЕШНОСТЕЙ ИЗМЕРЕНИЙ.....	19
2.1. Виды измерений и причины ошибок	19
2.2. Классификация погрешностей измерений	19
2.3. Вероятностная оценка случайной погрешности и построение функциональных зависимостей при многократных измерениях	21
3. ОБРАБОТКА РЕЗУЛЬТАТОВ ЭКСПЕРИМЕНТА СТАТИСТИЧЕСКИМИ МЕТОДАМИ. ПОСТРОЕНИЕ ФУНКЦИОНАЛЬНЫХ ЗАВИСИМОСТЕЙ.....	25
3.1. Основные законы распределения случайных величин, применяемые при обработке результатов эксперимента статистическими методами.....	25
3.2. Построение функциональных зависимостей при многократных измерениях	30
4. ПЛАНИРОВАНИЕ ЭКСПЕРИМЕНТА. ОБЩИЕ ПОЛОЖЕНИЯ	35
4.1. История планирования эксперимента.....	35
4.2. Основные понятия и определения теории планирования эксперимента.....	36
4.3. Основы планирования многофакторного эксперимента.....	39
4.4. Критерии оптимальности и типы планов	45
5. ТЕСТЫ.....	48
ГЛОССАРИЙ.....	60
ПРИЛОЖЕНИЕ 1	62
ПРИЛОЖЕНИЕ 2	65
БИБЛИОГРАФИЧЕСКИЙ СПИСОК	70

Учебное издание

Фокичева Елена Александровна,
Алексеев Михаил Иванович

**ПЛАНИРОВАНИЕ ЭКСПЕРИМЕНТА И ОБРАБОТКА
РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЙ**
Учебное пособие

Подписано в печать 23.06.2014. Формат 60 × 84/16
Бумага писчая. Печать офсетная.
Усл. п.л. 4,5. Тираж 20 экз. Заказ №

Отпечатано: РИО, ВоГУ 160000, г. Вологда, ул. С. Орлова, 6