Министерство образования и науки Российской Федерации

АМУРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

И.В.Абакумова

МЕТОДЫ И СРЕДСТВА ИССЛЕДОВАНИЯ ТЕХНОЛОГИЧЕСКИХ ПРОЦЕССОВ

Учебное пособие

Рекомендовано ДВ РУМЦ Министерства образования и науки в качестве учебного пособия для студентов высших учебных заведений специальностей 260704, 260901, 260902.

Благовещенск
Издательство АмГУ
2010

ББК 37.23я73
А 13

Рекомендовано учебно-методическим советом АмГУ

Рецензенты:

Розанова Е.А., профессор кафедры сервиса и моды Владивостокского государственного университета экономики и сервиса, канд.техн.наук;

Самуйло В.В., профессор кафедры ЭМТиАП Дальневосточного государственного аграрного университета, д-р.техн.наук

Абакумова И.В
А 13 Методы и средства исследования технологических процессов**:** учебное пособие/ И.В.Абакумова. – Благовещенск: Изд-во АмГУ, 2010. – 114 с.

Учебное пособие посвящено основным методам и средствам исследования технологических процессов и материалов в текстильной и легкой промышленности. В нем рассматриваются статистические характеристики и их признаки, изложены математические методы контроля и исследования технологических процессов и текстильных материалов, методы получения линейных и нелинейных регрессионных и корреляционных математических моделей при обработке результатов эксперимента, а также методы анализа полученных математических моделей.

Пособие может быть использовано при изучении дисциплины «Методы и средства исследования», а также для выполнения курсовых и дипломных работ студентами инженерных специальностей 260901 «Технология швейных изделий», 260902 «Конструирование швейных изделий», 260704 «Технология текстильных изделий».

1. ВИДЫ И ЭТАПЫ НАУЧНО-ИССЛЕДОВАТЕЛЬСКИХ РАБОТ

Цель научного исследования – всестороннее изучение объекта, процесса или явления, их структуры, связей и отношений на основе разработанных в науке принципов и методов познания, а также внедрение в производство полученных результатов.

Любое научное исследование изучает объект или предмет.

Объектом научного исследования является любая система, например, «человек-одежда-среда», швейное предприятие, цех, технологический процесс и пр.

Предмет исследования – это элементы системы, их закономерности, различные свойства, качества. Например, новые виды материалов, свойства тканей и т.д.

1.1. Классификация научно-исследовательских работ

Научно-исследовательские работы (НИР) подразделяются на теоретические, экспериментальные и теоретико-экпериментальные.

В *теоретических работах* на основе аналитических исследований сущности изучаемого процесса с использованием законов природы устанавливаются его закономерности и прогнозируются оптимальные условия работы действующего или вновь создаваемого процесса.

В *экспериментальных работах* все перечисленные выше задачи решаются путем эксперимента в лабораторных или производственных условиях.

Сочетание теоретических и экспериментальных частей научно-исследовательской работы способствует более глубокому решению задач научного исследования.

По решаемым задачам НИР в легкой и текстильной промышленности подразделяются на следующие виды:

1) Теоретико-экспериментальные, раскрывающие закономерности технологических процессов и определяющие оптимальный режим работы машин и механизмов с целью повышения эффективности технологических процессов,

улучшения качества продукции, совершенствования конструкции машин и механизмов, автоматизации производства, внедрения САПР одежды.

2) Экспериментальные работы по испытанию вновь разработанных материалов, видов одежды с целью определения надежности, долговечности, удобства носки, а также для испытания новых машин и механизмов с целью повышения их производительности, удобства обслуживания и качества вырабатываемой продукции.

3) Поисково-исследовальские работы, направленные на разработку новых технологий и технологических процессов на основе более эффективного использования применяемых в промышленности видов энергии – механической, тепловой и др., а также на основе новых принципов использования этих и других видов энергии и применения достижений современной науки.

4) Поисковые работы направлены на создание новых материалов, нового ассортимента нитей, тканей, трикотажа и других текстильных изделий, к поисковым работам относятся также работы по рациональному использованию сырья.

5) Исследовательские работы, связанные с изучением факторов, определяющих качество и эксплуатационные свойства изделий, а также работы по улучшению методов испытания материалов и разработке новых методов и приборов с целью создания новых стандартов или технических условий.

6) Экспериментально-исследовательские работы направлены на разработку новых методов исследования технологических процессов и средств для измерения параметров одежды и материалов, а также основных свойств, характеризующих технологически процесс.

Экспериментальные исследовательские работы взависимости от условий проведения и принятого объема делятся на лабораторные и производственные. Лабораторные исследования отличаются малым объемом испытаний и малым фронтом наблюдений.

Методы исследований

Метод – это способ теоретического исследования или практического осуществления чего-либо. Метод – это способ достижения цели.

Методика – совокупность методов исследования для практического применения чего-либо.

Любой метод имеет субъективную и объективную стороны. Метод исследования объективен, поскольку в разрабатываемой теории позволяет отражать действительность и ее взаимосвязи, т.е. является программой построения и практического применения теории.

Одновременно метод субъективен, так как зависит от субъективных особенностей исследователей.

Существует 4 уровня методов научного исследования:

1) эмпирический;
2) экспериментально-теоретический;
3) теоретический;
4) мето-теоретический.

Эмпирический – это исследовательский метод, основанный на описании фактов, без последующих заключений и теоретических обобщений. Методы эмпирического уровня: наблюдения, сравнения, счет, измерения, анкетный опрос, собеседование, тестирование. Методы этой группы конкретно связаны с изучаемыми явлениями и используются на этапе формирования научной гипотезы.

Методы экспериментально-теоретического уровня: эксперимент, анализ и синтез, индукция и дедукция, моделирование, логические методы. Эти методы позволяют обнаруживать достоверные факты, накапливать их и производить тщательную проверку.

Первоначальная систематизация фактов и их анализ проводятся уже в процессе наблюдений, собеседования, экспериментов.

Анализ и синтез – взаимосвязанные методы, первый метод расчленяет сложные явления на составляющие части с рассмотрением свойств этих частей, синтез же соединяет отдельные свойства явления в единое целое.

Индукция и дедукция – методы формальной логики, в первом методе делают умозаключение о явлении, исходя из накопленных фактов об элементах этого явления, во втором – делают умозаключение о свойствах одного из элементов явления на основании знания общих свойств всего множества.

Методы теоретического уровня сводятся к логическим обоснованиям собранных фактов, выработке понятий, суждений, умозаключений. К методам этого уровня относятся абстрагирование, идеализация, формализация, аксиоматика, обобщение и пр.

Абстрагирование – мысленное отвлечение от несущественных связей, отношений, выделение несколько сторон явления (например, замена объекта исследования более простым, представляющим упрощенную модель).

Идеализация – это мысленное конструирование объектов, которые практически не существуют (например, идеальный газ, абсолютно черное тело и т.д.).

Формализация – это отображение объекта исследования с помощью какой-либо системы, в знаковой форме какого-либо языка (например, математики, химии и т.д.).

Аксиоматика – это способ получения научной теории, при котором некоторые утверждения принимаются без доказательств (аксиомы) и затем используются для получения остальных законов.

Обобщение – это систематизация выявленных свойств явления.

На теоретическом уровне познания решаются задачи дальнейшего согласования теоретически разработанных систем с полученным новым экспериментальным материалом.

К *методам мета-теоретического уровня* относятся диалектический метод, метод системного анализа и пр. С помощью этих методов исследуются сами теории и разрабатываются пути их построения, устанавливаются границы их применения, способы введения новых понятий, обосновываются пути объединения в единое целое несколько теорий.

1.3. Этапы научно-исследовательской работы

Любая научно-исследовательская работа (УИРС или дипломная работа по исследовательской тематике) выполняется в определенной последовательности.

I этап заключается в выборе объекта исследования, обосновании цели и задач исследования, формулировании темы исследования, общем ознакомлении с проблемой. На данном этапе указываются причины разработки (обоснование темы исследования), приводится краткий обзор литературных источников, в котором описывается уже достигнутый уровень исследований и полученные ранее результаты. Особое внимание уделяется еще не решенным вопросам, актуальности (т.е. важности) и значимости работы для отрасли (или для предприятия). Обзор литературы позволяет наметить методы решения, этапы исследования, определить конечные цели выполнения темы. Сюда также входит анализ патентной литературы по теме исследования.

II этап включает составление технико-экономического обоснования (ТЭО) темы НИР. На стадии составления технико-экономического обоснования устанавливается область использования ожидаемых результатов НИР, возможности их практической реализации в данной отрасли (на данном предприятии), предполагаемый эффект от внедрения результатов исследования. Кроме экономического эффекта, указывается предполагаемый социальный эффект (например, рост производительности труда, качества выпускаемой продукции, повышение уровня безопасности труда и производственной санитарии, экологический эффект и пр.). В результате составления ТЭО делается вывод о целесообразности и необходимости выполнения НИР.

Технико-экономическое обоснование утверждается заказчиком. После этого конкретизируются цели и задачи исследования, составляется библиографический список отечественной и зарубежной литературы, аннотация литературных источников.

III этап – проведение теоретических исследований. На данном этапе осуществляется изучение физической сущности изучаемого явления или предмета. В результате обосновывается физическая модель, разрабатывается математическая модель явления или предмета исследования, анализируются предварительные результаты. Получение математической модели, как правило, осуществляется с использованием современных компьютерных программ. Производится анализ полученной математической модели, определяются ее экстремальные условия с целью последующей оптимизации технологического процесса. Оценка оптимальных условий производится по критериям, выражающимся, например, в максимальном объеме вырабатываемой продукции с 1 м^2, минимальной стоимости продукции при определенной производительности, минимальном расходе материалов и других ресурсов.

IV этап – организация и проведение экспериментальных исследований по предварительно разработанной методике и программе эксперимента. Эффективность данного этапа существенно зависит от выбора средств измерения. При разработке методики проведения эксперимента составляется рабочий план, в котором указывается объем экспериментальных работ, методы, используемая техника и приборы, трудоемкость и сроки.

По завершению теоретических и экспериментальных исследований проводится общий анализ полученных результатов, осуществляется сопоставление гипотезы с результатами эксперимента. В случае расхождения теоретических представлений и экспериментальных данных упорядочиваются теоретические модели либо проводятся дополнительные эксперименты. Затем формируются научные и производственные выводы, составляется научно-технический отчет.

V этап – внедрение результатов исследования в производство и определение действительной экономической эффективности.

2. ТЕОРЕТИЧЕСКИЕ ИССЛЕДОВАНИЯ. МОДЕЛИРОВАНИЕ В НАУЧНОМ И ТЕХНИЧЕСКОМ ТВОРЧЕСТВЕ

2.1. Задачи и методы теоретического исследования

Цель теоретического исследования состоит в выделении в процессе изучения существующих связей между исследуемым объектом и окружающей средой, объяснении и обобщении результатов эмпирических исследований общих закономерностей и их формализация (описание с помощью формул).

Задачами теоретического исследования являются:

- обобщение результатов исследования;
- нахождение закономерностей путем обработки экспериментальных данных;
- расширение результатов исследования на ряд подобных объектов без повторения всего объема исследования;
- изучение объекта, недоступного для непосредственного экспериментального исследования;
- повышение надежности экспериментального исследования объекта (обоснование параметров и условий наблюдения, точности измерения).

Теоретические исследования включают:

- анализ физической сущности явления;
- формулирование гипотезы исследования;
- построение физической или математической модели;
- проведение математического исследования;
- анализ теоретических решений;
- формулирование выводов.

2.2. Структура решения задачи

Структурно любая задача включает условия и требования.

Условия – это определение информационной системы, из которой следует исходить при решении задачи.

Требования – это цель, к которой следует стремиться в результате решения. Условия и требования могут быть исходными, привлеченными и искомыми (рис. 1).

Исходные условия – даются в первоначальной формулировке задачи. Если их оказывается недостаточно для решения задачи, исследователь вынужден привлекать дополнительные данные – *привлеченные*. Искомые данные или *искомые условия* – это то, что требуется отыскать в процессе решения задачи.

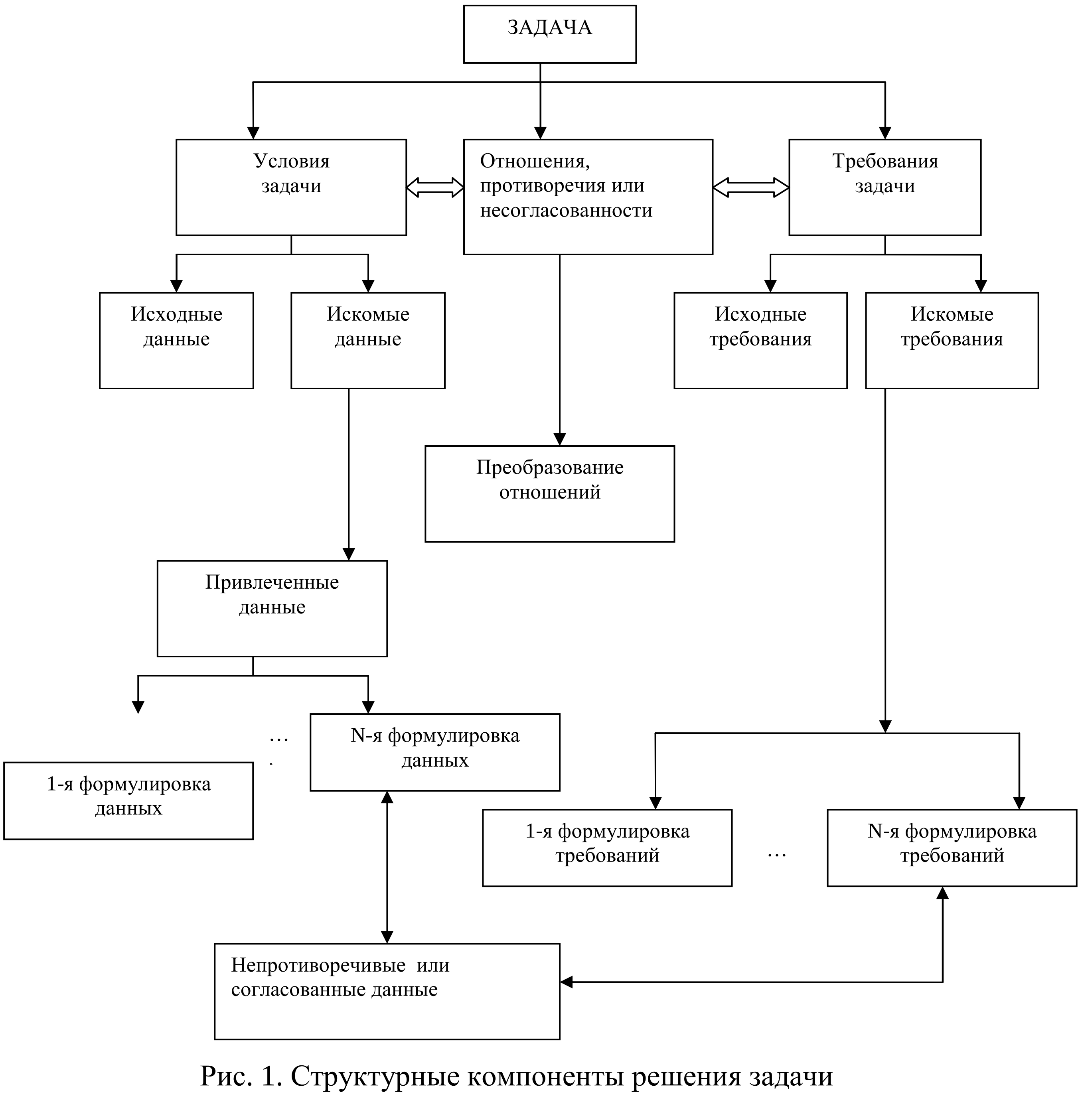


Рис. 1. Структурные компоненты решения задачи

Процесс проведения теоретических исследований включает в себя несколько стадий:

1) Оперативная стадия – включает проверку возможных изменений в среде, окружающей объект, анализ возможности устранения противоречий, анализ возможности переноса решения задачи из других отраслей знаний.

2) Синтетическая стадия – на этой стадии определяется влияние изменения одной части объекта на построения других его частей, определяются необходимые изменения других объектов, работающих совместно с данным объектом, оценивается возможность применения измененного объекта по-новому.

3) Стадия постановки задачи – здесь определяется конечная цель решения задачи, уточняются требования практической реализации полученного результата.

4) Аналитическая стадия – определение идеального конечного результата.

Решение теоретических задач должно носить творческий характер. Творческие решения часто не укладываются в заранее намеченные планы. Часто удачное решение, на которое не давит груз известных решений, возникают у специалистов смежных областей знаний. Собственные творческие мысли (оригинальные решения) возникают тем чаще, чем больше сил, труда, времени затрачивается на обдумывание путей решения теоретической задачи, чем глубже научный работник увлечен исследовательской работой.

При проведении теоретических исследований, основанных на общенаучных методах анализа и синтеза, широко используются методы расчленения и объединения исследуемой системы (объекта, явления).

Метод расчленения основан на выделении существенных и несущественных параметров, основных элементов и связей между ними. После расчленения объекта изучается вид взаимосвязи элементов и осуществляется моделирование этих элементов, после чего элементы объединяются в сложную модель объекта. На всех этапах построения модели объекта производится его упрощение, и вводятся допущения. Неверные допущения могут привести к серьезным ошибкам при формулировании теоретических выводов.

Любая модель – это естественный или искусственный объект, находящийся в соответствии с изучаемым объектом или какой-либо из его сторон. Модели позволяют проводить научные исследования различных процессов, проверять выводы и получать более наглядное и полное представление о процессе, чем это можно было бы сделать на основании расчета.

При построении моделей объекта исследования должны использоваться наиболее общие принципы и закономерности. Это позволяет учесть все допущения, принятые при получении формализованных теорий и точно определить область их применения.

Противоположным методом является *метод объединения* и связанный с ним комплексный подход к изучению объекта, которые чаще всего объединяются под названием «общая теория систем» или «системный подход».

Общая теория систем рассматривает подробно все составные элементы системы обязательно во взаимосвязи. Например, при изучении системы человек-одежда рассматриваются все свойства проектируемой одежды с точки зрения потребителя при обязательном рассмотрении защиты человека от неблагоприятных факторов среды.

Общая теория систем базируется на трех постулатах:

1) функционирование систем любой природы может быть описано на основе рассмотренных формальных структурно-функциональных связей между определенными элементами системы. Влияние материала, из которого состоят элементы системы, проявляются в характеристиках системы (ее структуре, динамике);

2) организация системы может быть определена на основе наблюдений, проведенных фиксированием состояний только тех элементов системы, которые непосредственно взаимодействуют с ее окружением;

3) организация системы полностью определяет ее функционирование и характер взаимодействия с окружающей средой.

Теоретическое исследование завершается формированием теории, не обязательно связанной с формированием математического аппарата. Теория проходит в своем развитии различные стадии – от качественного объяснения и количественного измерения процессов до их формализации и в зависимости от стадии может быть представлена как в виде качественных правил, так и в виде математических уравнений.

3. ОСНОВНЫЕ ПОЛОЖЕНИЯ НАУЧНОГО ЭКСПЕРИМЕНТА

3.1. Классификация, типы и задачи эксперимента

Важнейшей составной частью научных исследований является эксперимент, основа которого – научно поставленный опыт с точно учитываемыми и управляемыми условиями. Само слово эксперимент обозначает опыт, целенаправленное наблюдение.

Основной целью эксперимента является выявление свойств исследуемых объектов, проверка справедливости гипотез и на этой основе – широкое и глубокое изучение темы научного исследования.

Особенности постановки и организации эксперимента определяются его назначением. Эксперименты различаются:

- по способу формирования условий: естественные, искусственные;
- по целям исследований: преобразующие, констатирующие, контролирующие, поисковые, решающие;
- по организации проведения: лабораторные, натурные, полевые, производственные;
- по структуре изучаемых объектов и явлений: простые, сложные;
- по характеру внешних воздействий на объект исследования: вещественные, энергетические, информационные;
- по контролируемым величинам: активные, пассивные;
- по характеру изучаемых явлений: технологические, социологические.

Для проведения эксперимента любого типа необходимо:

- выбрать гипотезу, подлежащую проверке;
- разработать программу эксперимента;
- определить способы и приемы вмешательства в объект исследования;

- обеспечить условия для проведения экспериментальных работ;

- разработать пути для считывания результатов эксперимента;

- подготовить средства эксперимента (приборы, установки и т.д.);

- обеспечить эксперимент необходимым обслуживающим персоналом.

Особое значение имеет правильная разработка методики эксперимента. Методика – это совокупность математических и физических операций, размещенных в определенной последовательности, в соответствии с которой достигается цель исследования.

Перед каждым экспериментом составляется план (программа) эксперимента, который включает:

- цель и задачи эксперимента;

- выбор варьирующих параметров и факторов;

- определение объема эксперимента, числа опытов;

- порядок реализации опытов, определение последовательности изменения факторов, задание интервалов между будущими экспериментальными точками;

- обоснование средств применения, способов обработки и анализа результатов эксперимента.

Применение математической теории эксперимента позволяет уже при планировании определенным образом оптимизировать объем экспериментальных исследований и повысить их точность.

Перед началом эксперимента необходимо установить варьируемые факторы, т. е. установить основные и второстепенные характеристики, влияющие на исследуемый процесс, проанализировать расчетные схемы процесса. На основе этого анализа факторы устанавливаются в убывающий по важности для данного эксперимента ряд.

Иногда бывает трудно установить принадлежность тех или иных факторов к основным или второстепенным, и в таких случаях ставится предварительный эксперимент.

3.2. Подготовка и проведение предварительного эксперимента

Подготовка к проведению предварительного эксперимента включает ряд организационных и технических мероприятий.

Необходимо проверить свойства сырья и материалов, установить их соответствие задачам исследования, проверить состояние оборудования, стендов, приборов. Если в работе применяются новые средства исследований, то проводятся пробные опыты по разработанной и принятой методике, выявляются возможности осуществления принятой методики, а также неучтенные особенности эксперимента и возможные ошибки или погрешности. При использовании новых измерительных устройств проводится их тарировка и определяется точность показаний.

По результатам пробных опытов выявляется необходимость в доработке конструкции стенда, измерительных и регистрирующих устройств, вносятся соответствующие поправки в методику эксперимента.

После проведения серии опытов для каждой изучаемой закономерности необходимо обрабатывать результаты опытов с тем, чтобы в случае необходимости можно было исправить и дополнить методику исследования.

3.3. Средства и методы измерения

Важным моментом в проведении эксперимента является выбор средств измерений. Они должны соответствовать методике научно-исследовательской работы, ее целям и задачам, обеспечивать высокую производительность труда, обеспечивать требуемое качество экспериментальных работ (т.е. заданную степень точности, высокую воспроизводимость и надежность).

Средства измерения – это совокупность технических средств, имеющих нормируемые погрешности, дающих экспериментатору необходимую информацию об изучаемом объекте. В качестве средств измерения могут быть использованы меры, измерительные приборы, измерительные преобразователи, измерительные информационные системы.

Мерой называется средство измерения, предназначенное для воспроизведения заданного размера какой-либо физической величины (разновесы, линейки, магазины сопротивлений и др.). Измерение с помощью меры производится методом сравнения – измеряемая величина сравнивается с мерой.

Измерительным прибором называется средство измерения, предназначенное для выработки сигналов измерительной информации, т.е. сигналов, функционально связанных с измеряемыми величинами (например, массой, плотностью, линейной скоростью и т.п.). Эти сигналы измерительный прибор представляет в форме, доступной для непосредственного восприятия наблюдателем.

Измерительным преобразователем называется средство измерения, предназначенное для выработки сигнала измерительной информации в форме удобной для передачи и хранения. Это различного рода датчики физических величин и усилительно-преобразовательные устройства.

Измерительной информационной системой называется совокупность функционально объединенных средств измерения и вспомогательных устройств, предназначенных для автоматического сбора измерительной информации от нескольких источников, а также каналов связи.

Методы измерения можно разделить на прямые, косвенные и совокупные. При *прямых измерениях* искомую величину устанавливают непосредственно из опыта.

При *косвенных измерениях* величину определяют как функцию от иных величин, измеренных прямым методом. Этот способ используется в случае, если прямое измерение искомой величины слишком сложно или невозможно. Например, такую величину, как плотность вещества, чаще всего невозможно непосредственно измерить, и для определения плотности измеряют две другие величины – массу и объем.

Совокупные измерения состоят из ряда прямых измерений, выполненных или при различных условиях, или при различных сочетаниях измеряемых величин. Например, для измерения количества радиоактивного вещества, содержащегося в некотором растворе месяц назад, никакие прямые или косвенные измерения ответа не дадут. Однако, измеряя изменение концентрации раствора во времени (с течением времени радиоактивность падает), можно получить закон, по которому она меняется, и с помощью полученного уравнения определить концентрацию раствора на любой момент времени.

И косвенные, и совокупные измерения неотъемлемой частью включают в себя прямое измерение, выполненное тем или иным измерительным прибором. Само измерение может быть выполнено двумя методами - непосредственной оценки и сравнения.

Метод непосредственной оценки соответствует определению значения величины непосредственно по отсчетному устройству измерительного прибора. Метод непосредственной оценки применяется в антропометрических исследованиях при измерении различных признаков фигуры. Для этого используются сантиметровая лента, толстотный циркуль, голиометр, антропометр, набор специальных линеек, медицинские весы.

Метод сравнения заключается в том, что измеряемую величину сравнивают с величиной, воспроизводимой мерой (например, взвешивание). Метод сравнения может быть нулевым, дифференциальным или методом замещения.

Нулевой метод заключается в том, что на индикатор прибора действует одновременно две физических величины – измеряемая и воспроизводимая мерой. При этом измерение производится в момент, когда суммарное действие этих мер равно нулю, т.е. когда действие измеряемой меры компенсируется действием воспроизводимой меры.

Дифференциальный метод отличается от нулевого только тем, что мера не полностью компенсирует измеряемую величину, и эта нескомпенсированная разность фиксируется индикатором. При этом нескомпенсированная разность обычно составляет доли процента от измеряемой величины.

Наиболее точным из методов сравнения является *метод замещения*, при котором измеряемая величина замещается в измерительной схеме мерой. Мера после установки настраивается таким образом, чтобы измерительная схема пришла в состояние, идентичное состоянию до замены, т.е. чтобы показания приборов до и после замены были одинаковыми.

Различают также абсолютные и относительные измерения..

Абсолютные – это прямые измерения в единицах измеряемой величины. *Относительные измерения* представляют собой отношение измеряемой величины к одноименной величине, и представляются в процентах.

Точностью измерения называется степень соответствия результата измерения действительному значению измеряемой величины. Оценивается точность двумя показателями:

1) *ошибка измерения* – это разность между результатом измеряемой величины и действительным ее значением;

2) *надежность* – это вероятность того, что действительное значение измеряемой величины отличается от результата измерения не более чем на значение указанной ошибки.

В основу теории ошибок положено предположение о том, что при большом числе измерений случайные погрешности измеряемой величины встречаются одинаково часто. Большие погрешности встречаются реже. Чем больше число измерений, тем вероятнее, что среднее арифметическое всех результатов измерений приближается к истинному значению измеряемой величины. Появление того или иного результата измерения описывается законом нормального распределения.

4. СТАТИСТИЧЕСКИЕ СОВОКУПНОСТИ И ИХ ПРИЗНАКИ

Совокупность, охватывающая все массовое явление, называется *генеральной совокупностью*. Генеральная совокупность состоит из множества однородных элементов, имеющих общие свойства, оцениваемые статистическими характеристиками (например, вся продукция, все сырье некоторой фабрики, все население, которое обеспечивается одеждой и т.п.).

Признаком статистической совокупности объектов называется то или иное свойство, характеризующее все элементы этой совокупности. Например, признаками волокон хлопка данной кипы являются: длина волокна, прочность, зрелость, цвет и т.д.; признаками образцов пряжи являются: прочность, линейная плотность, сорт и т.д.; признаками совокупности людей некоторой области являются рост, обхват груди, длина стопы, возраст, пол и т.д.

Признаки могут быть:

1) *количественными*, т.е. поддающимися измерению (например, прочность пряжи, рост человека);

2) *качественными*, когда можно фиксировать лишь наличие или отсутствие некоторого признака (например, сортность продукции, пол человека).

Количественные признаки могут изменяться или *непрерывно*, или *дискретно* в зависимости от того, могут ли они принимать любые значения в некотором интервале или лишь только какие-то отдельные значения. В частности, длина волокна и рост человека являются непрерывно изменяющимися признаками, а число обрывов нити в час или количество выпускаемых в день изделий – признаки дискретные.

Задание

В задачах приведены различные статистические совокупности объектов и признаки, по которым их можно изучать. Определить какие из этих признаков являются количественными, качественными, непрерывными и дискретными.

1) Початки с пряжей. Признаки: вес, суровая или крашеная пряжа, длина пряжи, средняя толщина, средняя плотность, сорт.
2) Куски трикотажного полотна, выпускаемые фабрикой за сутки. Признаки: метраж, сортность, наличие или отсутствие брака в куске, гладкое полотно или с рисунком, суровое или крашеное, ширина, стоимость.
3) Трикотажные машины на фабрике. Признаки: отечественные или импортные, скорость, ширина игольницы, диаметр цилиндра, кругловязальные или плосковязальные, вид вырабатываемого переплетения, класс, стоимость.
4) Партия изделий. Признаки: размер, мужское или женское, цена, рост, сорт, модель.

5) Студенты вуза. Признаки: пол, возраст, рост, количество отличных отметок, стипендиаты или не стипендиаты, материальное положение, семейное положение, курс, группа, факультет.

Ответы на задания оформите в виде табл.1.

Таблица 1

Признаки статистических совокупностей

<table>
<tr><th rowspan="2">Статистическая совокупность</th><th rowspan="2">Признаки</th><th colspan="2">Количественные</th><th rowspan="2">Качественные</th></tr>
<tr><th>непрерыв-ные</th><th>дискрет-ные</th></tr>
<tr><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td><td></td></tr>
</table>

Основным методом математической статистики является *выборочный метод*, состоящий в том, что все массовое явление, т.е. всю генеральную совокупность, изучают с помощью *выборки* не очень большого объема, называемой *выборочной совокупностью*. Выборочная совокупность состоит из части объектов генеральной совокупности, отобранной определенным образом.

В частности, из такой выборочной совокупности определяют *выборочные характеристики* с тем, чтобы по ним судить о *генеральных характеристиках*.

Так для определения различных признаков от партии материала (генеральной совокупности) отбирают небольшую выборку (пробу), в которой должны правильно отражаться свойства всей партии (рис.2).

5. МЕТОДЫ ОТБОРА ВЫБОРОК

Выборка должна, возможно более походить на генеральную совокупность для того, чтобы по ней можно было судить о последней. Поэтому если в генеральной совокупности некоторые признаки встречаются в определенных пропорциях, то выборка будет тем лучше, чем ближе в ней пропорции тех же признаков будут к соответственным пропорциям в генеральной совокупности.

Выборка, достаточно точно воспроизводящая пропорции генеральной совокупности, называется *репрезентативной*, т.е. представляющей генеральную совокупность.

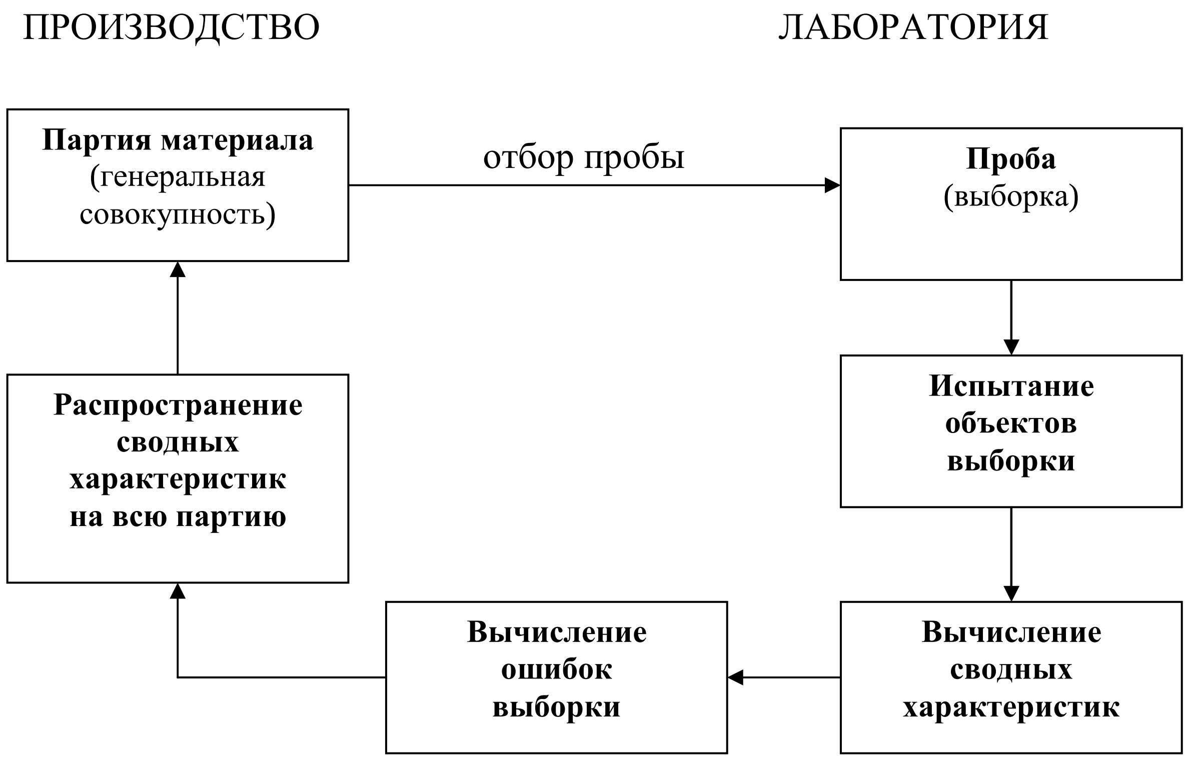


Рис.2. Общая схема лабораторного анализа и оценки свойств текстильных материалов

Требования к выборочной совокупности:

1) должна быть репрезентативной;

2) должна иметь достаточный объем, чтобы по результатам испытания выборочной совокупности можно было судить о генеральной совокупности;

3) должна быть отобрана по определенным правилам.

На рис. 3 приведена классификации методов отбора выборок.

5.1. Одностепенные методы отбора.

Эти методы отбора предусматривают выборку из всей генеральной совокупности, без предварительного деления ее на части.

Случайный метод отбора

В действительных условиях мы обычно не знаем численности и пропорции различных классов генеральной совокупности, поэтому не можем сделать репрезентативную выборку, учитывая эти пропорции.

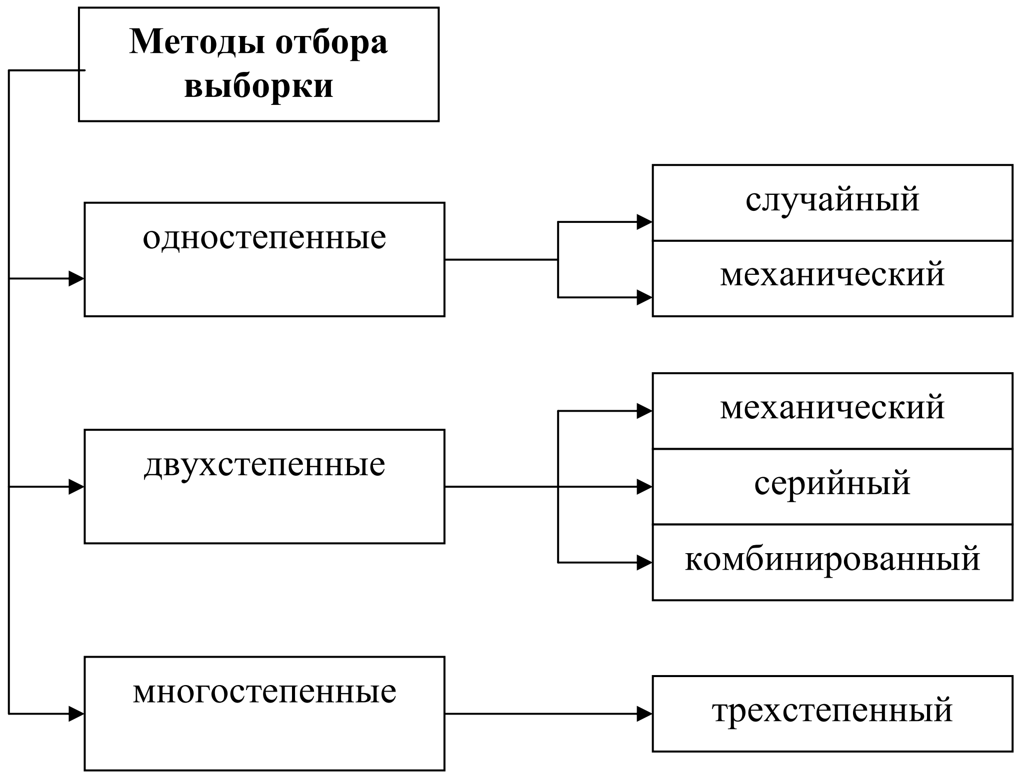


Рис.3. Схема классификации методов отбора выборок

Чтобы тем не менее, получить такую выборку, мы делаем ее случайной и отбираем ее таким способом, который обеспечивает каждому отдельному члену генеральной совокупности одинаковую возможность попасть в выборочную совокупность. Тогда, как показывает опыт, пропорции исследуемых признаков в выборочной совокупности будут воспроизводить более или менее точно те же пропорции в генеральной совокупности и выборочная совокупность будет репрезентативной.

Ошибочно полагать, что выполнение только одного из этих условий позволяет получить репрезентативную выборку. Например, при отборе наугад нескольких паковок пряжи из вскрытого ящика соблюдается первое условие случайного отбора, но не осуществляется одинаковая возможность попадания каждой паковки ящика в выборку. Очевидно, что имеется значительно большая вероятность того, что паковки будут отобраны с поверхности открытого ящика, и очень малая вероятность их извлечения со дна.

Образование случайных выборок при помощи карточек и таблиц случайных чисел

Наиболее простой способ получения случайной выборочной совокупности заключается в следующем. Члены генеральной совокупности нумеруются один за другим, и затем из полученных номеров наудачу берется столько, сколько членов должно войти в выборочную совокупность. Члены генеральной совокупности, имеющие отобранные таким образом номера, и образуют случайную выборку.

Чтобы выбирать наудачу номера, можно применять два способа:

1) способ карточек;

2) таблица случайных чисел.

При способе *карточек* каждый номер записывается на карточку. Карточки делаются одинаковыми и тщательно перемешиваются, после чего из них вынимается одна какая-нибудь карточка, и номер ее записывается. Затем она возвращается в пачку, которая снова тщательно перемешивается, и вновь вынимают одну карточку.

С помощью карточек можно получить как повторную выборку, или выборку с повторением, так и выборку без повторения (в последнем случае каждый вынутый номер откладывается в сторону, и дальнейшие номера вынимаются уже без него).

Однако практический опыт показывает, что невозможно перемешать карточки так, чтобы при каждом извлечении равновозможность появления любой из них была гарантирована. Основываясь на практике статистических опытов с карточками, шарами и т.д., во многих случаях можно сказать, что они дают не те результаты, которые должны наблюдаться согласно законам случая.

Таблица случайных чисел используется для составления случайных выборок. Этот способ лишен недостатков, присущих способу карточек, шаров и т.д. Таблицы случайных чисел впервые были составлены в Англии в 1927г. Типпетом (учеником Пирсона). В СССР подобные таблицы, но только небольшого объема были составлены в 1931 г. А.К.Митропольским, а затем в 1936г. были составлены таблицы очень большого объема М.Кадыровым (часть таблиц М.Кадырова приведена в литературе [1, 2] и в приложении 7).

Таблицы случайных чисел М.Кадырова состоят из четырехзначных чисел, составленных случайно из чисел 0,1...9. Случайность расположения чисел 0,1...9 в таблицах заключается в том, что нет никакого закона в их расположении, но вместе с тем каждое из этих десяти чисел на странице встречается приблизительно одинаковое число раз.

Покажем на примерах, как пользоваться таблицами случайных чисел.

Пример 1. Пусть имеется 1000 бобин пряжи из них 30% (300 штук) бракованные (иной толщины), причем эти дефектные бобины более или менее равномерно распределены среди 1000 бобин.

В этом примере нам известен объем генеральной совокупности и процент бракованных бобин. Покажем теперь, что выборка, составленная с помощью таблиц случайных чисел, является репрезентативной. Составим выборку из 100 бобин и посмотрим, какой процент дефектных бобин будет в выборке.

Составлять выборку будем в такой последовательности:

1) пронумеруем все наши 1000 бобин цифрами от 0 до 999, при этом нумеровать бобины будем так, чтобы номера 0,1...299 имели бы бобины бракованные, а номера 300,301...999 – хорошие бобины;

2) открываем таблицу случайных чисел и с любого ее места выписываем 100 трехзначных чисел (т.к. нумерация наших бобин трехзначная), опуская одинаковые номера, если они встречаются.

Пусть этими номерами будут следующие:

857	457	499	762	431	698	038	558	653	573
609	179	974	011	098	805	516	296	149	815
070	692	696	203	350	900	451	318	798	111
933	199	183	421	338	104	190	150	449	320
165	617	369	069	248	960	652	367	168	261
549	627	832	609	577	805	999	218	878	535
097	389	524	134	388	970	030	033	712	775
814	301	167	551	566	585	781	822	903	417
253	537	100	994	830	516	029	223	644	249
085	732	673	851	146	830	265	973	317	016

Из 100 выписанных чисел 33 приходятся на номера от 0 до 299 (эти числа подчеркнуты в таблице), т.е. в выборке оказалось 33% бракованных бобин, что достаточно близко к проценту бракованных бобин в генеральной совокупности. Если мы возьмем 100 трехзначных чисел в любом другом месте таблицы, то увидим, что процент бракованных бобин тоже будет близки к 30%.

Во взятом примере мы заранее знали, что в генеральной совокупности 30% бракованных бобин, т.е. выборка оказалась репрезентативной.

Этот пример показывает, что можно судить достаточно точно о проценте бракованных бобин в генеральной совокупности, состоящей из 1000 бобин, по их проценту в выборке, состоящей из 100 бобин, выбранных при помощи таблицы случайных чисел, и в том случае, когда этот процент нам не был известен заранее до опыта.

В таком случае (при неизвестном заранее проценте бракованных бобин в генеральной совокупности) обратная задача была бы такова: найти процент бракованных бобин в генеральной совокупности по проценту их в выборке, состоящей из 100 бобин.

Для решения этой задачи нам надо было бы отобрать с помощью таблиц случайных чисел для испытания те 100 бобин, номера которых попали в выборку. При этом мы бы получили процент бракованных бобин, близкий к 30%. А так как мы ранее на примере убедились, что выборка, составленная с помощью таблиц случайных чисел, является репрезентативной, то, следовательно, мы можем утверждать, что и в генеральной совокупности процент бракованных бобин близок к 30%.

Пример 2. Отобрать из 100 некоторых предметов (например, изделий) случайным образом 20.

Для этого необходимо выполнить следующее:

1) пронумеруем наши предметы от 0 до 99;

2) из любого места таблицы случайных чисел выписываем 20 различных двухзначных чисел, опуская одинаковые, если они встречаются, числа 01, 02 и т.п. считаются двухзначными.

Выпишем эти числа, опуская одинаковые, если они встречаются, с первой строки таблицы случайных чисел, приведенной в приложении 7:

89	83	56	35	88	12	38	60	69	20
22	77	65	55	79	24	78	44	50	87

Эти числа и дадут нам номера тех предметов из 100, которые нужно считать отобранными наудачу. Отобранные таким образом предметы составят случайную выборку без повторения объема 20 из нашей генеральной совокупности объема 100.

Изучив эти 20 предметов, мы изучим свойства выборки, по которым можно судить с некоторым приближением о свойствах генеральной совокупности.

Задание

Составить случайную выборку из 100 бобин, используя таблицу случайных чисел, приведенную в приложении 7. Определить процент бракованных бобин в выборке, если партия бобин состоит из 1000 штук (аналогично примеру Сделать выводы.

Одностепенный механический метод

Этот метод основан на нумерации всех объектов генеральной совокупности и отборе их части через определенный интервал. Например, в выборку берут 5, 10, 15, 20-й и т.д. объекты. Этот метод практически неприменим при очень большом числе объектов в генеральной совокупности.

5.2. Двухстепенные методы отбора

Двухстепенные методы отбора применяют при разделении генеральной совокупности на отдельные, примерно равные части и фиксации этого разделения в выборке, а также при записи и обработке результатов испытаний.

Партии большинства материалов состоят из отдельных частей (например, партия пряжи – из коробок и отдельных бобин). Однако если партия материала разделена на части, а отобранные из разных ее частей объекты объединены в одну выборку, такая выборка называется одностепенной.

Механический двухстепенный метод отбора

Предусматривает деление генеральной совокупности на равные серии (группы) объектов. От каждой серии отбирают случайным методом по одному объекту или проводят по одному испытанию.

Серийный метод отбора

Такой метод отбора начинают с предварительного разделения генеральной совокупности на равные серии (группы) объектов; затем случайным методом отбирают несколько серий, которые испытывают полностью. Подобный отбор на практике используют очень редко, поскольку серии обычно достаточно большие, а полное испытание даже одной серии требует значительного времени.

Комбинированный метод отбора

Особенно часто используется при оценке качества нитей, когда каждая паковка является серией с большим числом возможных испытаний. В выборку случайно отбирают несколько серий (паковок), которые испытывают не полностью, а частично. При этом методе наименьшую ошибку выборки среднего получают при проведении одного испытания на паковку и при возможно большем числе паковок в выборке.

5.3. Трехстепенный метод отбора

Этот метод отбора предусматривает разделение генеральной совокупности на равные группы, которые, в свою очередь, подразделяются на серии, состоящие из одинакового числа объектов. Из нескольких групп отбирают по одинаковому числу серий, а из каждой серии испытывают по одинаковому числу объектов.

Например, при испытании нитей из партии отбирают несколько ящиков (групп), из каждого ящика берут одинаковое количество паковок (серий), а из каждой паковки испытывают одно и то же число отрезков нити.

6. ТАБЛИЦЫ И ГРАФИКИ РАСПРЕДЕЛЕНИЯ

Всякое статистическое исследование состоит из:

1) отбора (выборки) какого-то количества членов генеральной совокупности;
2) испытания их по какому-нибудь признаку;
3) записи результатов измерения.

В итоге получается выборочная совокупность в виде первоначальной таблицы вариантов.

Вторым этапом статистического исследования является упорядочение первоначальной таблицы. Достигают этого двумя способами:

1) путем составления *вариационного ряда*, т.е. переписи всех членов первоначальной таблицы в порядке возрастания (убывания) вариантов;

2) путем составления *таблицы распределения частот или частостей.*

Частотой m при дискретном изменении признака называется число одинаковых вариантов в выборочной совокупности, а при непрерывном изменении признака – число вариантов, попадающих на тот или иной из частных интервалов Δx, на которые разбивается общий интервал изменения признака в данной выборке.

Сумма всех частот равна объему выборки n:

$$\sum_{i=1}^{k} m_i = n \qquad (1)$$

где k – число частных интервалов.

Относительной частотой (частостью) w называется отношение частоты m к объему n.

$$w = \frac{m}{n} \qquad (2)$$

Сумма относительных частот равна единице или 100%:

$$\sum_{i=1}^{k} w_i = 1 \qquad (3)$$

Таблицей распределения частот при дискретном изменении признака называется таблица, состоящая из отличных один от другого вариантов, записанных в порядке возрастания или убывания, с указанием их частот (относительных частот); в случае непрерывного изменения признака – таблица, состоящая из частных интервалов Δx_i или их середин x_i с указанием частот (относительных частот) вариантов, приходящихся на каждый из этих интервалов.

Пример

В табл.2 приведены результаты измерения линейной плотности (текс) крученой нити.

Таблица 2

Результаты измерения линейной плотности (текс) крученой нити

112	104	109	113	110	108	113	111	102	99
127	110	102	109	105	95	136	129	110	126
111	123	120	101	111	115	109	112	120	119
112	108	115	98	112	107	111	112	101	120
118	104	116	114	96	121	116	114	108	95
109	115	98	111	113	103	110	111	117	97
101	99	112	111	108	113	114	112	121	95
112	105	112	125	111	116	106	115	103	90
105	105	98	105	123	121	104	101	111	98
120	112	120	109	139	109	96	113	112	98

Для составления таблицы распределения частот необходимо:

1) Выявить по первоначальной таблице (табл.2) наибольший x_{max} и наименьший x_{min} варианты и определить общий интервал изменения признака в данной выборке.

$$R = x_{\max} - x_{\min} \qquad (4)$$

$$R = 139 - 90 = 49.$$

$$x_{max}=139, \quad x_{min}=90.$$

2) Разделить общий интервал на k частных интервалов или классов Δx.

Число классов k зависит от объема выборки n и определяется по формулам:

$$k = 3{,}332 \lg n + 1 \qquad \text{при} \quad 50 < n < 200 \qquad (5)$$

$$k = 4\sqrt[5]{0{,}75(n-1)^2} \qquad \text{при} \quad n > 200 \qquad (6)$$

или по табл.3.

Таблица 3

Зависимость числа классов *k* от объема выборки *n*

Объем выборки, n	40-60	60-100	100-200	200-300	300-500	500
Число классов, k	5-7	7-10	10-14	14-15	15-18	18-25

Классовый интервал Δx, или разницу между нижними (верхними) пределами смежных классов, определяют по формуле:

$$\Delta x = \frac{R}{k}, \qquad (7)$$

где k – число классов.

Желательно, чтобы величина классового интервала была кратной 10, 5 или 2, так как тогда значительно облегчается разноска первичных данных в таблице распределения частот.

Для данных табл.2, при объеме выборки n=100 и числе классов k=10, определяют классовый интервал:

$$\Delta x = \frac{49}{10} = 4{,}9 \cong 5$$

3) Определить границы каждого классового интервала, начиная с наименьшего (или наибольшего) варианта. Полученные классовые интервалы и их середины внести в составляемую таблицу распределения частот.

Для данных табл.2 границы каждого классового интервала записывают в первую графу таблицы распределения (табл.4). В первой строке слева записывают наименьший результат из табл.2 x_{min}=90. Каждое число, записываемое слева в первой графе других строк, определяют последовательным прибавлением интервала Δx до получения наибольшего результата x_{max}=139 или числа меньшего его. В итоге находят все нижние границы классов. Верхние границы классов определяют, прибавляя к нижним границам интервал Δx и вычитая величину низшего разряда r первичных результатов. В рассматриваемом случае Δx=5 и r=1; поэтому верхние границы классов больше нижних на $\Delta x – r$=4. Верхние границы записывают в первой графе справа, отделяя их от нижних чертой. Далее определяют среднее значение для каждого классового интервала и записывают его во вторую графу таблицы распределения.

Таблица 4

Таблица распределения частот

<table>
<tr><th>Границы классов, Δx</th><th>Среднее значение класса, X</th><th>Отметки числа результатов в границах класса</th><th>Частота, m</th><th>Относительная частота, w</th><th>Плотность относительной частоты, y</th><th>Накопленная частота, $S(m)$</th></tr>
<tr><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td></tr>
<tr><td>90-94</td><td>92</td><td>|</td><td>1</td><td>0,01</td><td>0,002</td><td>1</td></tr>
<tr><td>95-99</td><td>97</td><td>|_|_|_| |_|_|_| |||</td><td>13</td><td>0,13</td><td>0,026</td><td>14</td></tr>
<tr><td>100-104</td><td>102</td><td>|_|_|_| |_|_|_| |</td><td>11</td><td>0,11</td><td>0,022</td><td>25</td></tr>
<tr><td>105-109</td><td>107</td><td>|_|_|_| |_|_|_| |_|_|_| ||</td><td>17</td><td>0,17</td><td>0,034</td><td>42</td></tr>
<tr><td>110-114</td><td>112</td><td>|_|_|_| |_|_|_| |_|_|_| |_|_|_|
|_|_|_| |_|_|_| ||</td><td>32</td><td>0,32</td><td>0,064</td><td>74</td></tr>
<tr><td>115-119</td><td>117</td><td>|_|_|_| |_|_|_|</td><td>10</td><td>0,10</td><td>0,02</td><td>84</td></tr>
<tr><td>120-124</td><td>122</td><td>|_|_|_| |_|_|_|</td><td>10</td><td>0,10</td><td>0,02</td><td>94</td></tr>
<tr><td>125-129</td><td>127</td><td>||||</td><td>4</td><td>0,04</td><td>0,008</td><td>98</td></tr>
<tr><td>130-134</td><td>132</td><td></td><td>0</td><td>0</td><td>0</td><td>98</td></tr>
<tr><td>135-139</td><td>137</td><td>||</td><td>2</td><td>0,02</td><td>0,004</td><td>100</td></tr>
<tr><td>Сумма</td><td></td><td></td><td>100</td><td>1</td><td></td><td></td></tr>
</table>

4) Определить принадлежность каждого значения первоначальной таблицы к тому или иному классовому интервалу.

Для упрощения подсчетов числа результатов в графе 3 табл.4 можно первые четыре результата отмечать вертикальными черточками, а пятый – поперек их горизонтальной черточкой. По черточкам подсчитать частоту *m*, приходящуюся на каждый класс. При необходимости определить относительную частоту *w* по формуле (2) для каждого классового промежутка.

Таблицы распределения частот (табл.4) представляют собой таблично-заданные функции, связывающие частоты *m* или относительные частоты *w* с частными интервалами Δx изменения признака X или с их серединами *x*. Указанные функции, характеризующие распределение признака X в выборочных совокупностях, называются *эмпирическими законами распределения признака*.

Эмпирические законы распределения признака могут быть изображены графически в виде *полигонов* или *гистограмм*.

Полигоном частот, или *относительных частот* называется ломаная линия, получаемая путем откладывания по оси абцисс значений признака X (например, границ частных интервалов Δx), построения над серединами частных интервалов соответствующих частот m или *относительных частот* w и соединения получаемых точек (рис.4).

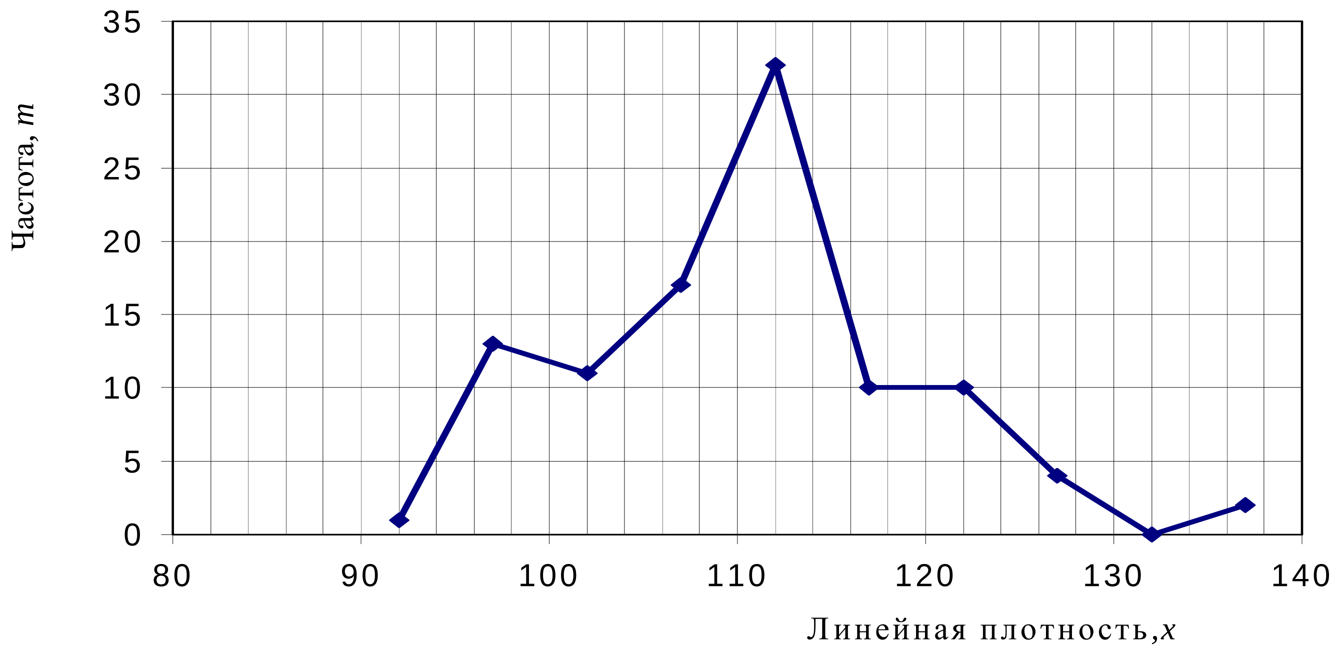


Рис.4. Полигон частот

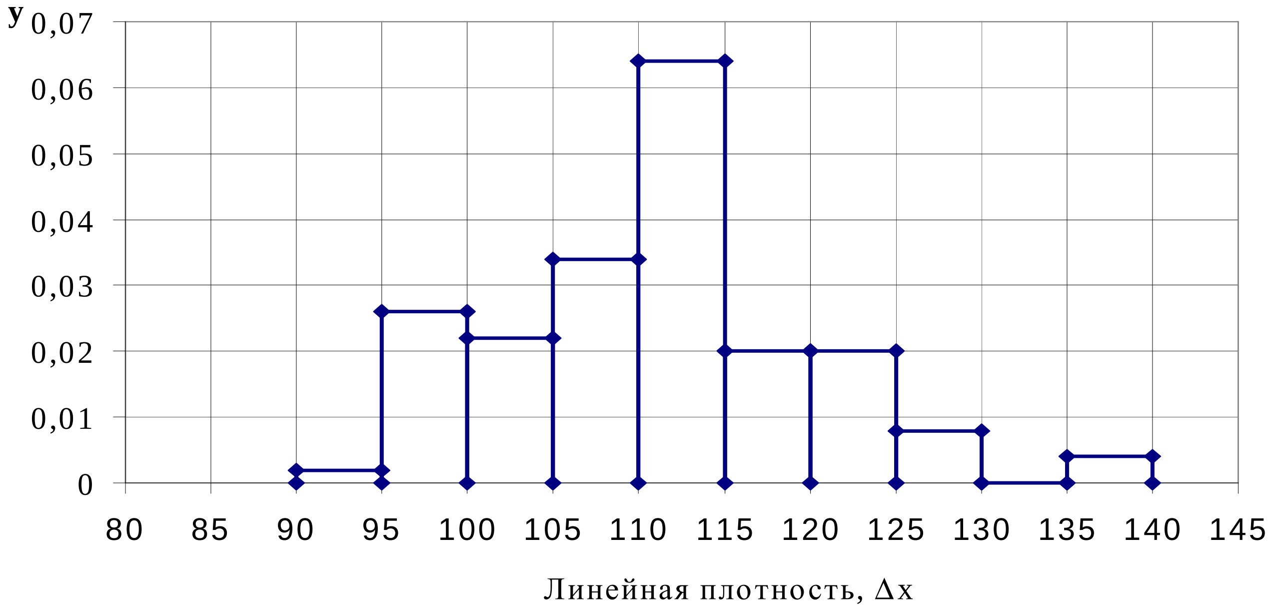
Гистограммой называется ступенчатый график, состоящий из прямоугольников, основаниями которых служат частные интервалы Δx_i, а площади равны частотам m_i или относительным частотам w_i.

Из того, что $m_i=y_i\Delta x_i$ или $w_i=y_i\Delta x_i$,

$$y_i = \frac{m_i}{\Delta x_i} \quad \text{или} \quad y_i = \frac{w_i}{\Delta x_i} \qquad (8)$$

где y_i – высоты прямоугольников, следует что

Такие выражения называются *плотностями частоты или относительной частоты*. На рис.5 изображена гистограмма, соответствующая таблице



распределения (табл.4), в таблице в столбце 5 подсчитаны плотности относительных частот y_i по формуле(8).

Рис.5 Гистограмма относительных частот

Свойства гистограмм относительных частот следующие:

1) Площадь части гистограммы относительных частот, соответствующая интервалу от $x=a$ до $x=b$ равна относительной частоте вариантов на этом интервале:

$$S_{ab}=w\ (a<X<b)$$

2)Площадь всей диаграммы относительных частот равна единице: $S_{общ}=1$.

В некоторых случаях применяют гистограммы, на которых по оси абцисс откладывают значения частоты (относительной частоты). На рис.6 приведена такая гистограмма частот, построенная по данным табл.4.

Кроме законов распределения в виде таблиц распределения частот m (относительных частот w), существует также суммарный закон распределения *накопленных* частот $S(m)$.

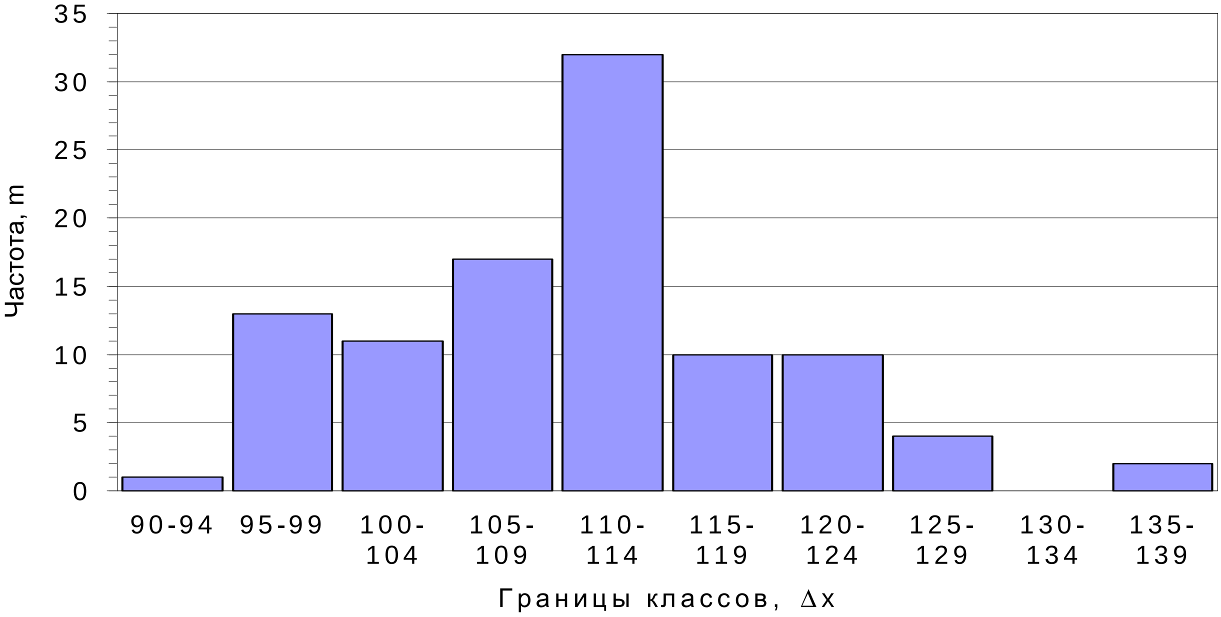


Рис.6. Гистограмма частот

Накопленной частотой (*или кумулятивной частотой)* для варианта x_p называется сумма частот $S(m_p)$.вариантов, не превышающих x_p, т.е.

$$S(m_p) = m_1 + m_2 + \ldots + m_p \qquad (9)$$

График суммарного распределения накопленных частот называется *кумулятой*. На рис.7 изображена кумулята, построеная по данным табл.4, в таблице в столбце 7 подсчитаны накопленные частоты $S(m)$по формуле (9).

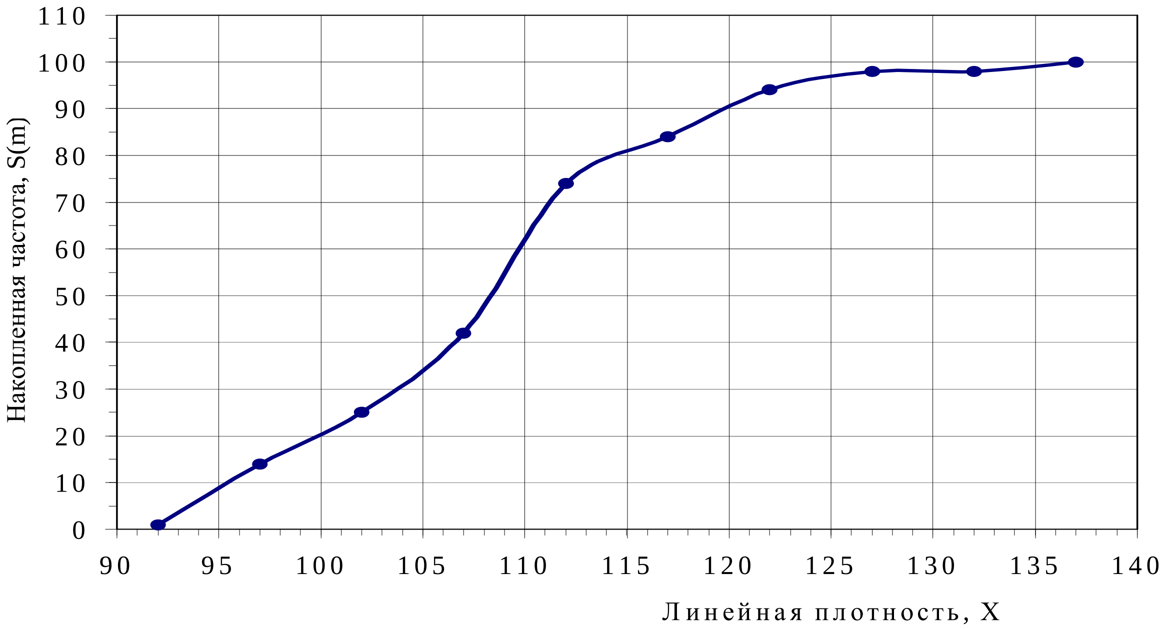


Рис.7. Кумулята

7. СВОДНЫЕ ВЫБОРОЧНЫЕ ХАРАКТЕРИСТИКИ

Для уплотнения первичной информации рекомендуется применять сводные характеристики, заменяющие совокупность первичных результатов отдельных измерений.

7.1.Средняя величина признака в выборке

Статистическую совокупность можно охарактеризовать различными средними величинами. *Среднее значение* определяет центр распределения случайных величин, около которого группируется большая их часть. Этот центр распределения характеризуется средней арифметической, средней квадратической, средней гармонической и средней геометрической.

Среднюю арифметическую определяют, складывая все результаты и деля

$$\overline{X} = \frac{X_1 + X_2 + ... + X_n}{n} = \frac{\sum_{i=1}^{n} X_i}{n}, \qquad (10)$$

сумму на число испытаний:

где n – численность выборочной совокупности.

Если среднюю арифметическую *X* используют для расчета других характеристик, то ее записывают на один разряд ниже низшего разряда в отдельных измерениях. Если *X* характеризует итоговый результат испытаний, ее округляют до низшего разряда, одинакового с первичными данными.

Средней квадратической называют квадратный корень из частного от деления суммы квадратов отдельных значений X_i совокупности на их число.

$$X_k = \sqrt{\frac{\sum_{i=1}^{n} X_i^2}{n}} \qquad (11)$$

Выборочная *средняя гармоническая* определяется по формуле:

$$X_h = \frac{n}{\sum_{i=1}^{n} \frac{1}{X_i}}. \qquad (12)$$

Среднюю гармоническую применяют при вычислении среднего номера и при определении зависимости между номером и линейной плотностью (текс).

Выборочная *средняя геометрическая* определяется по формуле:

$$X_г = \sqrt[n]{X_1 \cdot X_2 \cdot \ldots \cdot X_n}. \qquad (13)$$

Средняя геометрическая применяется в комплексных оценках.

Между рассмотренными средними существует следующее соотношение:

$$X_h \leq X_г \leq X \leq X_k.$$

В практике обработки результатов эксперимента при исследовании свойств продуктов или параметров технологических процессов текстильной промышленности чаще всего используется средняя арифметическая.

В практике текстильной промышленности применяют еще один вид средней величины – *моду* (напрмер, модальную длину волокна).

В случае дискретного изменения признака модой $x_{мод}$ называется наиболее часто встречающийся вариант. При непрерывном изменении признака *мода* – это значение признака с наибольшей плотностью.

Если статистическая совокупность дана в виде таблицы распределения частот (или относительных частот), то приближенно за моду $x_{мод}$ можно принять середину интервала с наибольшей частотой (относительной частотой), т.е. абсциссу середины наиболее высокой ступеньки гистограммы. Более точно *мода* $x_{мод}$ – это абсцисса максимума кривой, сглаживающей гистограмму в области ее вершины.

7.2. Мера рассеяния признака в выборке

Существует несколько видов меры рассеяния.

Размах варьирования определяют как разность между наибольшим показателем x_{max} и наименьшим x_{min} по формуле (4).

Неравномерность материала оценивают по среднему размаху варьирования из m выборок:

$$\overline{R}_{пв} = \frac{1}{m}\sum_{i=1}^{m} R_{ni}, \qquad (14)$$

где $R_{n1}, R_{n2,\ldots}, R_{nm}$– значение размахов для m выборок по n испытаний в каждой.

При небольшом числе измерений (n≤10) размахом пользуются, когда нужно быстро и не очень точно определить рассеяние отдельных значений случайных величин. Чем больше число измерений, тем менее надежным критерием рассеяния становится размах.

Рассеяние значений статистической величины может быть также охарактеризовано *средним абсолютным отклонением*:

$$\theta = \frac{\sum_{i=1}^{n} \left| X_i - \overline{X} \right|}{n}. \tag{15}$$

Если статистическая совокупность задана таблицей распределения частот m_i (относительных частот w_i), то среднее абсолютное отклонение рассчитывается:

$$\theta = \frac{\sum_{i=1}^{k} m_i \left| X_i - \overline{X} \right|}{n} = \sum_{i=1}^{r} w_i \left| X_i - \overline{X} \right|. \tag{16}$$

Величина θ обладает существенным недостатком: она плохо выявляет рассеяние при небольшом количестве больших отклонений X_i от X на фоне большого количества малых отклонений.

Отношение среднего абсолютного отклонение к средней арифметической, выраженное в процентах, называется *коэффициентом неровноты*:

$$H = \frac{\theta}{\overline{X}} \cdot 100, \%. \tag{17}$$

Сумма всех положительных отклонений от средней по абсолютной величине равна сумме всех отрицательных отклонений. При определении общей суммы абсолютных отклонений достаточно подсчитать только суммы одних положительных или отрицательных отклонений, а затем ее удвоить. На основании отмеченного свойства коэффициент неровноты можно более просто определить по формуле проф. В.П.Добычина

$$H = \left(\frac{n_1}{n} - \frac{S_1}{S} \right) \cdot 200 \tag{18}$$

или по формуле Зоммера

где n_1 – число результатов, меньших общей средней X;

n – общее число испытаний;

$$H = \frac{(\bar{X} - \bar{X}_1) \cdot n_1}{n \cdot \bar{X}} \cdot 200, \qquad (19)$$

S_1 – сумма результатов, меньших X;

S – сумма всех n результатов;

X_1 – средняя из вариантов, меньших X.

Численное значение коэффициента неровноты, определенное по формулам (17), (18), (19) одно и то же. Наиболее простой из них является формула проф. В.П.Добычина. Коэффициент неровноты обладает тем же недостатком, что и среднее абсолютное отклонение.

Абсолютной характеристикой рассеяния случайной величины X_i около центра распределения X является *дисперсия* $S^2\{X\}$. При малом объеме выборки (n≤30) дисперсия определяется по формуле:

$$S^2\{X\} = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2. \qquad (20)$$

Выражается дисперсия $S^2\{X\}$ в квадратных единицах признака X.

Величина, равная положительному значению корня квадратного из дисперсии, называется *средним квадратическим отклонением* величины X:

$$S\{X\} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2}. \qquad (21)$$

Среднее квадратическое отклонение, выраженное в процентах к средней арифметической, называется *коэффициентом вариации*:

$$C\{X\} = \frac{S\{X\}}{\bar{X}} \cdot 100, \%. \qquad (22)$$

Коэффициент вариации является относительной характеристикой рассеяния случайной величины.

Величина среднего квадратического отклонения $S\{X\}$, а следовательно, и величины $S^2\{X\}$ и $C\{X\}$, с ней связанные лучше выявляют большие отклонения на фоне малых, так как все отклонения возводятся в квадрат и тем самым влияние больших отклонений усиливается.

7.3. Меры асимметрии и крутовершинности кривой распределения

Эмпирические кривые распределения (рис.3) не всегда следуют нормальному закону распределения. Для многих распределений характерен сдвиг кривой влево или вправо.

Асимметрия характеризует степень несимметричности распределения относительно его среднего. Для вычисления показателя асимметрии использу-

$$A = \frac{\mu_3}{S^3\{X\}} = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^3}{n \cdot S^3\{X\}}, \qquad (23)$$

ют сумму кубов отклонений от средней:

где μ_3 – эмпирический (выборочный) момент третьего порядка.

Чем больше абсолютная величина показателя асимметрии, тем больше скошена кривая, тем больше ее асимметричность. Если величина показателя асимметрии положительна (A>0), а значит преобладает влияние положительных отклонений, то кривая распределения более полога справа ($x_{мод}$<X). При отрицательной асимметрии (A<0) кривая распределения более полога слева ($x_{мод}$ >X).

Эксцесс характеризует “крутость”, т.е. больший или меньший подъем кривой распределения по сравнению с нормальной. Для вычисления эксцесса используют сумму отклонений от средней в четвертой степени:

$$E = \frac{\mu_4}{S^4\{X\}} - 3 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^4}{n \cdot S^4\{X\}} - 3, \qquad (24)$$

где μ_4 – эмпирический (выборочный) момент четвертого порядка.

Крутовершинность оценивается по сравнению с нормальной кривой распределения, для которой $\frac{\mu_4}{S^4\{X\}} = 3$ и, следовательно, E=0. Если эксцесс положительный (E>0), то кривая имеет более высокую и “острую” вершину. Распределение с отрицательным эксцессом (E<0) имеют более “плосковершинные” и низкие кривые распределения.

7.4. Способ произведений для приближенного расчета характеристик

В тех случаях, когда выборка имеет большой объем (n>30) для упрощенного расчета сводных выборочных характеристик применяют *способ произведений (моментов)* и *способ сумм.*

Способ произведений (моментов) назван так потому, что в вычислительные формулы входят моменты разных степеней, а также при ычислениииях преимущественно пользуются произведением различных чисел.

Рассмотрим на примере обработку методом произведений приведенных в табл.2 результатов измерения линейной плотности (текс) крученой нити. Для этого все результаты распределяют по k классам (интервалам) аналогично примеру, рассмотренному в предыдущем разделе (см. табл.4).

Для удобства расчетов все результаты сводят в табл.5, в которой графы 1, 2, 3 рассчитываются аналогично табл.4.

Таблица 5

Расчет выборочных характеристик способом произведений

Границы классов, Δx	Среднее значение класса, X	Частота, m	Условное отклонение, α	$m\alpha$	$m\alpha^2$	$m\alpha^3$	$m\alpha^4$
1	2	3	4	5	6	7	8
90-94	92	1	–4	–4	16	–64	256
95-99	97	13	–3	–39	117	–351	1053
100-104	102	11	–2	–22	44	–88	176
105-109	107	17	–1	–17	17	–17	17
110-114	112	32	0	0	0	0	0
115-119	117	10	+1	10	10	10	10
120-124	122	10	+2	20	40	80	160
125-129	127	4	+3	12	36	108	324
130-134	132	0	+4	0	0	0	0
135-139	137	2	+5	10	50	250	1250
Сумма		100		–30	330	–72	3246
Условные моменты				–1,5	82,5	–90	20287,5

Условный центр x_0 намечают против одной из наибольших частот. Центру соответствует условное отклонение α=0, которое фиксируется в графе 4. В этой графе записывают постепенно изменяющиеся на единицу отклонения α: вверх от нуля – отрицательные, вниз – положительные. После этого для всех интервалов вычисляют $m\alpha$, $m\alpha^2$, $m\alpha^3$, $m\alpha^4$, которые записывают соответственно в графах 5, 6, 7, 8. В этих графах подсчитывают соответствующие суммы $\Sigma m\alpha$, $\Sigma m\alpha^2$, $\Sigma m\alpha^3$, $\Sigma m\alpha^4$.

Далее вычисляют условные моменты по формулам:

$$m_1 = \frac{\sum_{i=1}^{k} m_i \cdot \alpha_i}{n} \Delta x \qquad (25)$$

$$m_2 = \frac{\sum_{i=1}^{k} m_i \cdot \alpha_i^2}{n} \Delta \mathrm{x}^2 \qquad (26)$$

$$m_3 = \frac{\sum_{i=1}^{k} m_i \cdot \alpha_i^3}{n} \Delta \mathrm{x}^3 \qquad (27)$$

$$m_4 = \frac{\sum_{i=1}^{k} m_i \cdot \alpha_i^4}{n} \Delta \mathrm{x}^4, \qquad (28)$$

где Δx– классовый интервал.

Условные моменты записывают в соответствующих графах табл.5.

Определив значение x_0 как среднее границ класса с α=0, вычисляют выборочное среднее X по формуле:

$$\overline{X} = x_0 + \frac{\Delta x \cdot \sum_{i=1}^{k} m_i \cdot \alpha_i}{n} = x_0 + m_1, \qquad (29)$$

Для данных табл.5 получаем:

$$x_0 = \frac{110 + 114}{2} = 112$$

$$\overline{X} = 112 - 1{,}5 = 110{,}5.$$

Дисперсия рассчитывается по формуле:

$$S^2\{X\} = (m_2 - m_1^2), \qquad (30)$$

где m_1, m_2– соответствующие условные моменты.

Среднее квадратическое отклонение и коэффициент вариации определяется по формулам (21) и (22).

Для данных табл.5 получаем:

$$S^2\{X\} = \left(82{,}5 - (-1{,}5)^2\right) = 80{,}25$$
$$S\{X\} = \sqrt{80{,}25} \approx 8{,}96$$
$$C\{X\} = \frac{8{,}96 \cdot 100}{110{,}5} = 8{,}1\%$$

Показателей асимметрии и эксцесса рассчитываются по следующим формулам:

$$A = \frac{\left(m_3 - 3m_1m_2 + 2m_1^3\right)}{S^3\{X\}} \qquad (31)$$

$$\text{E} = \frac{\left(m_4 - 4m_3m_1 + 6m_2m_1^2 - 3m_1^4\right)}{S^4\{X\}} - 3. \qquad (32)$$

Для распределения, приведенного в табл.5 , показатели асимметрии и эксцесса имеют следующие значения:

$$A = \frac{-90 - 3\cdot(-1{,}5)\cdot 82{,}5 + 2\cdot(-1{,}5)^3}{8{,}96^3} = 0{,}38$$

$$E = \frac{20287{,}5 - 4\cdot(-90)\cdot(-1{,}5) + 6\cdot 82{,}5\cdot(-1{,}5)^2 - 3\cdot(-1{,}5)^4}{8{,}96^4} - 3 = 0{,}23.$$

Таким образом, распределение имеет положительную асимметрию и положительный эксцесс, т.е. кривая распределения более полога справа и имеет более высокую и «острую» вершину по сравнению с нормальным распределением. Это подтверждает график кривой распределения, который приведен для данной таблицы распределения на рис.3.

7.5. Способ сумм для приближенного расчета характеристик

Сводные выборочные характеристики можно приближенно определить по *способу сумм* путем кратного сложения частот по схеме, приведенной в табл.6. В качестве условного среднего выбирают обычно интервал или класс, имеющий наибольшую частоту и расположенный в средней части колонки классов. Против класса с наибольшей частотой в графе 4 ставят прочерк (см. табл.6). В графе 4 проставляют накопленные частоты сверху вниз и снизу вверх до прочерка. Далее подсчитывают суммы нижней части a_1 и верхней части b_1 графы 4:

a_1 – сумма чисел графы 4 снизу до прочерка;

b_1 – сумма чисел графы 4 сверху до прочерка.

Таблица 6

Расчет выборочных характеристик способом сумм

Границы классов, Δx	Среднее значение класса, X	Частота, m	b_1=82	b_2=56	b_3=17	b_4=1
1	2	3	4	5	6	7
90-94	92	1	1	1	1	1
95-99	97	13	14	15	16	–
100-104	102	11	25	<u>40</u>	–	–
105-109	107	17	<u>42</u>	–	–	–
110-114	112	32	–	–	–	–
115-119	117	10	<u>26</u>	–	–	–
120-124	122	10	16	<u>26</u>	–	–
125-129	127	4	6	10	16	–
130-134	132	0	2	4	6	8
135-139	137	2	2	2	2	2
Сумма		100	a_1=52	a_2=42	a_3=24	a_4=10

В графе 5 ставят три прочерка: средний из них – против класса с наибольшей частотой, один строчкой выше и один строчкой ниже, далее подсчитывают накопленные частоты сверху и снизу до прочерков. В графе 6 ставят пять прочерков, в графе 7 – соответственно семь прочерков, аналогично подсчитывая накопленные частоты, определяют суммы нижней части a_2, a_3, a_4 и верхней части b_2, b_3, b_4 граф:

a_2 – сумма чисел графы 5 снизу до прочерка;

b_2 – сумма чисел графы 5 сверху до прочерка.

a_3 – сумма чисел графы 6 снизу до прочерка;

b_3 – сумма чисел графы 6 сверху до прочерка.

a_4 – сумма чисел графы 7 снизу до прочерка;

b_4 – сумма чисел графы 7 сверху до прочерка.

Суммы a_1 и b_1 контролируют крайними значениями расположенными около черточек (эти цифры в табл.6 подчеркнуты):

$$a_1 = 26 + 26 = 52,$$
$$b_1 = 40 + 42 = 82.$$

$$m_1 = (a_1 - b_1)\frac{\Delta x}{n} \qquad (33)$$

$$m_2 = [a_1 + b_1 + 2(a_2 + b_2)]\frac{\Delta x^2}{n} \qquad (34)$$

$$m_3 = [(a_1 - b_1) + 6(a_2 - b_2) + 6(a_3 - b_3)]\frac{\Delta x^3}{n} \qquad (35)$$

Условные моменты определяют по формулам:

Для приведенного примера (табл.6) условные моменты, рассчитанные способом сумм, полностью совпадают с условными моментами, определенными способом произведений:

Далее аналогично рассчитываются сводные выборочные характеристики

$$m_4 = [(a_1 + b_1) + 14(a_2 + b_2) + 36(a_3 + b_3) + 24(a_4 + b_4)]\frac{\Delta x^4}{n}. \qquad (36)$$

$$m_3 = [52 - 82 + 6 \cdot (42 - 56) + 6 \cdot (24 - 17)] \cdot \frac{5^3}{100} = -90$$

$$m_4 = [52 + 82 + 14 \cdot (42 + 56) + 36 \cdot (24 + 17) + 24 \cdot (10 + 1)] \cdot \frac{5^4}{100} = 20287{,}5$$

по формулам (29)–(32).

$$m_1 = (52 - 82) \cdot \frac{5}{100} = -1{,}5$$

$$m_2 = [52 + 82 + 2 \cdot (42 + 56)] \cdot \frac{5^2}{100} = 82{,}5$$

8. АКТИВНЫЙ ЭКСПЕРИМЕНТ. ЛИНЕЙНАЯ ОДНОФАКТОРНАЯ РЕГРЕССИОННАЯ МОДЕЛЬ

Планирование эксперимента – это постановка опытов по некоторой заранее составленной схеме, обладающей какими-то оптимальными свойствами. В задачу планирования эксперимента входит: выбор необходимых для эксперимента опытов, т.е. построение матрицы планирования, и выбор методов математической обработки результатов эксперимента.Матрица планирования эксперимента представляет собой таблицу, в которой указаны значения уровней факторов в различных сериях опытов. Число опытов определяется задачами исследования и методами планирования эксперимента.

При определении регрессионной модели для объекта с одним выходным параметром Y проводят активный эксперимент в широком диапазоне изменения фактора X. Обычно применяют число уровней фактора, т.е. число опытов в матрице планирования эксперимента, N=5…6. Для повышения точности определения выходного параметра Y каждый опыт матрицы повторяется несколько раз (m>2).

Рассмотрим операции, которые совершает исследователь при обработки данных однофакторного эксперимента, на примере, в котором изучалось влияние X – натяжения нити (сН) на Y – силу трения нити в глазке ремизки (сН) ткацкого станка. В таблице 7 приведены значения выходного параметра Y_{uv} в v–м опыте каждого u–го опыта матрицы, когда число опытов N=6 и число повторов каждого опыта m=5.

Таблица 7

Матрица планирования эксперимента

X_u	u	v				
		Y_{uv}				
		1	2	3	4	5
10	1	2,1	2,5	2,2	1,9	1,6
20	2	4,2	4,8	4,4	3,7	4,3
30	3	6,2	6,6	6,3	6,1	5,7
40	4	8	8,2	8,4	7,7	8,5
50	5	10,5	9,8	10,7	10,2	9,5
60	6	12,7	12,6	12,7	12,2	12,1

8.1. Исключение резко выделяющихся данных

Рассмотрим эту операцию при анализе первого опыта матрицы u=1, когда X=10, Y_{uvmax}=2,5, Y_{uvmin}=1,6. Рассчитанные по формулам (10) и (20) значения среднего арифметического $\overline{Y}_u$ и дисперсии $S^2_u\{Y\}$ приведены в таблице 8.

Для исключения резко выделяющихся данных необходимо определить расчетные значения критерия Смирнова-Грабса по формулам:

при подозрении резко выделяющегося максимального значения $Y_{i\,max}$:

$$V_{R\max} = \frac{(Y_{i\max} - \overline{Y})}{S\{Y\}} \sqrt{\frac{m}{m-1}} \tag{37}$$

при подозрении резко выделяющегося минимального значения $Y_{i\,min}$:

$$V_{R\min} = \frac{(\overline{Y} - Y_{i\min})}{S\{Y\}} \sqrt{\frac{m}{m-1}} \tag{38}$$

где $\overline{Y}$ – среднее значение выходного параметра для u–го опыта матрицы;

$S_u\{Y\}$ – среднее квадратическое отклонение для u–го опыта матрицы.

Таблица 8

Расчет основных статистических характеристик

X_u	v									
	u	Y_{uv}					$\sum_{v=1}^{m} Y_{uv}$	$\overline{Y}_u$	$S^2_u\{Y\}$	W_R
		1	2	3	4	5				
10	1	2,1	2,5	2,2	1,9	1,6	10,3	2,06	0,0904	4,97
20	2	4,2	4,8	4,4	3,7	4,3	21,4	4,28	0,1256	4,84
30	3	6,2	6,6	6,3	6,1	5,7	30,9	6,18	0,0856	4,88
40	4	8	8,2	8,4	7,7	8,5	40,8	8,16	0,0824	4,79
50	5	10,5	9,8	10,7	10,2	9,5	50,7	10,14	0,1944	4,80
60	6	12,7	12,6	12,7	12,2	12,1	62,3	12,46	0,0664	4,06

Используя формулы (37) и (38) определяем для первого опыта матрицы:

$$V_{R\max 1} = \frac{2{,}5 - 2{,}06}{0{,}3} \sqrt{\frac{5}{5-1}} = 1{,}636$$

$$V_{R\min 1} = \frac{2{,}06 - 1{,}6}{0{,}3} \sqrt{\frac{5}{5-1}} = 1{,}711$$

По приложению 1 находим, что $V_T[p_D=0{,}95;\ m=5]=1{,}869$. Так как $V_{Rmax}<V_T$ и $V_{Rmin}<V_T$, то рассмотренные значения $Y_{uvmax}=2{,}5$ и $Y_{uvmin}=1{,}6$ не являются резко выделяющимися и остаются для дальнейшей статистической обработки.

Аналогично рассчитываются расчетные значения критерия Смирнова-Грабса V_{Rmax} и V_{Rmin} для других опытов матрицы, затем они сравниваются с табличным значением критерия V_T. Если расчетные значения критерия Смирнова-Грабса превышают табличное, то соответствующие значения Y_{uvmax} или Y_{uvmin} исключаются из дальнейшего расчета, а среднее значение и дисперсия пересчитываются для соответствующего опыта матрицы.

8.2. Проверка гипотезы о нормальном распределении случайных величин Y_{uv}

Проверка этой гипотезы для каждого *u*–го опыта матрицы состоит в определении расчетного значения критерия W_R по формуле:

$$W_R = \frac{Q^2}{S_u^2\{Y\}}, \tag{39}$$

где

$$Q = q_m(Y_m - Y_1) + \ldots + q_{m-k+1}(Y_{m-k+1} - Y_k), \tag{40}$$

$$Y_m \le Y_{m-1} \le \ldots \le Y_2 \le Y_1$$

$$k = \frac{m}{2} \text{ - при четном числе } m;$$

$$k = \frac{m\text{-}1}{2} \text{ - при нечетном числе } m.$$

Значения q_{m-i+1} для i=1…k и m =3…50 приведены в приложении 10 [1].

Для 1-го опыта матрицы при *u*=1 и *X*=10 располагаем значения Y_{uv} по возрастанию: 2,5 ≥2,2≥2,1 ≥1,9 ≥1,6.

$$k = \frac{5-1}{2} = 2$$

Используя приложение 10, находим q_1=0,6646 и q_2=0,2413 и вычисляем значение Q_1 и W_{R1}:

$$Q_1 = 0,6646(2,5-1,6)+0,2413(2,2-1,9)=0,671$$

$$W_{R1} = \frac{0,671^2}{0,0904} = 4,97$$

Расчетное значение W_{R1} сравнивают с табличным W_T, которое определяется по приложению 6. W_T определяется для заданной доверительной вероятности p_D и известного числа повторных опытов m. Для рассматриваемого примера $W_T[p_D=0,95;\ m=5]=0,762$.

Так как расчетное значение W_{R1} превышает табличное значение W_T для выбранной доверительной вероятности, то гипотеза о нормальном распределении случайных величин не отвергается.

В табл. 8 приведены значения W_R и для других опытов матрицы. Эти значения также превышают табличное, и поэтому первое условие о возможности применения регрессионного анализа удовлетворяется.

8.3. Проверка гипотезы об однородности дисперсий в опытах матрицы

Так как число повторных опытов (m=5) одинаково для всех опытов матрицы, то для проверки однородности дисперсий применяется критерий Кочрена, расчетное значение которого равно:

$$G_R = \frac{S^2_{u\max}\{Y\}}{\sum\limits_{u=1}^{N} S_u^2\{Y\}} \qquad (41)$$

$$G_R = \frac{0,1944}{0,6448} = 0,301$$

Расчетное значение G_R сравнивается с табличным значением G_T, которое определяют по приложению 4 в зависимости от числа опытов в матрице N и числа степеней свободы дисперсии $f\{S^2_u\}=m-1$ для заданной доверительной вероятности p_D. В рассматриваемом примере $G_T[p_D=0,95;\ N=6;\ f=5-1=4]=0,4803$.

Так как G_R < $G_Т$, то гипотеза об однородности дисперсий, т.е. равноточности и воспроизводимости опытов, не отвергается.

Если G_R > $G_Т$, то дисперсии в N рядах измерений неоднородны. После отбрасывания $S^2_{u\ max}\{Y\}$ описанную выше процедуру следует повторить для N-1 рядов измерений.

Если число повторных опытов неодинаково при различных уровнях факторов, то для проверки однородности дисперсий используется критерий Бартлера, расчетное значение которого равно:

$$B_R = \frac{2{,}303}{C}\left[f \lg S^2_{(1)}\{Y\} - \sum_{u=1}^{N} f_u \lg S^2_u\{Y\} \right], \qquad (42)$$

где
$$C = 1 + \frac{1}{3(N-1)}\left(\sum_{u=1}^{N} \frac{1}{f_u} - \frac{1}{f} \right) \qquad (43)$$

$S^2_{(1)}$ - средняя дисперсия выходного параметра в опытах матрицы, определяемая по формуле:

$$S^2_{(1)}\{Y\} = \frac{1}{f}\sum_{u=1}^{N} f_u S^2_u\{Y\} \qquad (44)$$

Число степеней свободы этой дисперсии равно:

$$f = \sum_{u=1}^{N} f_u \qquad (45)$$

$$f_u = m_u - 1$$

Если f_u>2, следовательно, величина B_R распределена как χ^2-критерий с числом степеней свободы N-1, который определяют по приложению 2. Если B_R < $B_Т = \chi^2[p_D;\ f=N-1]$, то это свидетельствует об отсутствии значимого различия между дисперсиями $S^2_u\{Y\}$, т.е. об их однородности.

8.4. Определение средней дисперсии выходного параметра в опытах матрицы

Если в опытах матрицы дисперсии однородны и число повторных опытов одинаково, то средняя дисперсия определяется по формуле:

$$S^2_{(1)} = \frac{1}{N}\sum_{u=1}^{N} S^2_u\{Y\} \qquad (46)$$

Число степеней свободы этой дисперсии равно:

$$f\left\{S^2_{(1)}\right\} = N(m-1) \qquad (47)$$

Средняя дисперсия характеризует средний разброс значений выходного параметра относительно его средних значений при каждом уровне факторов, т.е. ошибку опытов в эксперименте. В рассматриваемом примере эта дисперсия, или, как ее называют дисперсия воспроизводимости, равна:

$$S^2_{(1)} = \frac{0{,}6448}{6} = 0{,}107$$

$$f\left\{S^2_{(1)}\right\} = 6(5-1) = 24.$$

8.5. Определение подходящего вида регрессионной модели

Для определения подходящего вида регрессионной модели используют следующую информацию:

1) графическую взаимосвязь $\overline{Y} = f(X)$ между средними значениями выходного параметра для каждого уровня факторов и значением фактора по данным эксперимента. При сопоставлении этого графика с графиками известных функций устанавливают вид уравнения;
2) характер изменения разделенных и неразделенных разностей первого порядка, определяемых по данным эксперимента.

Если в результате эксперимента получены следующие пары значений $X_1\overline{Y_1},\ldots,X_u\overline{Y_u},\ldots X_N\overline{Y_N}$, то разделенными разностями первого порядка называются величины:

$$\Delta_{R1} = \frac{\overline{Y}2 - \overline{Y}1}{X_2 - X_1},\ \ldots,\ \Delta_{Ru} = \frac{\overline{Y}u+1 - \overline{Y}u}{X_{u+1} - Xu},\ \ldots,\ \Delta_{R(N-1)} = \frac{\overline{Y}_N - \overline{Y}_{N-1}}{X_N - X_{N-1}} \qquad (48)$$

и неразделенными разностями первого порядка – величины:

$$\Delta_{H1} = \overline{Y}2 - \overline{Y}1, \ldots, \Delta_{Hu} = \overline{Y}u+1 - \overline{Y}u, \ldots, \Delta_{H(N-1)} = \overline{Y}_N - \overline{Y}_{N-1} \qquad (49)$$

Неразделенные разности первого порядка используют, когда интервал варьирования факторов постоянный, т.е. $I_X=X_2-X_1=X_{u+1}-X_u=X_N-X_{N-1}=const$.

В рассматриваемом примере графическая взаимосвязь $\overline{Y}=f(X)$ между средними значениями выходного параметра $\overline{Y}$ для каждого уровня факторов и значением фактора X приведена на рис.8. При сопоставлении этого графика с графиками известных функций можно сделать вывод, что для описания экспериментальных данных наиболее подходит линейная модель.

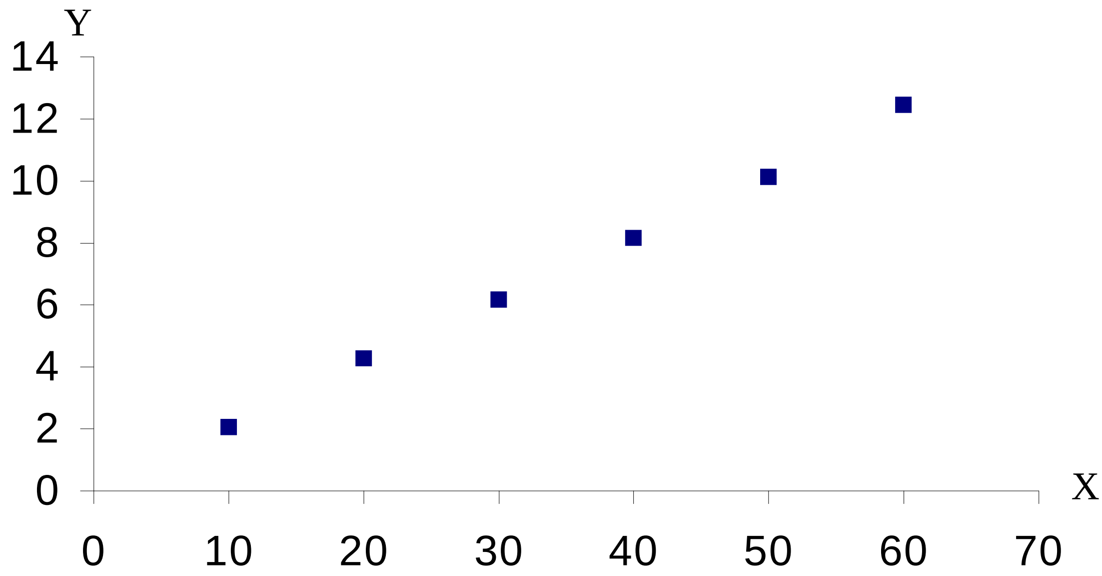


Рис.8 Зависимость силы трения нити в глазке ремизки от ее натяжения

В рассматриваемом примере интервал варьирования факторов постоянный и равен I_X=20-10=30-20=40-30=50-40=60-50=10. Поэтому определяем неразделенные разности первого порядка по формуле (49):

$$\Delta_{H1} = 4{,}28 - 2{,}06 = 2{,}22$$
$$\Delta_{H2} = 6{,}18 - 4{,}28 = 1{,}90$$
$$\Delta_{H3} = 8{,}16 - 6{,}18 = 1{,}98$$
$$\Delta_{H4} = 10{,}14 - 8{,}16 = 1{,}98$$
$$\Delta_{H5} = 12{,}46 - 10{,}14 = 2{,}32$$

Ввиду малого различия неразделенных разностей первого порядка $\Delta_{H\max} - \Delta_{H\min} = 2{,}32 - 1{,}9 = 0{,}42$, не превышающего удвоенной величины среднеквадратической ошибки эксперимента $2S_{(1)}\{Y\}=0{,}656$, можно считать что они тождественны и поэтому для описания экспериментальных данных можно принять уравнение прямой линии:

$$Y_R = a_0 + a_1 X \tag{50}$$

или
$$Y_R = d_0 + d_1(X - \overline{X}), \tag{51}$$

где
$$\overline{X} = \frac{1}{N}\sum_{u=1}^{N} X_u \tag{52}$$

Использование уравнения (51) позволяет упростить статистические расчеты при обработке экспериментальных данных, так как коэффициенты регрессии d_0 и d_1 не коррелированны.

8.6. Определение коэффициентов регрессии

Если дисперсии выходного параметра для каждого уровня фактора однородны, то для определения коэффициентов регрессии в уравнении (51) можно применять метод наименьших квадратов. Используя условие $\sum_{u=1}^{N}(Y_u - Y_{Ru})^2 = \min$, устанавливают следующие нормальные уравнения:

$$\left.\begin{aligned} & d_0 N + d_1 \sum_{u=1}^{N}(X_u - \overline{X}) = \sum_{u=1}^{N}\overline{Y}_u \\ & d_0 \sum_{u=1}^{N}(X_u - \overline{X}) + d_1 \sum_{u=1}^{N}(X_u - \overline{X})^2 = \sum_{u=1}^{N}(X_u - \overline{X})\overline{Y}_u \end{aligned}\right\} \tag{53}$$

Так как $\sum_{u=1}^{N}(X_u - \overline{X}) = 0$, то решая эти уравнения, получаем :

$$d_0 = \frac{1}{N}\sum_{u=1}^{N}\overline{Y}_u = \overline{Y} \tag{54}$$

$$d_1 = \frac{\sum_{u=1}^{N}(X_u - \overline{X})\overline{Y}_u}{\sum_{u=1}^{N}(X_u - \overline{X})^2} \tag{55}$$

Определим по формулам (54) и (55) коэффициенты регрессии для рассматриваемого примера. Расчеты необходимых сумм сводим в табл.9.

Таблица 9

Расчет сумм для определения коэффициентов регрессии

u	X_u	$X_u - \overline{X}$	$(X_u - \overline{X})^2$	$\overline{Y}u$	$(X_u - \overline{X})\overline{Y}u$
1	10	-25	625	2,06	-51,5
2	20	-15	225	4,28	-64,2
3	30	-5	25	6,18	-30,9
4	40	5	25	8,16	40,8
5	50	15	225	10,14	152,1
6	60	25	625	12,46	311,5
$\sum_{u=1}^{N}$	210	0	1750	43,28	357,8

По формуле (52) находим:

$$\overline{X} = \frac{210}{6} = 35.$$

$$d_0 = \overline{Y} = \frac{43,28}{6} = 7,21$$

$$d_1 = \frac{357,8}{1750} = 0,2.$$

По формулам (54) и (55) определяем:

Поэтому искомое уравнение имеет вид:

$$Y_R = 7,21 + 0,2(X - 35)$$

или

$$\mathrm{Y_R} = 0,21 + 0,2X.$$

График этой функции изображен на рис.9.

8.7. Определение адекватности полученного уравнения

Для определения адекватности полученного уравнения используют критерий Фишера, расчетное значение которого определяют по формуле:

$$F_R = \frac{S_{(2)}^2\{Y\}}{S_{(1)}^2\{Y\}} \qquad (56)$$

где $S_{(1)}^2\{Y\}$ - средняя дисперсия, или дисперсия воспроизводимости, определяемая по формуле (46);

$S_{(2)}^2\{Y\}$ - дисперсия, ххарактерзующая рассеивание средних экспериментальных значений $\overline{Y}_u$ относительно прямой ллинии определяемой уравнением регрессии Y_R.

Дисперсия $S^2_{(2)}\{Y\}$ характеризует точность аппроксимации зависимости $\overline{Y} = f(X)$ прямой линией и определяется по формуле:

$$S_{(2)}^2\{Y\} = \frac{m}{N-2}\sum_{u=1}^{N}(\overline{Y}_u - Y_{Ru})^2 \qquad (57)$$

Число степеней свободы этой дисперсии равно:

$$f\{S_{(2)}^2\} = N-2 \qquad (58)$$

Расчетное значение F_R сравнивают с табличным значением критерия Фишера F_T, которое определяют по приложению 3 в зависимости от доверительной вероятности p_D=0,95 и числа степеней свободы дисперсий $f\{S_{(1)}^2\}$ и $f\{S_{(2)}^2\}$. Если F_R<F_T , то гипотеза об адекватности линейного уравнения опытным данным не отвергается.

Расчет суммы в формуле (57) сведен в табл.10.

Используя данные табл.10 находим:

$$S_{(2)}^2\{Y\} = \frac{5}{6-2}0{,}0982 = 0{,}123$$

$$f\{S_{(2)}^2\} = 6-2 = 4$$

Таблица 10

Расчет суммы для определения дисперсии $S^2_{(2)}\{Y\}$

u	X_u	d_1X_u	Y_{Ru}	$\overline{Y}_u$	$\overline{Y}_u - Y_{Ru}$	$(\overline{Y}_u - Y_{Ru})^2$
1	10	2	2,21	2,06	-0,15	0,0225
2	20	4	4,21	4,28	0,07	0,0049
3	30	6	6,21	6,18	-0,03	0,0009
4	40	8	8,21	8,16	-0,05	0,0025
5	50	10	10,21	10,14	-0,07	0,0049
6	60	12	12,21	12,46	0,25	0,0625
$\sum_{u=1}^{N}$	210			43,28	0,02	0,0982

Подставляя найденные значения дисперсий в формулу (56), получаем:

Если F_R<1, то определяем обратное значение отношения дисперсий:

$$F_R = \frac{S^2_{(1)}\{Y\}}{S^2_{(2)}\{Y\}}$$

$$F_R = \frac{0{,}123}{0{,}107} = 1{,}14.$$

Сравниваем полученное значение критерия Фишера F_R с табличным значением, которое равно $F_T[p_D=0{,}95;\ f\{S^2_{(1)}\} = 24; f\{S^2_{(2)}\} = 4] = 2{,}78$. В рассматриваемом примере F_R=1,14<F_T=2,78, поэтому гипотеза об адекватности линейной модели не отвергается.

8.8. Определение значимости коэффициентов регрессии и их доверительных интервалов

Для оценки значимости коэффициентов регрессии используется критерий Стьюдента, расчетное значение которого определяется по формуле:

$$t_R\{d_i\} = \frac{|d_i|}{S\{d_i\}} \tag{59}$$

где $S\{d_i\}$ – оценка среднего квадратического отклонения коэффициента регрессии d_i.

Для оценки дисперсий коэффициентов регрессии d_0 и d_1 в уравнении (51) используют формулы:

$$S^2\{d_0\} = \frac{S^2\{Y\}}{mN} = \frac{S^2\{\overline{Y}\}}{N} \tag{60}$$

$$S^2\{d_1\} = \frac{S^2\{Y\}}{m\sum_{u=1}^{N}(X_u - \overline{X})^2} = \frac{S^2\{\overline{Y}\}}{\sum_{u=1}^{N}(X_u - \overline{X})^2} \tag{61}$$

В формулы (60) и (61) входит дисперсия $S^2\{Y\}$, которая является сводной оценкой дисперсии случайной величины Y_u выходного параметра при условии линейной связи (51). Эта дисперсия определяется по формуле:

$$S^2\{Y\} = \frac{(m-1)NS^2_{(1)}\{Y\} + (N-2)S^2_{(2)}\{Y\}}{mN-2} \tag{62}$$

Число степеней этой дисперсии определяется :

$$f\{S^2\} = mN - 2 \tag{63}$$

Подставив в формулу (62) ранее определенные значения $S^2_{(1)}\{Y\}$ и $S^2_{(2)}\{Y\}$, найдем сводную дисперсию случайной величины. Для рассматриваемого примера:

$$S^2\{Y\} = \frac{(5-1)\cdot 6\cdot 0{,}107 + (6-2)\cdot 0{,}123}{5\cdot 6 - 2} = 0{,}11$$

$$f\{S^2\} = 5\cdot 6 - 2 = 28.$$

По формулам (60) и (61) определяем дисперсии коэффициентов регрессии:

$$S^2\{d_0\} = \frac{0{,}11}{5\cdot 6} = 0{,}09, \qquad S\{d_0\} = 0{,}3$$

$$S^2\{d_1\} = \frac{0{,}11}{5\cdot 1750} = 0{,}000013, \qquad S\{d_1\} = 0{,}0035.$$

Расчетные значения критерия Стьюдента определяем по формуле (59):

$$t_R\{d_0\} = \frac{7,21}{0,3} = 23,85$$
$$t_R\{d_1\} = \frac{0,2}{0,0035} = 56,5.$$

По приложению 2 находим табличное значение критерия Стьюдента при условии, что доверительная вероятность p_D=0,95 и число степеней свободы, определяемое по формуле (63) $f\{S^2\}$=28. Следовательно, $t_T[p_D=0,95;f=28]=2,048$.

Так как $t_R\{d_0\}$=23,85>>t_T=2,048 и $t_R\{d_1\}$=56,5>> t_T=2,048, полученные коэффициенты значимы и, следовательно, связь между *Y* и *X* значима.

Доверительные абсолютные ошибки коэффициентов регрессии вычисляем по формуле:

$$\varepsilon\{d_i\} = S\{d_i\} \cdot t_T[p_D; f\{S^2\}] \quad (64)$$

Для данного примера эти ошибки равны:

$$\varepsilon\{d_0\} = 0,3 \cdot 2,048 = 0,61$$
$$\varepsilon\{d_1\} = 0,0035 \cdot 2,048 = 0,007$$

Доверительные интервалы для истинных значений коэффициентов регрессии d_0, d_1 в линейном уравнении (51) определяются неравенством:

$$d_i - \varepsilon\{d_i\} \le d_i \le d_i + \varepsilon\{d_i\} \quad (65)$$

Для рассматриваемого примера доверительные интервалы коэффициентов регрессии при p_D=0,95следующие:

$$7,21 - 0,61 \le d_0 \le 7,21 + 0,61$$
$$6,6 \le d_0 \le 7,82$$
$$0,2 - 0,007 \le d_1 \le 0,2 + 0,007$$
$$0,193 \le d_1 \le 0,0207$$

8.9. Определение доверительных интервалов средних значений выходного параметра при фиксированном значении фактора

Чтобы определить степень отклонения расчетных значений выходного параметра Y_{Ru} от истинного его значения при каждом уровне фактора X_u, определяем доверительные ошибки $\varepsilon\{Y_{Ru}\}$ расчетного значения выходного параметра и доверительные интервалы среднего значения выходного параметра.

Доверительные ошибки расчетных значений выходного параметра для каждого уровня фактора рассчитываются по формуле:

$$\varepsilon\{Y_{Ru}\} = S_m\{Y_{Ru}\} \cdot t_T[p_d; f\{S^2\}], \qquad (66)$$

где $S_m\{Y_{Ru}\}$ - оценка среднего квадратического отклонения расчетного значения выходного параметра Y_{Ru} для каждого значения X_u, определяемая по формуле:

$$S_m\{Y_{Ru}\} = \sqrt{S^2\{d_0\} + S^2\{d_1\} \cdot (X_u - \overline{X})^2} \qquad (67)$$

Для данного примера:

$$S_m\{Y_{Ru}\} = \sqrt{0{,}09 + 0{,}000013 \cdot (X_u - \overline{X})^2}.$$

Расчеты значений $S_m\{Y_{Ru}\}$для каждого u–го уровня фактора сведены в табл.11.

В рассматриваемом примере табличное значение критерия Стьюдента (см. пункт 5.8) равно $t_T[p_D=0{,}95;f=28]=2{,}048$. Подставляя это значение в формулу (66), получаем:

$$\varepsilon_m\{Y_{Ru}\} = 2{,}048 \cdot S_m\{Y_{Ru}\}.$$

Таблица 11

Расчет доверительных интервалов средних значений выходного параметра

u	X_u	$(X_u - \overline{X})^2$	$S^2_m\{Y_{Ru}\}$	$S_m\{Y_{Ru}\}$	$\varepsilon_m\{Y_{Ru}\}$	Y_{Ru}	$Y_{mR}^{(u)}(X)$	$Y_{mR}^{(0)}(X)$
1	10	625	0,099	0,315	0,65	2,21	1,56	2,86
2	20	225	0,094	0,307	0,63	4,21	3,58	4,84
3	30	25	0,092	0,303	0,62	6,21	5,59	6,83
4	40	25	0,092	0,303	0,62	8,21	7,59	8,83
5	50	225	0,094	0,307	0,63	10,21	9,58	10,84
6	60	625	0,099	0,315	0,65	12,21	11,56	12,86

В таблице 11 приведены полученные значения $S^2_m\{Y_{Ru}\}$, $S_m\{Y_{Ru}\}$ и $\varepsilon\{Y_{Ru}\}$ для каждого уровня фактора. Зная доверительные ошибки расчетных значений, можно найти доверительные интервалы для истинных средних значений выходного параметра, используя следующее неравенство:

$$Y^{(u)}_{mR}(X) = Y_{Ru} - \varepsilon\{Y_{Ru}\} \le Y_{Ru} \le Y_{Ru} + \varepsilon\{Y_{Ru}\} = Y^{(0)}_{mR}(X) \qquad (68)$$

На основе приведенных в табл.11 значений границ доверительного интервала строим график функций $Y_{mR}{}^{(u)}(X)$ и $Y_{mR}{}^{(0)}(X)$ (см. рис.8). Графики этих двух функций образуют своеобразный “коридор”. Любое сечение его прямой, параллельной вертикальной оси, соответствует доверительному интервалу, в котором с заданной вероятностью будет находиться истинное среднее значение выходного параметра. Легко заметить, что в этот коридор попадают средние экспериментальные значения $\overline{Y}u$. Однако некоторые индивидуальные экспериментальные значения выходного параметра в него не попадают, так как интервалы построены для средних значений.

8.10. Определение доверительных интервалов для индивидуальных значений выходного параметра при каждом уровне фактора

Границы доверительного интервала для индивидуальных значений выходного параметра Y_{uv} при каждом уровне фактора X_u определяются по формулам:

$$Y^{(u)}_{eR}(X) = Y_R(X) - \varepsilon_e\{Y_R\} \qquad (69)$$

$$Y^{(0)}_{eR}(X) = Y_R(X) + \varepsilon_e\{Y_R\} \qquad (70)$$

$$\varepsilon_e\{Y_R\} = S_e\{Y_{Ru}\} \cdot t_T[p_D; f\{S^2\}] \qquad (72)$$

$$S_e\{Y_{Ru}\} = \sqrt{S^2_m\{Y_{Ru}\} + S^2\{Y\}} \qquad (73)$$

Используя значения $S^2_m\{Y_{Ru}\}$ из табл.11 и ранее определенные по формуле (62) $S^2\{Y\}=0,11$ и $t_T[p_D=0,95;f=28]=2,048$, все расчеты верхней границы и нижней границы доверительного интервала по формулам (69) и (70) сводим в табл.12. Используя данные табл.12, строим графики функций $Y_{eR}{}^{(u)}(X)$ и

$Y_{eR}^{(0)}(X)$, которые являются доверительными границами зоны индивидуальных значений Y_{uv} выходного параметра (см. рис.9). Вероятность попадания точек, соответствующих индивидуальным значениям выходного параметра, равна 0,95, т.е. из ста измерений выходного параметра при любом уровне варьирования фактора 95 измерений попадают в эту зону и только 5 не попадают.

Таблица 12

Расчет доверительных интервалов для индивидуальных значений выходного параметра

u	X_u	$S^2{}_m\{Y_{Ru}\}$	$S^2{}_e\{Y_{Ru}\}$	$S_e\{Y_{Ru}\}$	$\varepsilon_e\{Y_{Ru}\}$	Y_{Ru}	$Y_{eR}^{(u)}(X)$	$Y_{eR}^{(0)}(X)$
1	10	0,099	0,209	0,457	0,936	2,21	1,274	3,146
2	20	0,094	0,204	0,452	0,925	4,21	3,285	5,135
3	30	0,092	0,201	0,449	0,919	6,21	5,291	7,129
4	40	0,092	0,201	0,449	0,919	8,21	7,291	9,129
5	50	0,094	0,204	0,452	0,925	10,21	9,285	11,135
6	60	0,099	0,209	0,457	0,936	12,21	11,274	13,146

Рассматривая индивидуальные значения Y_{uv} (см. табл.8) и границы зоны для каждого X_u (табл.12) замечаем, что все индивидуальные измерения попали в доверительную зону, т.е. располагаются между $Y_{eR}^{(u)}(X)$ и $Y_{eR}^{(0)}(X)$. На этом заканчивается статистическая обработка данных рассматриваемого однофакторного эксперимента.

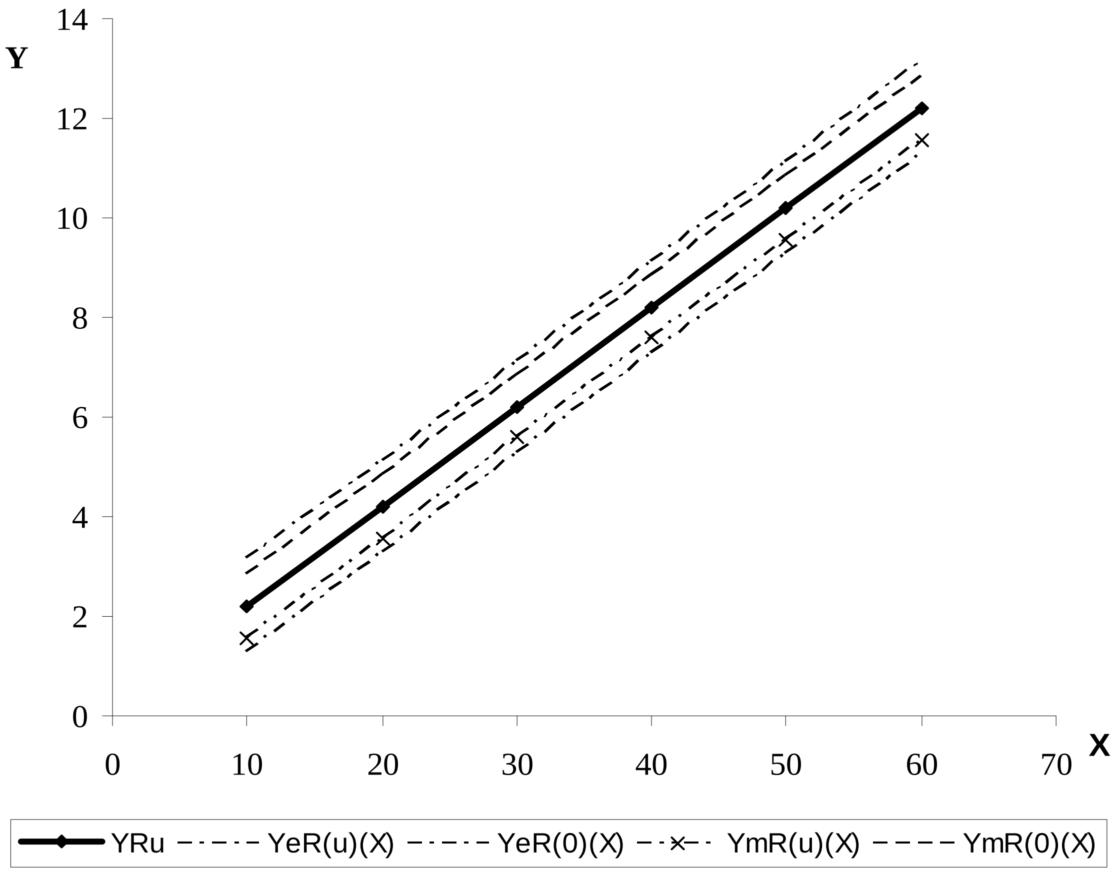


Рис.9. Линейная регрессионная однофакторная модель и ее доверительные интервалы

9. КВАДРАТИЧНАЯ ПАРАБОЛИЧЕСКАЯ ОДНОФАКТОРНАЯ РЕГРЕССИОННАЯ МОДЕЛЬ (МОДЕЛЬ ВТОРОГО ПОРЯДКА)

При определении параболической модели матрица планирования однофакторного эксперимента и условия проведения его одинаковы, как и при получении линейной регрессионной модели. Число уровней фактора, или число опытов в матрице планирования, обычно принимают N=5…12.

Рассмотрим операции, которые совершает исследователь при обработке данных этого эксперимента, на примере, в котором определялось влияние коэффициента крутки X (*число кручений на 1 м*) на разрывную нагрузку Y *(дан)* для льняной пряжи 333 текс. В табл.13 приведены значения X_u, $\overline{Y_u}$ и дисперсии $S^2_u\{Y\}$, полученные по данным пяти повторных опытов (m=5) при каждом уровне фактора X_u.

Таблица 13

Расчет основных статистических характеристик

X_u	v								
	u	Y_{uv}					$\sum_{v=1}^{m} Y_{uv}$	$\bar{Y}_u$	$S^2_u\{Y\}$	W_R
		1	2	3	4	5				
60	1	4,4	4,6	4,8	5,4	4,7	23,9	4,78	0,142	3,579
80	2	6,7	6,4	6,6	6,1	6	31,8	6,36	0,093	3,691
100	3	7,2	6,5	7,1	7	6,8	34,6	6,92	0,077	3,754
120	4	6,7	6,5	6,2	6,1	6,8	32,3	6,46	0,093	3,691
140	5	6,1	6,3	5,5	5,7	6,1	29,7	5,94	0,108	3,654

9.1. Исключение резко выделяющихся данных

Данная операции осуществляется аналогично как и для определения линейной однофакторной регрессионной модели. Для этого определяются расчетные значения критерия Смирнова-Грабса по формулам (37-38). Рассмотрим эту операцию при анализе первого опыта матрицы u=1, когда X=60, Y_{uvmax}=5,4, Y_{uvmin}=4,4:

$$V_{R\max 1} = \frac{5{,}4 - 4{,}78}{\sqrt{0{,}142}}\sqrt{\frac{5}{5-1}} = 1{,}839$$

$$V_{R\min 1} = \frac{4{,}78 - 4{,}4}{\sqrt{0{,}142}}\sqrt{\frac{5}{5-1}} = 1{,}127$$

По приложению 1 находим, что $V_T[p_D=0{,}95;\ m=5]=1{,}869$. Так как $V_{Rmax}<V_T$ и $V_{Rmin}<V_T$, то рассмотренные значения Y_{uvmax}=5,4 и Y_{uvmin}=4,4 не являются резко выделяющимися и остаются для дальнейшей статистической обработки.

Аналогично рассчитываются расчетные значения критерия Смирнова-Грабса V_{Rmax} и V_{Rmin} для других опытов матрицы, затем они сравниваются с табличным значением критерия V_T. Если расчетные значения критерия Смирнова-Грабса превышают табличное, то соответствующие значения Y_{uvmax}

или Y_{uvmin} исключаются из дальнейшего расчета, а среднее значение и дисперсия пересчитываются для соответствующего опыта матрицы.

9.2. Проверка гипотезы о нормальном распределении случайных величин Y_{uv}

Проверка этой гипотезы для каждого u–го опыта матрицы состоит в определении расчетного значения критерия W_R по формулам (39-40).

Для 1-го опыта матрицы при u=1 и X=60 располагаем значения Y_{uv} по возрастанию: 5,4 ≥4,8≥4,7 ≥4,6 ≥4,4.

$$k = \frac{m-1}{2} = \frac{5-1}{2} = 2$$

Используя приложение 5, находим q_1=0,6646 и q_2=0,2413 и вычисляем значение Q_1 и W_{R1}:

$$Q_1 = 0{,}6646(5{,}4-4{,}4)+0{,}2413(4{,}8-4{,}6)=0{,}713$$

$$W_{R1} = \frac{0{,}713^2}{0{,}142} = 3{,}579$$

Расчетное значение W_{R1} сравнивают с табличным W_T, которое определяется по приложению 5. W_T определяется для заданной доверительной вероятности p_D и известного числа повторных опытов m. Для рассматриваемого примера $W_T[p_D=0{,}95;\ m=5]=0{,}762$.

Так как расчетное значение W_{R1} превышает табличное значение W_T для выбранной доверительной вероятности, то гипотеза о нормальном распределении случайных величин не отвергается.

В табл. 13 приведены значения W_R и для других опытов матрицы. Эти значения также превышают табличное, и поэтому первое условие о возможности применения регрессионного анализа удовлетворяется.

9.3. Проверка гипотезы об однородности дисперсий в опытах матрицы

Так как число повторных опытов (m=5) одинаково для всех опытов матрицы, то для проверки однородности дисперсий применяется критерий Кочре-

$$G_R = \frac{0{,}142}{0{,}513} = 0{,}277$$

на, расчетное значение которого определяется по формуле (41):

Расчетное значение G_R сравнивается с табличным значением G_T, которое определяют по приложению 4 в зависимости от числа опытов в матрице N и числа степеней свободы дисперсии $f\{S^2_u\}=m-1$ для заданной доверительной вероятности p_D. В рассматриваемом примере $G_T[p_D=0{,}95; N=5; f=5-1=4]=0{,}5441$. Так как $G_R < G_T$, то гипотеза об однородности дисперсий, т.е. равноточности и воспроизводимости опытов, не отвергается.

Если $G_R > G_T$, то дисперсии в N рядах измерений неоднородны. После отбрасывания $S^2_{u\ max}\{Y\}$ описанную выше процедуру следует повторить для N-1 рядов измерений.

Если число повторных опытов неодинаково при различных уровнях факторов, то для проверки однородности дисперсий используется критерий Бартлера, расчетное значение которого определяется по формулам (42-45).

9.4. Определение средней дисперсии выходного параметра в опытах матрицы

Если в опытах матрицы дисперсии однородны и число повторных опытов одинаково , то средняя дисперсия определяется по формуле(46), а число степеней свободы этой дисперсии по формуле(47).

Средняя дисперсия характеризует средний разброс значений выходного параметра относительно его средних значений при каждом уровне факторов, т.е. ошибку опытов в эксперименте. В рассматриваемом примере эта дисперсия, или, как ее называют дисперсия воспроизводимости, равна:

$$S_{(1)}^2 = \frac{0,513}{5} = 0,1026$$

$$f\left\{S_{(1)}^2\right\} = 5 \cdot (5-1) = 20.$$

9.5. Определение подходящего вида регрессионной модели

Для определения подходящего вида регрессионной модели используют следующую информацию:

1) графическую взаимосвязь $\overline{Y} = f(X)$ между средними значениями выходного параметра для каждого уровня факторов и значением фактора по данным эксперимента. При сопоставлении этого графика с графиками известных функций устанавливают вид уравнения;

2) характер изменения разделенных и неразделенных разностей первого порядка, определяемых по данным эксперимента по формулам (48-49).

В рассматриваемом примере графическая взаимосвязь $\overline{Y} = f(X)$ между средними значениями выходного параметра $\overline{Y}$ для каждого уровня факторов и значением фактора X приведена на рис.10. При сопоставлении этого графика с графиками известных функций можно сделать вывод, что для описания экспериментальных данных наиболее подходит параболическая модель.

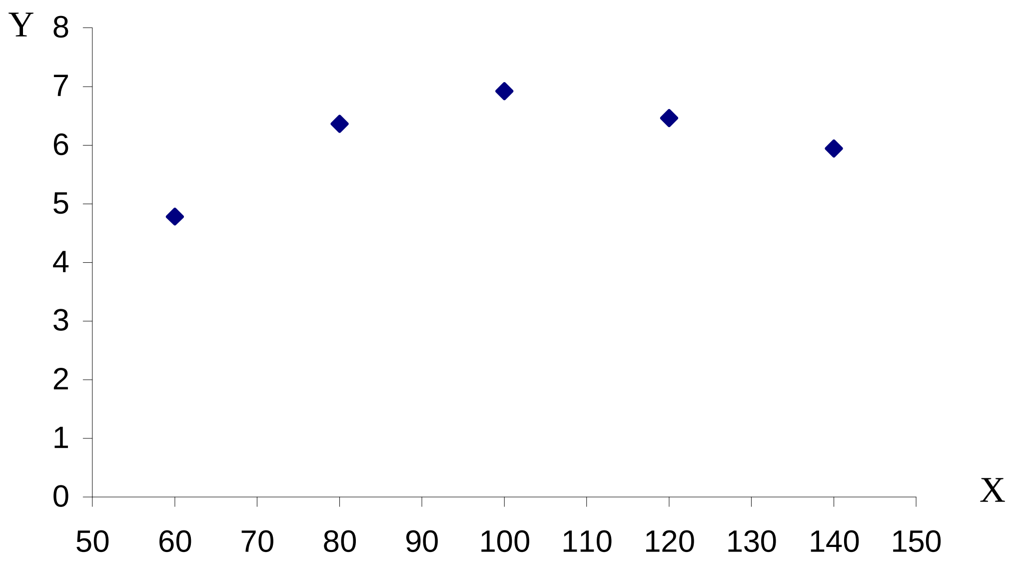


Рис.10 Зависимость разрывной нагрузки от крутки льняной пряжи

В рассматриваемом примере интервал варьирования факторов постоянный и равен I_X=80-60=100-80=120-100=140-120=20. Поэтому определяем неразделенные разности первого порядка по формуле (49):

$$\Delta_{H1} = 6{,}36 - 4{,}78 = 1{,}58$$
$$\Delta_{H2} = 6{,}92 - 6{,}36 = 0{,}56$$
$$\Delta_{H3} = 6{,}46 - 6{,}92 = -0{,}46$$
$$\Delta_{H4} = 5{,}94 - 6{,}46 = -0{,}52$$

Так как неразделенные разности первого порядка $\Delta_{H\max} - \Delta_{H\min} = 1{,}58 - 0{,}46 = 1{,}12$, превышают удвоенную величину среднеквадратической ошибки эксперимента $2S_{(1)}\{Y\}$=0,641, поэтому они являются нетождественными, и экспериментальные данные не могут быть описаны линейным уравнением.

Таким образом, для описания экспериментальных данных условно можно принять полином второй степени или квадратичную параболическую однофакторную модель:

$$Y_R = a_0 + a_1 X_1 + a_{11} X_1^2 \tag{74}$$

9.6. Определение коэффициентов регрессии

Если дисперсии выходного параметра для каждого уровня фактора однородны, то для определения коэффициентов регрессии в уравнении (74) можно применять метод наименьших квадратов. Используя условие $\sum\limits_{u=1}^{N}(Y_u - Y_{Ru})^2 = \min$, устанавливают следующие нормальные уравнения:

$$\begin{cases} a_0 N + a_1 \sum\limits_{u=1}^{N} X_u + a_{11} \sum\limits_{u=1}^{N} X_u^2 = \sum\limits_{u=1}^{N} \overline{Y_u} \\ a_0 \sum\limits_{u=1}^{N} X_u + a_1 \sum\limits_{u=1}^{N} X_u^2 + a_{11} \sum\limits_{u=1}^{N} X_u^3 = \sum\limits_{u=1}^{N} X_u \overline{Y_u} \\ a_0 \sum\limits_{u=1}^{N} X_u^2 + a_1 \sum\limits_{u=1}^{N} X_u^3 + a_{11} \sum\limits_{u=1}^{N} X_u^4 = \sum\limits_{u=1}^{N} X_u^2 \overline{Y_u} \end{cases} \tag{75}$$

где N – общее число опытов в матрице планирования эксперимента.

Величины $\sum X_u, \sum X_u^2$ и т.д., входящие в нормальные уравнения (75), могут быть определены по данным эксперимента и для рассматриваемого примера сведены в табл.14.

Таблица 14

Расчет сумм для определения коэффициентов регрессии

u	X_u	X_u^2	X_u^3	X_u^4	$\overline{Y_u}$	$X_u\overline{Y_u}$	$X_u^2\overline{Y_u}$
1	60	3600	216000	12960000	4,78	286,8	17208
2	80	6400	512000	40960000	6,36	508,8	40704
3	100	10000	1000000	100000000	6,92	692	69200
4	120	14400	1728000	207360000	6,46	775,2	93024
5	140	19600	2744000	384160000	5,94	831,6	116424
$\sum$	500	54000	6200000	745440000	30,46	3094,4	336560

Подставляя эти значения в уравнения (75) получаем систему уравнений:

Данную систему уравнений можно решить относительно a_0, a_1, a_{11} мето-

$$\begin{cases} 5a_0 + 500a_1 + 54000a_{11} = 30{,}46 \\ 500a_0 + 54000a_1 + 6200000a_{11} = 3094{,}4 \\ 54000a_0 + 6200000a_1 + 745440000a_{11} = 336560 \end{cases}$$

дом последовательного исключения неизвестных или с помощью матричного метода.

Определение коэффициентов регрессии в уравнении и статистических характеристик параметров уравнения упрощается, если в матрице планирования эксперимента используются кодированные значения факторов и опыты располагаются симметрично относительно основного уровня фактора. Эти условия удовлетворяются в рассматриваемом примере.

Значение основного уровня фактора определяется по формуле:

$$X_0 = \overline{X} = \frac{X_{\min} + X_{\max}}{2} \qquad (76)$$

Интервал варьирования фактора рассчитывается:

$$I_x = \frac{X_{\max} - X_{\min}}{N - 1} \qquad (77)$$

Кодированные значения уровней фактора определяют по формуле:

Для рассматриваемого примера эти значения равны:

$$x_1 = \frac{X_1 - X_0}{I_x} \tag{78}$$

$$x_2 = \frac{X_2 - X_0}{I_x}$$

и т.д

$$X_0 = \frac{60+140}{2} = 100$$

$$I_x = \frac{140-60}{5-1} = 20$$

$$x_1 = \frac{60-100}{20} = -2$$

$$x_2 = \frac{80-100}{20} = -1$$

$$x_3 = \frac{100-100}{20} = 0$$

$$x_4 = \frac{120-100}{20} = 1$$

$$x_5 = \frac{140-100}{20} = 2$$

Матрица планирования эксперимента для кодированных значений фактора X приведена в табл.15.

Таблица 15

Матрица планирования эксперимента для кодированных значений фактора

u	X_u	x_u	x_0	$\overline{Y}_u$	x_u^2	$x_u \overline{Y}_u$	$x_u^2 \overline{Y}_u$	x_u^4
1	60	-2	1	4,78	4	-9,56	19,12	16
2	80	-1	1	6,36	1	-6,36	6,36	1
3	100	0	1	6,92	0	0	0	0
4	120	1	1	6,46	1	6,46	6,46	1
5	140	2	1	5,94	4	11,88	23,76	16
$\sum$	500	0	5	30,46	10	2,42	55,7	34

Матрица планирования эксперимента (см. табл.15) обладает свойством ортогональности:

$$\sum_{u=1}^{N} x_0 x_u = 0$$

Поэтому решение системы уравнений (75) исключается, и расчет коэффициентов регрессии полинома второго порядка ведется по следующим формулам:

Подставляя в эти формулы соответствующие значения сумм из табл.15,

$$Y_R = b_0 + b_1 x + b_{11} x^2 \tag{79}$$

$$b_0 = \frac{1}{B}\sum_{u=1}^{N} x_u^4 \cdot \sum_{u=1}^{N} x_0 \overline{Y_u} - \frac{1}{B}\sum_{u=1}^{N} x_u^2 \cdot \sum_{u=1}^{N} x_u^2 \overline{Y_u} - \frac{1}{B} \tag{80}$$

$$b_1 = \frac{\sum_{u=1}^{N} x_u \overline{Y_u}}{\sum_{u=1}^{N} x_u^2} \tag{81}$$

$$b_{11} = \frac{N}{B}\sum_{u=1}^{N} x_u^2 \overline{Y_u} - \frac{1}{B}\sum_{u=1}^{N} x_u^2 \cdot \sum_{u=1}^{N} x_0 \overline{Y_u} \tag{82}$$

$$B = N\sum_{u=1}^{N} x_u^4 - \left(\sum_{u=1}^{N} x_u^2\right)^2 \tag{83}$$

получаем:

$$B = 5 \cdot 34 - 10^2 = 70$$

$$b_0 = \frac{1}{70} \cdot 34 \cdot 30{,}46 - \frac{1}{70} \cdot 10 \cdot 55{,}7 = 6{,}838$$

$$b_1 = \frac{2{,}42}{10} = 0{,}242$$

$$b_{11} = \frac{5}{70} \cdot 55{,}7 - \frac{1}{70} \cdot 10 \cdot 30{,}46 = -0{,}373$$

Уравнение (79) в кодированных значения фактора x имеет вид:

$$Y_R = 6{,}838 + 0{,}242x - 0{,}373x^2 \tag{84}$$

Переход от коэффициентов b_i при кодированных значениях фактора к коэффициентам a_i при натуральных значениях фактора для параболического уравнения осуществляется по формулам:

$$a_1 = \frac{b_1}{I_x} - \frac{2b_{11}}{I_x^2}\overline{X} \qquad (85)$$

$$a_{11} = \frac{b_{11}}{I_x^2} \qquad (86)$$

$$a_0 = b_0 - \frac{b_1}{I_x}\overline{X} + \frac{b_{11}}{I_x^2}\overline{X^2} \qquad (87)$$

Для рассматриваемого примера коэффициенты регрессии при натуральных значениях фактора рассчитываются:

$$a_1 = \frac{0{,}242}{20} - \frac{2\cdot(-0{,}373)}{20^2}\cdot 100 = 0{,}199$$

$$a_{11} = \frac{-0{,}373}{20^2} = -0{,}0009$$

$$a_0 = 6{,}838 - \frac{0{,}242}{20}\cdot 100 + \frac{(-0{,}373)}{20^2}\cdot 100^2 = -3{,}694$$

Уравнение в натуральных значениях фактора X имеет вид:

$$Y_R = -3{,}694 + 0{,}199X - 0{,}0009X^2 \qquad (88)$$

9.7. Определение адекватности полученного уравнения

Для определения адекватности полученного уравнения используют критерий Фишера, расчетное значение которого определяют по формуле:

$$F_R = \frac{S^2_{над}\{Y\}}{S^2\{\overline{Y}\}} \qquad (89)$$

где $S^2_{над}\{Y\}$ – дисперсия, характеризующая неадекватность, которая определяется по формуле (90);

$S^2\{\overline{Y}\}$ – дисперсия среднего значения выходного параметра, определяемая по формуле (92).

$$S^2_{над} = \frac{\sum\left(\overline{Y_u} - Y_{Ru}\right)^2}{N - N_k} \qquad (90)$$

где N_k – число коэффициентов в уравнении (87);

Y_{Ru} – значение выходного параметра для u-го опыта, определяемое по формуле (84).

Число степеней свободы для этой дисперсии определяется:

$$f\left\{S^2_{над}\right\} = N - N_k \tag{91}$$

$$f\left\{S^2_{над}\right\} = 5 - 3 = 2$$

Для определения дисперсии, характеризующей неадекватность, составляем таблицу 16.

Таблица 16

Расчет суммы для определения дисперсии $S^2_{над}\{Y\}$

u	x_u	$\overline{Y}_u$	Y_R	$\overline{Y}_u - Y_R$	$(\overline{Y}_u - Y_R)^2$
1	-2	4,78	4,862	-0,082	0,0067
2	-1	6,36	6,223	0,137	0,0188
3	0	6,92	6,838	0,082	0,0067
4	1	6,46	6,707	-0,247	0,061
5	2	5,94	5,830	0,11	0,0121
$\sum$	0	30,46	30,46		0,1053

Пользуясь данными табл.16, находим:

$$S^2_{ад}\{Y\} = \frac{0,1053}{5-3} = 0,0527$$

По данным табл.13 определяем дисперсию среднего значения выходного

$$S^2\left\{\overline{Y}\right\} = \frac{0,1026}{5} = 0,0205$$

$$S^2\left\{\overline{Y}\right\} = \frac{S^2_{1(1)}\{Y\}}{m} = \frac{\sum_{u=1}^{N} S^2_u\{Y\}}{m \cdot N} \tag{92}$$

параметра:

Расчетное значение критерия Фишера определяем по формуле:

$$F_R = \frac{0,0527}{0,0205} = 2,566$$

Табличное значение критерия Фишера находим по приложению 3 $F_T[p_D=0,95;\quad f\{S^2_{(1)}\} = 20; f\{S^2_{над}\} = 2] = 3,49$. В рассматриваемом примере $F_R=2,566<F_T=3,49$, поэтому гипотеза об адекватности линейной модели не отвергается.

9.8. Определение значимости коэффициентов регрессии и их доверительных интервалов

Для оценки значимости коэффициентов регрессии используется критерий Стьюдента, расчетное значение которого определяется по формуле (59).

Для оценки дисперсий коэффициентов регрессии b_0, b_1 и b_{11} в уравнении (84) используют формулы:

$$S^2\{b_0\} = \frac{S^2\{\overline{Y}\}}{B}\sum_{u=1}^{N} x_u^4 \qquad (93)$$

$$S^2\{b_1\} = \frac{S^2\{\overline{Y}\}}{\sum_{u=1}^{N} x_u^2} \qquad (94)$$

$$S^2\{b_{11}\} = \frac{N \cdot S^2\{\overline{Y}\}}{B} \qquad (95)$$

Определим дисперсии коэффициентов регрессии для рассматриваемого примера:

$$S^2\{b_0\} = \frac{0{,}0205}{70} \cdot 34 = 0{,}00997$$

$$S\{b_0\} = 0{,}0998$$

$$S^2\{b_1\} = \frac{0{,}0205}{10} = 0{,}00205$$

$$S\{b_1\} = 0{,}0453$$

$$S^2\{b_{11}\} = \frac{5 \cdot 0{,}0205}{70} = 0{,}0015$$

$$S\{b_{11}\} = 0{,}0383$$

Расчетные значения критерия Стьюдента определяем по формуле (59):

$$t_R\{b_0\} = \frac{6{,}838}{0{,}0998} = 68{,}49$$

$$t_R\{b_1\} = \frac{0{,}242}{0{,}0453} = 5{,}34$$

$$t_R\{b_{11}\} = \frac{0{,}373}{0{,}0383} = 9{,}74$$

.

По приложению 2 находим табличное значение критерия Стьюдента при условии, что доверительная вероятность p_D=0,95 и число степеней свободы, определяемое по формуле (47) $f\{S^2_{(1)}\}$=20. Следовательно, $t_T[p_D=0{,}95;f=20]=2{,}086$.

Так как $t_R\{b_i\}$>t_T, то полученные коэффициенты значимы и, следовательно, связь между Y и X значима.

Доверительные абсолютные ошибки коэффициентов регрессии вычисляем по формуле (64):

$$\varepsilon\{b_0\} = 0{,}0998 \cdot 2{,}086 = 0{,}208$$
$$\varepsilon\{b_1\} = 0{,}0453 \cdot 2{,}086 = 0{,}094$$
$$\varepsilon\{b_{11}\} = 0{,}0383 \cdot 2{,}086 = 0{,}08$$

Доверительные интервалы для истинных значений коэффициентов регрессии b_0, b_1, b_{11} в параболическом уравнении (84) определяются неравенством (65):

$$6{,}838 - 0{,}208 \le b_0 \le 6{,}838 + 0{,}208$$
$$6{,}63 \le b_0 \le 7{,}046$$
$$0{,}242 - 0{,}094 \le b_1 \le 0{,}242 + 0{,}094$$
$$0{,}148 \le b_1 \le 0{,}336$$
$$-0{,}373 - 0{,}08 \le b_{11} \le -0{,}373 + 0{,}08$$
$$-0{,}453 \le b_{11} \le -0{,}293$$

9.9. Определение доверительных интервалов средних значений выходного параметра при фиксированном значении фактора

Чтобы определить степень отклонения расчетных значений выходного параметра Y_{Ru} от истинного его значения при каждом уровне фактора X_u, определяем доверительные ошибки $\varepsilon\{Y_{Ru}\}$ расчетного значения выходного параметра и доверительные интервалы среднего значения выходного параметра.

Доверительные ошибки расчетных значений выходного параметра для каждого уровня фактора рассчитываются по формуле:

$$\varepsilon\{Y_{Ru}\} = S\{Y_{Ru}\} \cdot t_T[p_d; f\{S^2_{(1)}\}], \qquad (96)$$

где $S^2\{Y_{Ru}\}$ - дисперсия расчетного значения выходного параметра Y_{Ru} для каждого значения X_u, определяемая по формуле:

$$S^2\{Y_{Ru}\} = \left[\frac{1}{B}\sum_{u=1}^{N} x_u^4 + \left(\frac{1}{\sum\limits_{u=1}^{N} x_u^2} - \frac{2}{B}\sum_{u=1}^{N} x_u^2\right)x^2 + \frac{N}{B}x^4\right] S_u^2\{\overline{Y}\} \qquad (97)$$

Для данного примера:

$$S^2\{Y_{Ru}\} = \left[\frac{34}{70} + \left(\frac{1}{10} - \frac{2}{70}\cdot 10\right)x^2 + \frac{5}{70}x^4\right]\cdot 0{,}0205 =$$

$$= \left[0{,}4857 - 0{,}1857x^2 + 0{,}0714x^2\right]\cdot 0{,}0205$$

Расчеты значений $S^2\{Y_{Ru}\}$для каждого u–го уровня фактора сведены в табл.17.

Таблица 17

Расчет доверительных интервалов средних значений выходного параметра

u	X_u	$S^2\{Y_{Ru}\}$	$S\{Y_{Ru}\}$	Y_{Ru}	$\varepsilon\{Y_{Ru}\}$	$Y_{mR}^{(u)}(X)$	$Y_{mR}^{(0)}(X)$
1	-2	0,0182	0,1349	4,862	0,2814	4,581	5,143
2	-1	0,0076	0,0872	6,223	0,1819	6,041	6,405
3	0	0,01	0,1	6,838	0,2086	6,629	7,047
4	1	0,0076	0,0872	6,707	0,1819	6,525	6,889
5	2	0,0182	0,1349	5,830	0,2814	5,549	6,111

В рассматриваемом примере табличное значение критерия Стьюдента (см. пункт 6.8) равно $t_T[p_D=0{,}95;f=20]=2{,}086$. Подставляя это значение в формулу (96), получаем:

$$\varepsilon_m\{Y_{Ru}\} = 2{,}086 \cdot S_m\{Y_{Ru}\}.$$

В таблице 17 приведены полученные значения $S^2\{Y_{Ru}\}$, $S\{Y_{Ru}\}$ и $\varepsilon\{Y_{Ru}\}$ для каждого уровня фактора. Зная доверительные ошибки расчетных значений, можно найти доверительные интервалы для истинных средних значений выходного параметра, используя формулу (68):

$$Y_{mR}^{(u)}(X) = Y_{Ru} - \varepsilon\{Y_{Ru}\} \le Y_{Ru} \le Y_{Ru} + \varepsilon\{Y_{Ru}\} = Y_{mR}^{(0)}(X)$$

На основе приведенных в табл.17 значений границ доверительного интервала строим график функций $Y_R(X)$, ${Y_{mR}}^{(u)}(X)$ и ${Y_{mR}}^{(0)}(X)$ (см. рис.11). Графики этих функций образуют своеобразный “коридор”. Любое сечение его прямой, параллельной вертикальной оси, соответствует доверительному интервалу, в котором с заданной вероятностью будет находиться истинное среднее значение выходного параметра. Легко заметить, что в этот коридор попадают средние экспериментальные значения $\overline{Y}u$. Однако некоторые индивидуальные экспериментальные значения выходного параметра в него не попадают, так как интервалы построены для средних значений.

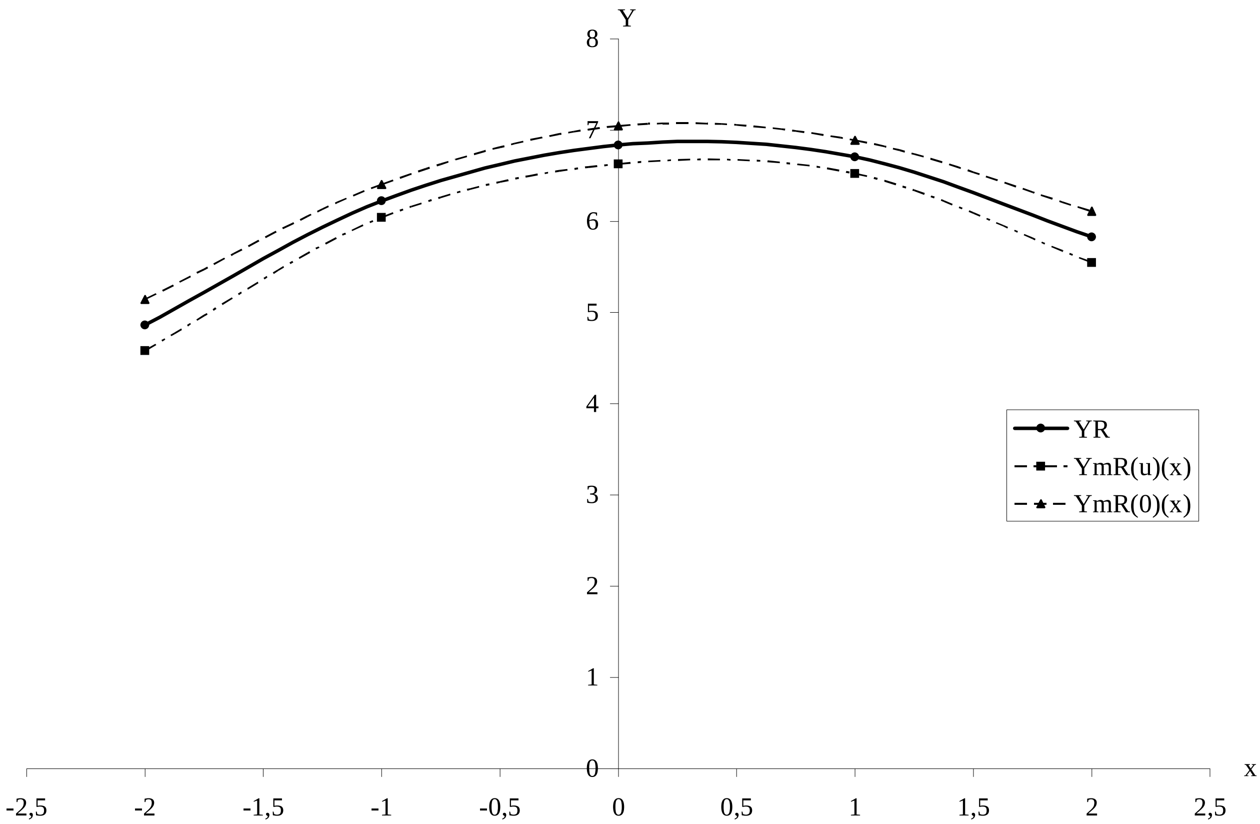


Рис.11. Параболическая регрессионная однофакторная модель и ее доверительные интервалы

10. ОПРЕДЕЛЕНИЕ СТАТИСТИЧЕСКИХ РЕГРЕССИОННЫХ МНОГОФАКТОРНЫХ МОДЕЛЕЙ ПО ДАННЫМ ЭКСПЕРИМЕНТА С ФАКТОРНЫМ ПЛАНИРОВАНИЕМ.

10.1. Полный факторный эксперимент

Полный факторный эксперимент (ПФЭ) – это эксперимент, который реализует все возможные неповторяющиеся комбинации уравнений исследуемых факторов.

Регрессионная многофакторная модель, получаемая по результатам ПФЭ, имеет вид линейного полинома:

$$Y_R = b_0 + b_1 x_1 + b_2 x_2 + \ldots + b_i x_i + \ldots + b_M x_M \tag{98}$$

или неполного полинома второго порядка:

$$Y_R = b_0 + b_1 x_1 + \ldots + b_i x_i + \ldots + b_M x_M + b_{12} x_1 x_2 + \ldots + b_{ij} x_i x_j + \ldots + b_{M-1} x_{M-1} x_M \tag{99}$$

где $У_R$ – расчетное значение выходного параметра;

x_i – кодированные значения уравнений факторов;

b_i, b_{ij} – выборочные значения коэффициентов регрессии;

$i = 1 \ldots M$; $j = 1 \ldots M$ $(i \neq j)$ – номер фактора.

Определение регрессионной модели на базе ПФЭ включает следующие этапы:

1) проведение предварительного эксперимента;
2) планирование ПФЭ;
3) нахождение условий для проведения ПФЭ;
4) проведение основного эксперимента по матрицам планирования эксперимента;
5) обработка результатов эксперимента;
6) анализ полученной модели.

Предварительный эксперимент проводится с целью определение точности измерения выходного параметра и доверительного объема его измерений и для проверки гипотезы о нормальности распределения случайной величины $У_{uv}$.

Планирование эксперимента включает:

- составление матрицы планирования;
- планирование повторных опытов;

- рандомизация опытов (последовательность их проведения).

При факторном планировании в отличие от традиционного (однофакторного) планирования по величине коэффициентов регрессии b_i, b_{ij} в регрессионной модели можно судить не только о влиянии каждого фактора x_i, но и их взаимодействия $x_i x_j$, т.е. изменения влияния одного фактора при переходе второго на другой уровень.

Число коэффициентов регрессии в регрессионной модели (98) и (99) определяется по формуле:

$$N_k = M + 1 + C_M^2 \qquad (100)$$

где M – число факторов в эксперименте;

C_M^2 – число сочетаний из М элементов по n=2, которое для регресионной модели (98) равно нулю, а для модели (99) определяется по формуле:

$$C_M^2 = \frac{M(M-1)}{2} \qquad (101)$$

Если для каждого фактора число уравнений одинаково и равно k, то число возможных неповторяющихся комбинаций или число опытов в матрице ПЭФ равно:

$$N = k^M \qquad (102)$$

Расчеты по формулам (100-101) показывают, что все коэффициенты регрессии для моделей (98-99) можно определить при варьировании факторов на двух уровнях ($k=2$), так как для модели (98) $N>N_k$, а для модели (99) $N=N_k$. В случае $N=N_k$ (т.е., когда число степеней свободы $f=N-N_k=0$) нельзя оценить адекватность регрессионной модели, так как недостает степеней свободы.

Из формулы (102), что с увеличением числа факторов значительно увеличится число опытов в матрице ПФЭ. Поэтому использование опытов ПФЭ дает эффект в том случае, когда выходной параметр $У$ зависит не более, чем от четырех факторов, т.е. при $M \leq 4$.

Матрицу планирования ПФЭ обозначают 2^M. Матрица планирования для двухфакторного процесса 2^2 с кодированными значениями факторов представлена в табл. 18.

Таблица 18

Матрица планирования эксперимента ПФЭ 2^2

u	факторы		
	x_0	x_1	x_2
1	+1	-1	-1
2	+1	+1	-1
3	+1	-1	+1
4	+1	+1	+1

где u – номер опыта в матрице;

$x_{вi}$ =+1 = +– кодированное значение верхнего уровня фактора;

$x_{нi}$ =-1 = – – кодированное значение нижнего уровня фактора;

x_0 = +1 – уровень фиктивного фактора.

Кодированные значения уравнений фактора определяются по формулам:

$$x_{вi} = \frac{X_{вi} - X_{0i}}{I_i} = +1$$

$$x_{нi} = \frac{X_{нi} - X_{0i}}{I_i} = -1$$

где X_{0i} – натуральное значение основного уровня i-го фактора;

$X_{вi}$ – натуральное значение верхнего уровня i-го фактора;

$X_{нi}$ – натуральное значение нижнего уровня i-го фактора;

I_i – интервал варьирования i-го фактора.

Каждый опыт матрицы планирования эксперимента обычно повторяется не менее двух раз с целью определения дисперсии воспроизводимости, характеризующей ошибку опыта. Порядок проведения опытов по матрице планирования должен быть *рандомизирован*, т.е. соответствовать последовательности случайных чисел, определяемой по специальной таблице.

Рандомизацию опытов проводят для предупреждения влияния различных воздействий (например, изменение температуры и влажности воздуха, изменение состава и характеристик сырья и т.д.).

В табл. 19 представлена матрица планирования эксперимента по определению влияния заправочных параметров вязания на поверхностную плотность

Y основовязального полотна, выработанного переплетением трико-цепочка-уток. Эксперимент ПФЭ 2^3 проводился при трех повторных опытах (m=3) с рандомизированным порядком их проведения.

Таблица 19

Матрица планирования эксперимента ПФЭ 2^3

u	Факторы				Рандомизация опытов			$У_{uv}$			$\overline{Y}_u$	$S_u^2\{Y\}$
	x_0	x_1	x_2	x_3	1	2	3	$У_{u1}$	$У_{u2}$	$У_{u3}$		
1	+	-	-	-	13	24	12	553	548	552	551,00	7,00
2	+	+	-	-	4	19	14	557	558	561	558,67	4,33
3	+	-	+	-	3	9	22	558	559	564	560,33	10,33
4	+	+	+	-	23	5	1	568	565	571	568,00	9,00
5	+	-	-	+	15	7	20	407	402	400	403,00	13,00
6	+	+	-	+	18	21	17	410	408	405	407,67	6,33
7	+	-	+	+	8	10	6	415	410	408	411,00	13,00
8	+	+	+	+	16	2	11	418	415	421	418,00	9,00
Σ											3877,67	72,00

Для рандомизации всех N·m=24 повторных опытов (см. табл.) находят случайный порядок 24 чисел по таблице случайных чисел (приложение 7), из которых первые 8 чисел заносят в 1 столбец, вторые 8 – во 2 столбец, третьи 8 – в 3 столбец. Таким образом, согласно порядку рандомизации повторных опытов первым проведен опыт при условиях, соответствующих опыту u=4 матрицы планирования, и получено значение выходного параметра Y_{43}=571. Рандомизацию опытов матрицы можно проводить также раздельно для каждой серии повторных опытов, т.е. по столбцам Y_{uv}.

Условия проведения эксперимента представлены в табл.20.

Таблица 20

Факторы и интервалы их варьирования

Наименование фактора	Кодированное значение фактора	Интервал варьирования фактора	Уровни варьирования факторов		
			-1	0	+1
Входное натяжение нитей цепочки, сН	x_1	2,5	3	5,5	8
Входное натяжении нитей трико, сН	x_2	2	2,5	4,5	6,5
Линейная плотность уточной нити, текс	x_3	40	540	580	620

Обработка результатов ПФЭ, представленных в табл. 19 проводится в следующей последовательности.

10.2. Исключение резко выделяющихся данных

Исключение резко выделяющихся значений выходного параметра осуществляется по методике, изложенной в главе 8. Для каждого опыта матрицы эксперимента рассчитываются значения критерия Смирнова-Грабса по формулам (37-38). Для первого опыта матрицы:

$$V_{R\max 1} = \frac{553-551}{2{,}64}\sqrt{\frac{3}{3-1}} = 0{,}926$$

$$V_{R\min 1} = \frac{551-548}{2{,}64}\sqrt{\frac{3}{3-1}} = 1{,}389$$

По приложению 1 находим, что $V_T[p_D=0{,}95;\ m=3]=1{,}412$. Так как $V_{Rmax}<V_T$ и $V_{Rmin}<V_T$, то рассмотренные значения для первого опыта Y_{uvmax}=553 и Y_{uvmin}=548 не являются резко выделяющимися и остаются для дальнейшей статистической обработки.

Аналогично рассчитываются расчетные значения критерия Смирнова-Грабса V_{Rmax} и V_{Rmin} для других опытов матрицы, затем они сравниваются с табличным значением критерия V_T.

10.3. Проверка гипотезы об однородности дисперсий в опытах матрицы

Проверка гипотезы об однородности дисперсии в опытах матрицы с помощью критерия Кочрена, расчетное значение которого определяется по формуле (41). Используя данные табл. 19, находим:

$$G_R = \frac{S^2_{u\max}\{Y\}}{\sum\limits_{u=1}^{N} S^2_u\{Y\}} = \frac{13}{72} = 0{,}18$$

Табличное значение критерия Кочрена (см. приложение 4):

$G_T[p_D=0{,}95;\ N=8;\ f=3\text{-}1=2]=0{,}5157.$

Т.к. $G_R = 0{,}18 < G_T = 0{,}5157$, то дисперсии однородны и проведенный ПФЭ 2^3 обладает свойствами воспроизводимости.

Если дисперсии неоднородны, то нужно увеличивать число повторных опытов.

10.4. Определение оценок коэффициентов регрессии

Коэффициенты регрессии в регрессионных многофакторных моделях (98-99) определяются по следующим формулам:

$$b_i = \frac{1}{N}\sum_{u=1}^{N} x_{iu}\overline{Y}_u \qquad (103)$$

$$i = 0,1,2,\ldots M$$

$$b_{ij} = \frac{1}{N}\sum_{u=1}^{N} x_{iu}x_{ju}\overline{Y}_u \qquad (104)$$

$$i \neq j$$

$$b_{ijl} = \frac{1}{N}\sum_{u=1}^{N} x_{iu}x_{ju}x_{lu}\overline{Y}_u \qquad (105)$$

$$i \neq j \neq l$$

Пользуясь данными табл. 19, находим по формуле (103):

$$b_0 = \frac{1}{8}(551 + 558{,}67 + 560{,}33 + 568 + 403 + 407{,}67 + 411 + 418) = 484{,}71$$

$$b_1 = \frac{1}{8}(-551 + 558{,}67 - 560{,}33 + 568 - 403 + 407{,}67 - 411 + 418) = 3{,}375$$

$$b_2 = \frac{1}{8}(-551 - 558{,}67 + 560{,}33 + 568 - 403 - 407{,}67 + 411 + 418) = 4{,}625$$

$$b_3 = \frac{1}{8}(-551 - 558{,}67 - 560{,}33 - 568 + 403 + 407{,}67 + 411 + 418) = -74{,}79$$

По формуле (104) определяем коэффициенты регрессии при двойных взаимодействиях факторов:

$$b_{12} = \frac{1}{8}(551 - 558{,}67 - 560{,}33 + 568 + 403 - 407{,}67 - 411 + 418) = 0{,}29$$

$$b_{13} = \frac{1}{8}(551 - 558{,}67 + 560{,}33 - 568 - 403 + 407{,}67 - 411 + 418) = -0{,}46$$

$$b_{23} = \frac{1}{8}(551 + 558{,}67 - 560{,}33 - 568 - 403 - 407{,}67 + 411 + 418) = 0{,}04$$

По формуле (105) определяем коэффициент регрессии при тройном взаимодействии факторов:

$$b_{123} = \frac{1}{8}(-551 + 558{,}67 + 560{,}33 - 568 + 403 - 407{,}67 - 411 + 418) = 0{,}29$$

В результате расчетов получаем регрессионную многофакторную модель:

$$Y_R = 484{,}71 + 3{,}375x_1 + 4{,}625x_2 - 74{,}79x_3 + 0{,}29x_1x_2 - 0{,}46x_1x_3 - 0{,}04x_2x_3 + 0{,}29x_1x_2x_3$$

Данная модель не является окончательной моделью изучаемого процесса. После проверки значимости коэффициентов регрессии эта модель уточняется.

10.5. Определение значимости коэффициентов регрессии и их доверительных интервалов

Проверка значимости коэффициентов регрессии осуществляется по критерию Стьюдента, расчетное значение которого $t_R\{b_i\}$ сравнивают с табличным t_T. Если $t_R\{b_i\} > t_T$, то гипотеза о значимости коэффициентов регрессии не отвергается.

Расчетное значение критерия Стьюдента равно:

$$t_R\{b_i\} = \frac{|b_i|}{S\{b_i\}} \qquad (106)$$

где $S\{b_i\}$ – среднее квадратическое отклонение выборочного коэффициента регрессии.

Т.к. матрица планирования эксперимента обладает свойством ортогональности, поэтому дисперсии коэффициентов регрессии одинаковы, т.е. $S^2\{b_i\}= S^2\{b_{ij}\}= S^2\{b_{ijl}\}$ и определяются по формуле:

$$S^2\{b_i\} = \frac{1}{N} S^2\{\overline{Y}\} \tag{107}$$

$$S^2\{\overline{Y}\} = \frac{1}{m} S^2\{\mathrm{Y}\} \tag{107}$$

Если доказано, что дисперсии $S_u^2\{Y\}$ однородны, то дисперсия воспроизводимости равна:

$$S^2\{Y\} = \frac{1}{N} \sum_{u=1}^{N} S_u^2\{Y\} \tag{108}$$

Пользуясь данями табл. 19, для рассматриваемого примера находим дисперсии коэффициентов регрессии и расчетные значения критерия Стьюдента:

$$S^2\{Y\} = \frac{72}{8} = 9$$

$$S^2\{\overline{Y}\} = \frac{9}{3} = 3$$

$$S^2\{b_i\} = \frac{3}{8} = 0{,}375 \qquad\qquad S\{b_i\} = 0{,}61$$

$$t_R\{b_1\} = \frac{3{,}375}{0{,}61} = 5{,}51$$

$$t_R\{b_2\} = \frac{4{,}625}{0{,}61} = 7{,}55$$

$$t_R\{b_3\} = \frac{74{,}79}{0{,}61} = 122{,}1$$

$$t_R\{b_{12}\} = \frac{0{,}29}{0{,}61} = 0{,}48$$

$$t_R\{b_{13}\} = \frac{0{,}46}{0{,}61} = 0{,}75$$

$$t_R\{b_{23}\} = \frac{0{,}04}{0{,}61} = 0{,}07$$

$$t_R\{b_{123}\} = \frac{0{,}29}{0{,}61} = 0{,}48$$

Табличное значение критерия Стьюдента определяют по приложению 2 при условии, что $p_D=0{,}95$ и число степеней свободы $f\{S_u^2\}=N(m-1)$:

$t[p_D=0{,}95;f=8(3-1)=16]=2{,}12$

Сопоставляя расчетные и табличные значения критерия Стьюдента, находим, что только коэффициенты регрессии b_1, b_2, b_3, значимы. Значимость коэффициентов регрессии зависит не только от удельного влияния данного фактора на выходной параметр, но и от интервала варьирования уровней фактора. Незначимость может быть обусловлена также малым интервалом варьирования фактора, большой дисперсией воспроизводимости вследствие наличия неуправляемых и неконтролируемых факторов, а также расположением основного уровня фактора X_{0i} близко к точке частного экстремума Y по этому фактору.

Границы доверительных интервалов коэффициентов регрессии определяются неравенством:

$$b_i - \varepsilon\{b_i\} \le b_i \le b_i + \varepsilon\{b_i\} \tag{109}$$

где $\varepsilon\{b_i\}$ – ошибка коэффициента регрессии, определяемая по формуле:

$$\varepsilon\{b_i\} = t_T \cdot S\{b_i\} \tag{110}$$

Для ортогональных матриц коэффициенты регрессии независимы и доверительные границы определяются раздельно для каждого коэффициента. Поэтому в случае незначимости какого-то коэффициента регрессии он может быть отброшен без пересчета всех остальных.

Для рассматриваемого примера искомая регрессионная многофакторная модель процесса включает только значимые коэффициенты:

$$Y_R = 484{,}71 + 3{,}375x_1 + 4{,}625x_2 - 74{,}79x_3 \tag{111}$$

10.6. Определение адекватности полученного уравнения

Проверку адекватности полученной модели можно проводить только при условии $N - N_k > 0$, т.е., когда на базе матрицы ПФЭ определяют регрессионную многофакторную модель по формуле (98) или (99), в которой отсутствует хотя бы один член. Для проверки гипотезы об адекватности используют критерий Фишера, расчетное значение которого F_R сравнивают с табличным F_T. Если $F_R < F_T$, то с вероятностью $p_D=0{,}95$ гипотеза об адекватности полученной модели не отвергается. Расчетное значение критерия Фишера равно:

$$F_R = \frac{S_{ад}^2\{Y\}}{S^2\{Y\}} = \frac{S_{над}^2\{Y\}}{S^2\{\overline{Y}\}}, \qquad \text{если } S_{над}^2\{Y\} > S^2\{\overline{Y}\} \tag{112}$$

$$F_R = \frac{S^2\{Y\}}{S_{ад}^2\{Y\}} = \frac{S^2\{\overline{Y}\}}{S_{над}^2\{Y\}}, \qquad \text{если } S^2\{\overline{Y}\} > S_{над}^2\{Y\} \tag{113}$$

где $S^2\{\overline{Y}\}$ – дисперсия, характеризующая ошибку опыта и определяемая по формуле (107);

$S_{ад}^2\{Y\}$ – дисперсия, обусловленная неадекватностью регрессионной моделью и определяемая по формуле:

$$S_{ад}^2\{Y\} = mS_{над}^2\{Y\} = \frac{m\sum_{u=1}^{N}(\overline{Y}_u - Y_{Ru})^2}{N - N_k} \tag{114}$$

Для расчета $S_{над}^2\{Y\}$ данные сводятся в таблицу 21.

Таблица 21

Расчет адекватности регрессионной модели

u	Y_{Ru}	$\overline{Y}_u$	$\overline{Y}_u - Y_R$	$(\overline{Y}_u - Y_R)^2$
1	551,5	551,00	-0,50	0,2500
2	558,25	558,67	0,42	0,1736
3	560,75	560,33	-0,42	0,1736
4	567,5	568,00	0,50	0,2500
5	401,92	403,00	1,08	1,1736
6	408,67	407,67	-1,00	1,0000
7	411,17	411,00	-0,17	0,0278
8	417,92	418,00	0,08	0,0069
$\sum$	-	3877,67		3,0556

Значения рассчитывают по формуле (111) с использованием данных табл. 19. В рассматриваемом примере:

$$S_{над}^2\{Y\} = \frac{3{,}0556}{8-4} = 0{,}76$$

Так как $S_{над}^2\{Y\} < S^2\{\overline{Y}\}$, то для определения расчетного значения критерия Фишера используют формулу (113):

$$F_R = \frac{3}{0{,}76} = 3{,}93$$

Табличное значение критерия Фишера находят по приложению 3:

$F[p_D=0,95;f\{S_u^2\}=8(3\text{-}1)=16;\ f\{S_{над}^2\}=8\text{-}4=4]=5,86$

Так как. $F_R=3,93 < F_T=5,86$ гипотеза об адекватности регрессионной модели не отвергается.

Если гипотеза об адекватности отвергается, необходимо переходить к описанию процесса полиномом 2-го порядка на базе другого вида эксперимента или, если это возможно, проводить эксперимент с меньшим интервалом варьирования уровней факторов.

10.7. Анализ регрессионной многофакторной модели

После получения математической модели проводится ее анализ. В моделях вида (98) значения коэффициентов регрессии характеризуют вклад соответствующего фактора в величину выходного параметра при переходе фактора с основного уровня на верхний или нижний. Вклад фактора при переходе от нижнего к верхнему уровню в величину выходного параметра называют *эффектом фактора*. Эффект фактора равен $B_i=2b_i$.

Чем больше коэффициент регрессии, тем выше эффект этого фактора, т.е. тем сильнее проявляется влияние фактора на выходной параметр. Таким образом, по величине коэффициентов регрессии в модели (98) можно осуществить ранжировку факторов по силе их влияния на выходной параметр Y. Знак перед коэффициентом регрессии определяет характер влияния фактора на выходной параметр Y. Факторы, коэффициенты которых имеют знак плюс, повышают величину выходного параметра, а имеющие знак минус – снижают ее.

При указанном выше анализе моделей (98) необходимо учитывать следующее:

1) с увеличением интервала варьирования фактора абсолютное значение выходного параметра увеличивается;

2) знаки линейных коэффициентов регрессии не изменяются при движении по поверхности отклика до экстремальной зоны, но изменяются за этой зоной;

3) незначимость коэффициента регрессии для какого-либо фактора может быть обусловлена рядом указанных выше причин

Чтобы выяснить причину незначимости коэффициента регрессии b_i при x_i, вновь проводят ПФЭ с другими значениями X'_{0i} и I'_i, причем значения I_i и m увеличиваются (для уменьшения $S\{b_i\}$). Если в этом эксперименте коэффициент регрессии для того же фактора окажется незначим, можно сделать вывод о близости основного уровня этого фактора к оптимальному значению.

Усложнение анализа модели (99) объясняется тем, что влияние взаимодействия факторов на выходной параметр Y неоднозначно и зависит не только от величины и знака коэффициентов b_{ij}, но и от знаков коэффициентов b_i и b_j. Часто для наглядности используют регрессионную модель, в которой факторы имеют натуральные значения. Переход от коэффициентов регрессии b_i, найденных при кодированных значениях факторов x_i, к коэффициентам a_i при натуральных значениях факторов проводится подстановкой значений из формулы (78) в модель (98) или (99).

Для линейной модели (98) коэффициенты регрессии a_i в преобразованной модели $Y_R=a_0+a_1X_1+ a_2X_2+\ldots+ a_MX_M$ определяются по формулам:

$$a_0 = b_0 - \sum_{i=1}^{M} \frac{b_i X_{0i}}{I_i} \qquad (115)$$

$$a_i = \frac{b_i}{I_i} \qquad (116)$$

Если в матрицу ПФЭ подставить натуральные значения верхних и нижних уровней факторов, получим матрицу, которая не обладает свойством ортогональности, и коэффициенты регрессии, рассчитанные по формулам (103-105), для этой матрицы будут зависимы. Поэтому по величине и знакам коэффициентов a_i в преобразованной модели нельзя судить о силе и характере влияния факторов на выходной параметр Y, а следовательно, и нельзя проводить ранжировку факторов.

11. ПАССИВНЫЙ ЭКСПЕРИМЕНТ. ЛИНЕЙНАЯ ОДНОФАКТОРНАЯ КОРРЕЛЯЦИОННАЯ МОДЕЛЬ

Рассмотрим операции, которые совершает исследователь при обработки данных однофакторного пассивного эксперимента, на примере, в котором изучалось влияние X – обхвата груди (см) на Y – рост (см) у мужчин некоторого города. В результате эксперимента получено m = 26 парных значений, приведенных в табл.22.

Таблица 22

Результаты пассивного эксперимента

X	94	96	99	87	100	101	96	102	96	101	90	98	99
У	171	173	167	161	183	171	169	176	173	173	170	165	175
X	96	93	96	100	102	93	92	95	95	93	96	100	95
У	171	169	167	175	169	172	165	168	173	172	165	178	164

11.1. Составление корреляционной таблицы

Для построения корреляционной таблицы рассчитывают число интервалов K_y и K_x, используя эмпирическую формулу:

$$K_y \approx K_x \approx 1 + 3{,}32 \lg m \qquad (117)$$

$$K_y \approx K_x \approx 1 + 3{,}32 \cdot \lg 26 = 5{,}68$$

В рассматриваемом примере принимаем $K_y = K_x$, =5

Для определения величины интервалов для распределения совокупности случайных величин по ячейкам корреляционной таблицы используют формулы:

В корреляционной таблице (см. табл.23) показаны границы интервалов X_j

$$I_x = \frac{X_{\max} - X_{\min}}{K_x} = \frac{102 - 87}{5} = 3$$

$$I_y = \frac{Y_{\max} - Y_{\min}}{K_y} = \frac{183 - 161}{5} = 4{,}4$$

–X_{j+1}, Y_i – Y_{i+1} и средние значения этих интервалов $\overline{X}_j$ и $\overline{Y}_ш$.

В соответствии с принятыми интервалами все 26 пар значений распределяют по ячейкам корреляционной таблицы, отмечая в правом верхнем углу ячейки частоту m_{ji}.

11.2. Кодирование случайных величин

С целью упрощения расчетов проводят кодирование исходных случайных величин X и Y. Кодированные значения средних интервалов случайных величин определяют по формулам:

$$x_j = \frac{\overline{X}_j - X_0}{I_x} \tag{118}$$

$$y_i = \frac{\overline{Y}_i - Y_0}{I_y}$$

где X_0, Y_0 – координаты условного центра таблицы, который чаще всего соответствует ячейке с максимальным значением частоты.

Для рассматриваемого примера ячейка с максимальным значением частоты m_{ji}=4, ее координаты Y_0 = 172 и X_0 = 94,5.

Для первого интервала, у которого $\overline{X}_1$=88,5 кодированное значение середины интервала определяется:

$$x_1 = \frac{88{,}5 - 94{,}5}{3} = -2$$

Для первого интервала, у которого $\overline{Y}_1$=163,2 кодированное значение середины интервала определяется:

$$y_1 = \frac{163{,}2 - 172}{4{,}4} = -2$$

Для следующих интервалов кодированные значения равны:

$$x_2 = \frac{91{,}5 - 94{,}5}{3} = -1 \qquad y_2 = \frac{167{,}6 - 172}{4{,}4} = -1$$

$$x_3 = \frac{94{,}5 - 94{,}5}{3} = 0 \qquad y_3 = \frac{172 - 172}{4{,}4} = 0$$

$$x_4 = \frac{97{,}5 - 94{,}5}{3} = 1 \qquad y_4 = \frac{176{,}4 - 172}{4{,}4} = 1$$

$$x_5 = \frac{100{,}5 - 94{,}5}{3} = 2 \qquad y_5 = \frac{180{,}8 - 172}{4{,}4} = 2$$

Рассчитанные значения x y проставлены в соответствующие ячейки табл.23.

В левом верхнем углу ячейки с частотой m_{ji} – записывают значение $m_{ji}y_i$, а в левом нижнем углу – $m_{ji}x_j$.

11.3. Определение средних значений кодированных случайных величин

Определяют общие средние значения кодированных величин:

$$\bar{x}=\frac{1}{m}\sum_{j=1}^{h} m_{xj}x_j \qquad (119)$$

$$\bar{y}=\frac{1}{m}\sum_{i=1}^{q} m_{yi}y_i \qquad (120)$$

$$m_{xj}=\sum_{i=1}^{q} m_{ji}$$

$$m_{yi}=\sum_{j=1}^{p} m_{ji} \qquad (121)$$

$$m=\sum_{i=1}^{q} m_{yi}=\sum_{j=1}^{p} m_{xj} \qquad (122)$$

где $m_{xj}x_j$, $m_{yi}y_i$ – определяют по следующим формулам в соответствующих ячейках корреляционной таблицы:

где p, q – число интервалов для X и Y.

Для рассматриваемого примера:

$$\bar{x}=\frac{21}{26}=0{,}81$$

$$\bar{y}=\frac{-10}{26}=-0{,}38$$

11.4. Определение средних значений натуральных случайных величин

Определяют общие средние значения натуральных случайных величин:

$$\overline{X}=I_x\cdot\bar{x}+X_0 \qquad (123)$$

$$\overline{Y}=I_y\cdot\bar{y}+Y_0 \qquad (124)$$

Для данного примера:

$$\overline{X}=3\cdot 0{,}81+94{,}5=96{,}93$$

$$\overline{Y}=4{,}4\cdot(-0{,}38)+172=170{,}33$$

Таблица 23

Корреляционная таблица

$Y_i - Y_{i+1}$	$\overline{Y}_i$	y_i	87 – 90	90 – 93	93 – 96	96 – 99	99– 102	m_{yi}	$m_{yi}y_i$	$m_{yi}y_i^2$	$\sum m_{ji}x_j$	$y_i\sum m_{ji}x_j$	$(\sum m_{ji}x_j)^2$	$\frac{(\sum m_{ji}x_j)^2}{m_{yi}}$
			$\overline{X}_j$											
			88,5	91,5	94,5	97,5	100,5							
			x_j											
			-2	-1	0	1	2							
161,0 – 165,4	163,2	-2	-2 1 -2	-2 1 -1	-2 1 0	-4 2 2		5	-10	20	-10	2	1	0,2
165,4 – 169,8	167,6	-1			-2 2 0	-2 2 2	-2 2 4	6	-6	6	6	-6	36	6
169,8 – 174,2	172,0	0		0 1 -1	0 4 0	0 3 3	0 2 4	10	0	0	6	0	36	3,6
174,2 – 178,6	176,4	1					4 4 8	4	4	4	8	8	64	16
178,6 – 183,0	180,8	2					2 1 2	1	2	4	2	4	4	4
m_{xj}			1	2	7	7	9	26	-10	34	21	8	140	29,8
$m_{xj}x_j$			-2	-2	0	7	18	21						
$m_{xj}x_j^2$			4	2	0	7	36	49						
$\sum m_{ji}y_i$			-2	-2	-4	-6	4	-10						
$x_j\sum m_{ji}y_i$			4	2	0	-6	8	8						
$(\sum m_{ji}y_i)^2$			4	4	16	36	16	76						
$\frac{(\sum m_{ji}y_i)^2}{m_{xj}}$			4	2	2,29	5,14	1,78	15,21						

11.5. Определение дисперсии и среднего квадратического отклонения кодированных случайных величин

Определяем дисперсии кодированных случайных величин по формулам:

$$S^2\{x\} = \frac{1}{m}\left[\sum_{j=1}^{p} m_{xj} x_j^2 - m \cdot \left(\overline{x}\right)^2\right] \qquad (125)$$

$$S^2\{y\} = \frac{1}{m}\left[\sum_{i=1}^{q} m_{yi} y_i^2 - m \cdot \left(\overline{y}\right)^2\right] \qquad (126)$$

$$S^2\{x\} = \frac{1}{26}\left(49 - 26 \cdot (0{,}81)^2\right) = 1{,}23$$

$$S^2\{y\} = \frac{1}{26}\left(34 - 26 \cdot (-0{,}38)^2\right) = 1{,}16$$

Среднее квадратическое отклонение кодированных случайных величин:

$$S\{x\} = 1{,}11$$
$$S\{y\} = 1{,}08$$

11.6. Определение дисперсии и среднего квадратического отклонения натуральных случайных величин

Дисперсия натуральных случайных величин определяется по формулам:

$$S^2\{X\} = I_x^2 \cdot S^2\{x\} \qquad (127)$$

$$S^2\{Y\} = I_y^2 \cdot S^2\{y\} \qquad (128)$$

Определяем дисперсии натуральных случайных величин X и Y:

$$S^2\{X\} = 3^2 \cdot 1{,}23 = 11{,}07$$
$$S^2\{Y\} = 4{,}4^2 \cdot 1{,}16 = 22{,}46$$

Среднее квадратичное отклонение натуральных случайных величин:

$S\{X\} = 3{,}33$

$S\{Y\} = 4{,}74$

11.7. Определение коэффициента корреляции и коэффициента детерминации

Коэффициент корреляции определяют по формуле:

$$r_{yx}=\frac{\sum_{i=1}^{q}\sum_{j=1}^{p}m_{ji}y_ix_j-m\overline{y}\overline{x}}{m\cdot S\{y\}\cdot S\{x\}} \qquad (129)$$

Коэффициент корреляции для данного примера:

$$r_{yx}=\frac{8-26\cdot 0{,}81\cdot(-0{,}38)}{26\cdot 1{,}11\cdot 1{,}08}=0{,}51$$

Коэффициент детерминации равен $KD_r=r^2_{yx}$=0,26. Этот коэффициент показывает, что только 26% изменений выходного параметра обусловлено изменениями входного параметра и 74% – другими факторами и случайными воздействиями.

11.8. Определение значимости коэффициента корреляции

Определяем значимость коэффициента корреляции, используя критерий Стьюдента, расчетное значение которого равно:

$$t_R\{r_{yx}\}=\frac{r_{yx}\sqrt{m-2}}{\sqrt{1-r_{yx}^2}} \qquad (130)$$

Для рассматриваемого примера:

$$t_R\{r_{yx}\}=\frac{0{,}51\cdot\sqrt{26-2}}{\sqrt{1-0{,}51^2}}=2{,}9$$

Табличное значение критерия определяем по приложению 2:
t_T [p_D = 0,95; f = m – 2 = 24] = 2,064. Так как t_R = 2,9 > t_T = 2,064, то коэффициент корреляции значим, и гипотеза о наличии корреляционной зависимости *Y* и *X* не отвергается.

11.9. Определение дисперсионного и корреляционного отношения

Определяем дисперсионное и корреляционное отношение по данным табл.19:

$$h_{yx}^2=\frac{m\sum_{j=1}^{p}\frac{1}{m_{xj}}\left(\sum_{i=1}^{q}m_{ji}y_i\right)^2-\left(\sum_{j=1}^{p}\sum_{i=1}^{q}m_{ji}y_i\right)^2}{m\sum_{i=1}^{q}m_{yi}y_i^2-\left(\sum_{i=1}^{q}m_{yi}y_i\right)^2} \qquad (131)$$

$$h_{xy}^2 = \frac{m\sum_{i=1}^{q}\frac{1}{m_{yi}}\left(\sum_{j=1}^{p} m_{ji}x_j\right)^2 - \left(\sum_{j=1}^{p}\sum_{i=1}^{q} m_{ji}x_j\right)^2}{m\sum_{j=1}^{p} m_{xj}x_j^2 - \left(\sum_{j=1}^{p} m_{xj}x_j\right)^2} \tag{132}$$

Для рассматриваемого примера:

$$h_{yx}^2 = \frac{26\cdot 15{,}21 - 10^2}{26\cdot 34 - 10^2} = 0{,}377$$

$$h_{yx} = 0{,}61$$

$$h_{xy}^2 = \frac{26\cdot 29{,}8 - 21^2}{26\cdot 49 - 21^2} = 0{,}4$$

$$h_{yx} = 0{,}63$$

11.10. Определение значимости корреляционного отношения

Определяем значимость корреляционного отношения, используя критерий Стьюдента, расчетное значение которого равно:

$$t_R\{h_{yx}\} = \frac{h_{xy}\sqrt{m-2}}{\sqrt{1-h_{yx}^2}} \tag{133}$$

$$t_R\{h_{xy}\} = \frac{h_{yx}\sqrt{m-2}}{\sqrt{1-h_{xy}^2}} \tag{134}$$

Для рассматриваемого примера:

$$t_R\{h_{yx}\} = \frac{0{,}61\cdot\sqrt{26-2}}{\sqrt{1-0{,}377}} = 3{,}84$$

$$t_R\{h_{xy}\} = \frac{0{,}63\cdot\sqrt{26-2}}{\sqrt{1-0{,}4}} = 4{,}0$$

Табличное значение критерия определяем по приложению 2: t_T [p_D = = 0,95; f = m – 2 = 24] = 2,064.

Так как $t_R > t_T$, то корреляционные отношения значимы и гипотеза о наличии корреляционной связи между X и Y не отвергается.

11.11. Проверка гипотезы о линейной связи между Y и X

Гипотезу о линейной связи между X и Y проверяют с помощью крите-

$$F_R = \frac{\left(h_{yx}^2 - r_{yx}^2\right)/(p-2)}{\left(1-h_{yx}^2\right)/(m-p)} \tag{135}$$

рия Фишера, расчетное значение которого определяется по формуле:

Для рассматриваемого примера:

$$F_R = \frac{(0{,}377 - 0{,}26)/(5-2)}{(1-0{,}377)/(26-5)} = 1{,}3$$

Табличное значение критерия Фишера определяем по приложению 3: F_T [p_D = 0,95; $f_{числ}$ = p – 2=5 – 2 = 3; $f_{знам}$ = m – p =26 – 5 = 21] = 8,66.

Так как F_R = 1,3 < F_T = 8,66, то гипотеза о линейной взаимосвязи между случайными величинами *Y* и *X* не отвергается.

11.12. Определение коэффициентов в корреляционных уравнениях

Определяют коэффициенты в корреляционных уравнениях 2-х сопряженных прямых:

$$Y_R(X) = d_{0x} + d_{1x}\left(X - \overline{X}\right) \qquad (136)$$

$$X_R(Y) = d_{0y} + d_{1y}\left(Y - \overline{Y}\right) \qquad (137)$$

Коэффициенты линейной корреляционной однофакторной модели определяют по формулам:

$$d_{0x} = \overline{Y} \qquad (138)$$

$$d_{1x} = \frac{r_{yx} S\{Y\}}{S\{X\}} \qquad (139)$$

$$d_{0y} = \overline{X} \qquad (140)$$

$$d_{1x} = \frac{r_{yx} S\{X\}}{S\{Y\}} \qquad (141)$$

Для рассматриваемого примера:

$$d_{0x} = 170{,}33$$

$$d_{1x} = \frac{0{,}51 \cdot 4{,}74}{3{,}33} = 0{,}73$$

$$d_{0y} = 96{,}93$$

$$d_{1y} = \frac{0{,}51 \cdot 3{,}33}{4{,}74} = 0{,}36$$

Сопряженные корреляционные уравнения имеют вид:

$$Y_R = 170{,}33 + 0{,}73 \cdot (X - 96{,}33) = 170{,}33 + 0{,}73X - 70{,}32 = 100{,}01 + 0{,}73X$$

$$X_R = 96{,}93 + 0{,}36 \cdot (Y - 170{,}33) = 96{,}93 + 0{,}36X - 61{,}32 = 35{,}61 + 0{,}36Y$$

Расчетные значения $Y_R(\overline{X}_j)$ приведены в табл.24. На рис.12 приведен график линейной корреляционной однофакторной модели $Y_R(\overline{X}_j)$.

Таблица 24

Расчет линейной корреляционной однофакторной модели и ее доверительных интервалов

$\overline{X}_j$	88,5	91,5	94,5	97,5	100,5
$Y_R(X) = Y_R(\overline{X}_j)$	164,615	166,805	168,995	171,185	173,375
$\sum m_{ji} y_i$	-2	-2	-4	-6	4
$\frac{I_y}{m_{xj}} \sum m_{ji} y_i$	-8,8	-4,4	-2,51429	-3,77143	1,955556
$\overline{Y}_x = \overline{Y}(\overline{X}_j)$	163,2	167,6	169,486	168,229	173,956
$\overline{X}_j - \overline{X}$	-8,43	-5,43	-2,43	0,57	3,57
$\left(\overline{X}_j - \overline{X}\right)^2$	71,065	29,485	5,905	0,325	12,745
$S_m^2\left\{Y_R(\overline{X}_j)\right\}$	5,12	2,528	1,058	0,71	1,484
$S_m\left\{Y_R(\overline{X}_j)\right\}$	2,263	1,59	1,029	0,843	1,218
$\varepsilon_m\{Y_{Ri}\}$	4,67	3,282	2,123	1,739	2,515
$S_e^2\left\{Y_R(\overline{X}_j)\right\}$	23,06	20,468	18,998	18,65	19,424
$S_e\left\{Y_R(\overline{X}_j)\right\}$	4,802	4,524	4,359	4,319	4,407
$\varepsilon_e\{Y_{Ri}\}$	9,911	9,338	8,996	8,914	9,097

11.13. Определение значимости коэффициентов в корреляционных уравнениях

Определяют значимость коэффициентов в сопряженных корреляционных уравнениях, используя критерий Стьюдента:

$$t_R\{d_1\} = \frac{|d_1|}{S\{d_1\}} = \frac{|d_1| \cdot S\{X\}\sqrt{m-2}}{S\{Y\} \cdot \sqrt{1 - r_{yx}^2}} \qquad (142)$$

$$t_R\{d_{1x}\} = \frac{0{,}73 \cdot 3{,}33 \cdot \sqrt{26-2}}{4{,}74 \cdot \sqrt{1-0{,}26}} = 2{,}92$$

$$t_R\{d_{1y}\} = \frac{0{,}36 \cdot 4{,}743 \cdot \sqrt{26-2}}{3{,}33 \cdot \sqrt{1-0{,}26}} = 2{,}92$$

Табличное значение критерия Стьюдента определяем по приложению 2: t_T [p_D = 0,95; f = m – 2 = 26 – 2 =24] = 2,064.

Так как $t_R\{d_{1x}\} > t_T$ и $t_R\{d_{1y}\} > t_T$, то оба коэффициента значимы.

11.14. Определение доверительных интервалов коэффициентов искомых уравнений

Доверительные интервалы коэффициентов искомых уравнений определяют из неравенств:

$$d_{1x} - \varepsilon\{d_{1x}\} \le \delta_x \le d_{1x} + \varepsilon\{d_{1x}\} \tag{143}$$
$$d_{1y} - \varepsilon\{d_{1y}\} \le \delta_y \le d_{1y} + \varepsilon\{d_{yx}\}$$

где $\varepsilon\{d_{1x}\}$, $\varepsilon\{d_{1y}\}$ – абсолютные доверительные ошибки коэффициентов,

$$\varepsilon\{d_{1x}\} = \frac{S\{Y\}\sqrt{1-r_{yx}^2}}{S\{X\}\sqrt{m-2}} \cdot t_T[p_D = 0{,}95;\ f = m-2] \tag{144}$$
$$\varepsilon\{d_{1y}\} = \frac{S\{X\}\sqrt{1-r_{yx}^2}}{S\{Y\}\sqrt{m-2}} \cdot t_T[p_D = 0{,}95;\ f = m-2]$$

которые определяются по формулам:

Для рассматриваемого примера:

$$\varepsilon\{d_{1x}\} = \frac{4{,}74 \cdot \sqrt{1-0{,}51^2}}{3{,}33 \cdot \sqrt{26-2}} = 0{,}25$$
$$\varepsilon\{d_{1y}\} = \frac{3{,}33 \cdot \sqrt{1-0{,}51^2}}{4{,}74 \cdot \sqrt{26-2}} = 0{,}12$$
$$0{,}73 - 0{,}25 = 0{,}48 \le \delta_x \le 0{,}73 + 0{,}25 = 0{,}98$$
$$0{,}36 - 0{,}12 = 0{,}24 \le \delta_y \le 0{,}36 + 0{,}12 = 0{,}48$$

11.15. Определение условных средних выходных параметров для каждого фактора

Определяют условные средние выходные параметры для каждого значения $\overline{X}_j$ по формуле:

$$\overline{Y}(X) = \overline{Y}(\overline{X}_j) = \frac{1}{m_{xj}} \sum_{i=1}^{q} m_{ji} \overline{Y}_i = \frac{I_y}{m_{xj}} \sum_{i=1}^{q} m_{ji} y_i + Y_0 \tag{145}$$

где $\overline{Y}_i = I_y y_i + Y_0$

Подставляя в формулу (145) значения I_y и Y_0, получают:

Расчет значений $\overline{Y}(X)=\overline{Y}(\overline{X}_j)$ приведен в табл.24. По полученным значениям строят эмпирические ломанные линии $\overline{Y}(X)$, характеризующие изменения условных средних выходного параметра (см. рис. 12).

$$\overline{Y}(X)=\frac{4{,}4}{m_{xj}}\sum_{i=1}^{q} m_{ji}y_i+172$$

11.16. Определение дисперсии расчетного значения выходного параметра для фиксированного значения фактора

Определяют дисперсии расчетного значения выходного параметра для фиксированного значения фактора $\overline{X}_j$ по формуле:

$$S_m^2\{Y_R(X_j)\}=S^2\{d_{0x}\}+S^2\{d_{1x}\}\cdot(\overline{X}_j-\overline{X})^2 \qquad (146)$$

где $$S^2\{d_{0x}\}=S^2\{Y\}\frac{(1-r_{yx}^2)}{m-2}$$

$$S^2\{d_{1x}\}=\frac{S^2\{Y\}\cdot(1-r_{yx}^2)}{S^2\{X\}\cdot(m-2)}$$

Подставляя в формулу (127) значения $S^2\{d_{0x}\}$ и $S^2\{d_{1x}\}$, получают:

$$S^2\{d_{0x}\}=22{,}46\frac{(1-0{,}51^2)}{26-2}=0{,}69$$

$$S^2\{d_{0x}\}=\frac{22{,}46\cdot(1-0{,}51^2)}{11{,}07\cdot(26-2)}=\frac{0{,}69}{11{,}07}$$

$$S_m^2\{Y_R(\overline{X}_j)\}=0{,}69\left[1+\frac{(\overline{X}_j-\overline{X})^2}{11{,}07}\right]$$

Расчет дисперсий и средних квадратических отклонений выходного параметра приведен в табл.24.

11.17. Определение доверительных интервалов средних значений выходного параметра при фиксированном значении фактора

Доверительные интервалы истинного среднего значения выходного параметра для фиксированного значения фактора определяют следующим неравенством:

$$Y_{mR}^{(u)}(X)=Y_{mR}^{(u)}\left(\overline{X}_j\right)=Y_R\left(\overline{X}_j\right)-\varepsilon_m\left\{Y_{Rj}\right\}\le\eta_j\le Y_R\left(\overline{X}_j\right)+\varepsilon_m\left\{Y_{Rj}\right\}=Y_{mR}^{(0)}\left(\overline{X}_j\right)=Y_{mR}^{(0)}(X)$$

где абсолютная доверительная ошибка расчетного значения выходного параметра определяется:

$$\varepsilon_m\left\{Y_{Rj}\right\}=S_m\left\{Y_R\left(X_j\right)\right\}\cdot t_T\left[p_D=0{,}95;\ f=m-2\right] \tag{147}$$

Расчет доверительных интервалов средних значений выходного параметра приведен в табл.24. На рис.12 изображены графики $Y_{mR}^{(u)}(X)$ и $Y_{mR}^{(0)}(X)$, между которыми с вероятностью 0,95 должны располагаться точки, соответствующие условным средним $\overline{Y}\left(\overline{X}_j\right)$.

11.18. Определение доверительных интервалов для индивидуальных значений выходного параметра при каждом уровне фактора

Доверительные интервалы для индивидуальных значений выходного параметра при фиксированном значении фактора $\overline{X}_j$ определяют следующим неравенством:

$$Y_{eR}^{(u)}(X)=Y_{eR}^{(u)}\left(\overline{X}_j\right)=Y_R\left(\overline{X}_j\right)-\varepsilon_e\left\{Y_{Rj}\right\}\le\eta_j\le Y_R\left(\overline{X}_j\right)+\varepsilon_e\left\{Y_{Rj}\right\}=Y_{eR}^{(0)}\left(\overline{X}_j\right)=Y_{eR}^{(0)}(X)$$

где абсолютная доверительная ошибка единичного значения выходного параметра по отдельным экспериментальным значениям определяется:

$$\varepsilon_e\left\{Y_{Rj}\right\}=S_e\left\{Y_R\left(\overline{X}_j\right)\right\}\cdot t_T\left[p_D=0{,}95;\ f=m-2\right] \tag{148}$$

$$S_e^2\left\{Y_R\left(\overline{X}_j\right)\right\}=S^2\left\{d_{0x}\right\}\left[1+m+\frac{\left(\overline{X}_j-\overline{X}\right)^2}{S^2\{X\}}\right] \tag{149}$$

Подставляя в формулу (130) значения $S^2\{d_{0x}\}$, $\overline{X}$ и $S^2\{X\}$, получают:

$$S_e^2\left\{Y_R\left(\overline{X}_j\right)\right\}=0{,}69\cdot\left[1+26+\frac{\left(\overline{X}_j-96{,}93\right)^2}{11{,}07}\right]$$

Расчет доверительных интервалов для индивидуальных значений выходного параметра приведен в табл.24. На рис.12 изображены графики $Y_{eR}^{(u)}(X)$ и $Y_{eR}^{(0)}(X)$, между которыми с вероятностью 0,95 должны располагаться точки, соответствующие отдельным значениям выходного параметра Y_j.

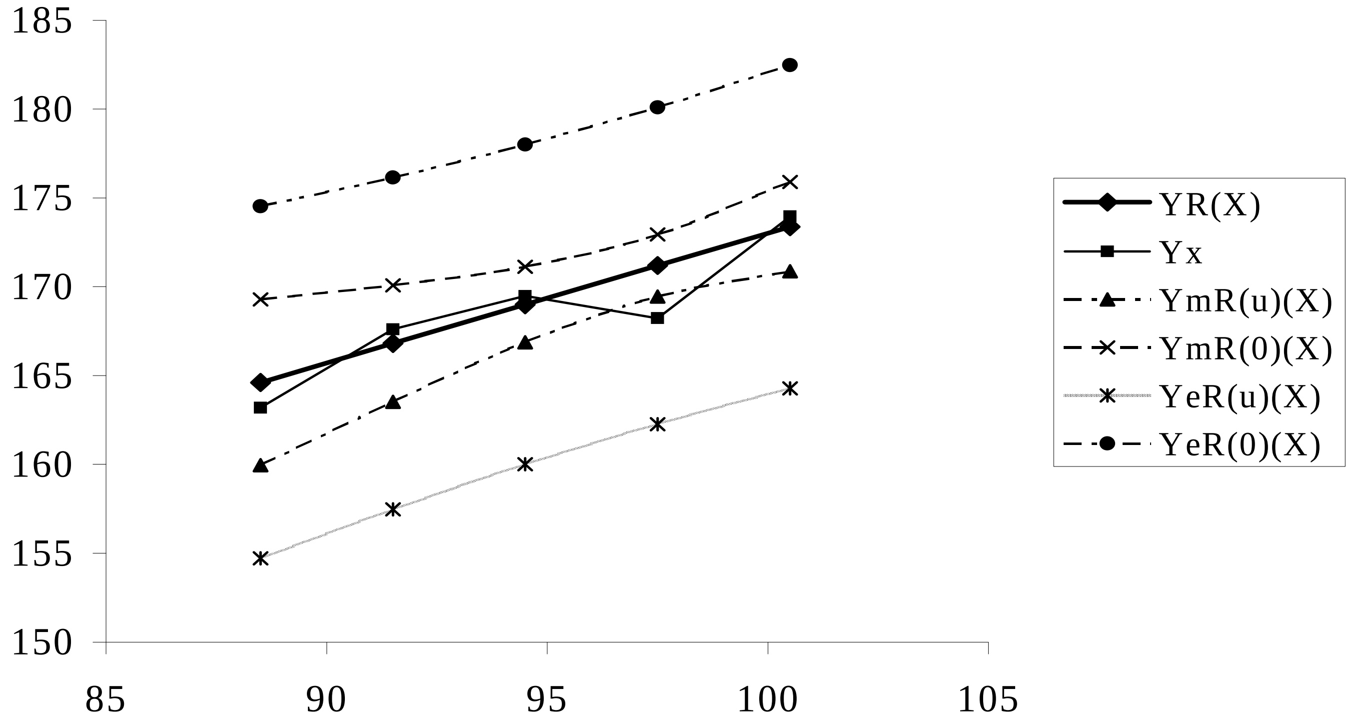


Рис.12. Графики линейной корреляционной однофакторной модели и ее доверительных интервалов

ПРИЛОЖЕНИЯ

Приложение 1

Критические значения V_T критерия исключения резко выделяющихся данных выборки

Повторности	$P_Д$		
m	**0,99**	**0,95**	**0,90**
3	1,414	1,412	1,406J
4	1,723	1,689	1,645
5	1,955	1,869	1.791
6	2,130	1,996	1,894
7	2,265	2,093	1,974
8	2,374	2,172	2,041
9	2,464	2,237	2,097
10	2,540	2,294	2,146
11	2,606	2,343	2,190
12	2,663	2,387	2,229
13	2,714	2,426	2,264
14	2,759	2,461	2,297
15	2,800	2,493	2,326
16	2,837	2,523	2,354
17	2,871	2,551	2,380
18	2,903	2,577	2,404
19	2,932	2,600	2,426
20	2,959	2,623	2,447
21	2,984	2,644	2,467
22	3,008	2,664	2.486
23	3,030	2,683	2,504
24	3,051	2,701	2,502
25	3,071	2,717	2,537

Значения $t_Т$ критерия Стьюдента $t_Т$ [$Р_Д$; f]

f	$Р_Д$				
	Двусторонний критерий				
	0,8	**0,9**	**0,95**	**0,99**	**0,999**
1	3,078	6,314	12,706	63,657	636,62
2	1,886	2,920	4,303	9,925	31,598
3	1,638	2,353	3,182	5,841	12,924
4	1,533	2,132	2,776	4,604	8,610
5	1,476	2,015	2,571	4,032	6,869
6	1,440	1,943	2,447	3,707	5,959
7	1,415	1,895	2,365	3,499	5,408
8	1,397	1,860	2,306	3,355	5,041
9	1,383	1,833	2,262	3,250	4,781
10	1,372	1,812	2,228	3,169	4,587
11	1,363	1,796	2,201	3,106	4,437
12	1,356	1,782	2,179	3,055	4,318
13	1,350	1,771	2,160	3,012	4,221
14	1,345	1,761	2,145	2,977	4,140
15	1,341	1,753	2,131	2,947	4,073
16	1,337	1,746	2,120	2,921	4,015
17	1,333	1,740	2,110	2,898	3,965
18	1,330	1,734	2,101	2,878	3,922
19	1,328	1,729	2,093	2,861	3,883
20	1,325	1,725	2,086	2,845	3,850
21	1,323	1,721	2,080	2,831	3,819
22	1,321	1,717	2,074	2,819	3,792
23	1,319	1,714	2,069	2,807	3,767
24	1,318	1,711	2,064	2,797	3,745
25	1,316	1,708	2,060	2,787	3,725
26	1,315	1,706	2,056	2,779	3,707
27	1,314	1,703	2,052	2,771	3,690
28	1,313	1,701	2,048	2,763	3,674
29	1,311	1,699	2,045	2,756	3,659
30	1,310	1,697	2,042	2,750	3,646
40	1,303	1,684	2,021	2,704	3,551
60	1,296	1,671	2,000	2,660	3,460
120	1,289	1,658	1,980	2,617	3,373
∞	1,282	1,645	1,960	2,576	3,291
$Р_Д$	**0,90**	**0,95**	**0,975**	**0,995**	**0,9995**
	Односторонний критерий				

Приложение 3

Таблица значений F_T критерия Фишера F_T [$P_Д$=0,95, f_2; f_1]

(f_2 — степень свободы для большей дисперсии,

f_1 - степень свободы для меньшей дисперсии)

$f_1 \backslash f_2$	1	2	3	4	5	6	7	8	9
1	161,4	199,5	215,7	224,6	230,2	234,0	236,8	238,9	240,5
2	18,51	19.00	19,16	19,25	19,30	19,33	19,35	19.37	19,38
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77
В	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68
8	5,32	4,16	4,07	3,84	3,69	3,58	3,50	3,44	3,39
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02
11	4,81	4,38	3,59	3,36	3,20	3,09	3.01	2,95	2,90
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28
26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,31	2,25
28	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24
29	4,18	3,33	2,93	2,70	2,55	2,43	2,35	2,28	2,22
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04
120	3,922	3,07	2,68	2,45	2,29	2,17	2,09	2,02	1,96
∞	3,84	3,00	2,60	2,37	2,21	2,10	2,01	1,94	1,88

Окончание приложения 3

$f_1 \backslash f_2$	10	12	15	20	24	*30*	40	60	120	∞
1	241,9	243,9	245,9	248,0	249,1	250,1	251,1	252,2	253,3	254,3
2	19,40	19,41	19,43	19,45	19,45	19,46	19,47	19,48	19,49	19,50
3	8,79	8,74	8,70	8,66	8,64	8,62	8,59	8,57	8,55	8,53
4	5,96	5,91	5,86	5,80	5,77	5,75	5,72	5,69	5,66	5,63
5	4,74	4,68	4,62	4,56	4,53	4,50	4,48	4,43	4,40	4,36
6	4,06	4,00	3,94	3,87	3,84	3,81	3,77	3,74	3,70	3,67
7	3,64	3,57	3,51	3,44	3,41	3,38	3,34	3,30	3,27	3,23
8	3,35	3,28	3,22	3,15	3,12	3,08	3,04	3,01	2,97	2,93
9	3,14	3,07	3,01	2,94	2,90	2,86	2,83	2,79	2,75	2,71
10	2,98	2,91	2,85	2,77	2,74	2,70	2,66	2,62	2,58	2,54
11	2,85	2,79	2,72	2,65	2,61	2,57	2,53	2,49	2,45	2.40
12	2,75	2,69	2,62	2,54	2,51	2,47	2,43	2,38	2,34	2,30
13	2,67	2,60	2,53	2,46	2,42	2,38	2,34	2,30	2,225	2,21
14	2,60	2,53	2,46	2,39	2,35	2,31	2,27	2,22	2,18	2,13
15	2,54	2,48	2,40	2,33	2,29	2,25	2,20	2,16	112,	2,07
16	2,49	2,42	2,35	2,28	2,24	2,19	2,15	2,11	2,06	2,01
17	2,45	2,38	2.31	2,23	2,19	2.15	2,10	2,06	2,01	1,96
18	2,41	2,34	2,27	2,19	2,15	2,11	2,06	2,02	1,97	1,92
19	2,38	2,31	2,23	2,16	2,11	2,07	2,03	1,98	1,93	1,88
20	2,35	2,28	2,20	2,12	2,08	2,04	1,99	1,95	1,90	1,84
21	2,32	2,25	2,18	2,10	2,05	2,01	1,96	1,92	1,87	1,81
22	2,30	2,23	2,15	2,07	2,03	1,98	1,94	1,89	1,84	1,78
23	2,27	2,20	2,13	2,05	2,01	1,96	1,91	1,86	1,81	1.76
24	2,25	2,18	2,11	2,03	1,98	1,94	1,89	1,84	1,79	1,73
25	2,24	2,16	2,09	2,01	1,96	1,92	1,87	1,82	1,77	1,69
26	2,22	2,15	2,07	1,99	1,95	1,90	1.85	1,80	1,75	1,67
27	2,20	2,13	2,06	1,97	1.93	1,88	1,84	1,79	1,73	1,65
28	2,19	2.12	2,04	1,96	1,91	1,87	1,82	1,77	1,71	1,64
29	2,18	2JO	2,03	1,94	1,90	1,85	1,81	1,75	1,70	1,62
30	2,16	2,09	2,01	1,93	1,89	1,84	1,79	1,74	1,68	1,51
40	2,08	2,00	1,92	1,84	1,79	1,74	1,69	1,64	1,58	1,39
60	1,99	1,92	1,84	1,75	1,70	1,65	1,59	1,53	1,47	1,25
120	1,91	1,83	1,75	1,66	1,61	1,55	1,50	1,43	1,35	1,00
∞	1,83	1,75	1,67	1,57	1,52	1,46	1,39	1,32	1,22	1,71

Приложение 4

Табличные значения критерия Кочрена, т. е. отношения наибольшей эмпирической дисперсии к сумме *N* эмпирических дисперсий G_T [p_D=0,95; *f* = *m*— 1, N]

N	***f***													
	Доверительная вероятность 0,95													
	1	2	3	4	5	6	7	8	9	10	16	36	144	∞
2	0,9985	0,9750	0,9392	0,9057	0,8772	0,8534	0.8332	0,8159	0,8010	0,7880	0,7341	0,5602	0,5813	0,5000
3	0,9669	0,8709	0,7977	0,7457	0,7071	0,6771	0,6530	0,6333	0,6167	0,6025	0,5466	0,4748	0,4031	0,3333
4	0,9065	0,7679	0,6841	0,6287	0,5859	0,5598	0,5365	0,5175	0,5017	0,4884	0,4366	0,3720	0,3093	0,2500
5	0,8412	0,6838	0,5981	0,5441	0,5065	0,4783	0,4564	0,4387	0,4241	0,4118	0,3645	0,3066	0,2513	0,2000
6	0,7808	0,6161	0,5321	0,4803	0,4447	0,4184	0,3980	0,3817	0,3682	0,3568	0,3135	0,2612	0,2119	0,1667
7	0,7271	0,5612	0,4800	0,4307	0,3974	0,3726	0,3535	0,3384	0,3259	0,3154	0,2756	0,2278	0,1833	0,1429
8	0,6798	0,51.57	0,4377	0,3910	0,3595	0,3362	0,3185	0,3043	0,2926	0,2829	0,2462	0.2022	0,1616	0,1250
9	0,6385	0,4775	0,4027	0,3584	0,3286	0,3067	0,2901	0,2768	0,2659	0,2568	0,2226	0,1820	0,1446	0,1111
10	0,6020	0,4450	0,3733	0,3311	0,3029	0,2823	0,2666	0,254!	0,2439	0,2353	0,2032	0,1655	0,1308	0,1000
12	0,5410	0,3924	0,3264	0,2880	0,2624	0,2439	0,2299	0,2187	0,2098	0,2020	0,1737	0,1403	0,1100	0,0833
15	0,4709	0,3346	0,2758	0,2419	0,2195	0,2034	0,1911	0,1815	0,1736	0,1671	0,1429	0,1144	0,0889	0,0667
20	0,3894	0,2705	0,2205	0,1921	0,1735	0,1602	0,1501	0,1422	0,1357	0,1303	0,1108	0,0879	0,0675	0,0500
24	0,3434	0,2354	0,1907	0,1656	0,1493	0,1374	0,1286	0,1216	0,1160	0,1113	0,0094	0,0743	0,0567	0,0417
30	0,2929	0,1980	0,1593	0,1377	0,1237	0,1137	0,1061	0,1002	0,0958	0,0921	0,0771	0,0604	0,0457	0,0333
40	0,2370	0,1576	0,1259	0,1082	0,0968	0,0887	0,0827	0,0780	0,0745	0,0713	0,0595	0,0462	0,0347	0,0250
60	0,1737	0,1131	0,0895	0,0765	0,0682	0,0623	0,0583	0,0552	0,0520	0,0497	0,0411	0,0316	0,0234	0,0167
120	0,0998	0,0632	0,0495	0,0419	0,0371	0,0337	0,0312	0,0292	0,0279	0,0266	0,0218	0,0165	0,0120	0,0083
∞	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Приложение 5

Коэффициенты q_{m-i+1}, используемые при проверке экспериментальных данных на нормальность с помощью критерия *W*, для *m*=3...50

i	*m*							
	3	4	5	6	7	8	9	10
1	0,7071	0,6872	0,6646	0,6431	0,6233	0,6052	0,5888	0,5739
2		0,1677	0,2413	0,2806	0,3031	0,3164	0,3244	0,3291
3				0,0875	0,1401	0,1743	0,1976	0,2141
4						0,0561	0,0947	0,1224
5								0,0399

i	*m*							
	11	12	13	14	15	16	17	18
1	0,5601	0,5475	0,5359	0,5251	0,5150	0,5056	0,4968	0,4886
2	0,3315	0,3325	0,3325	0,3318	0,3306	0,3290	0,3273	0,3253
3	0,2260	0,2347	0,2412	0,2460	0,2495	0,2521	0,2540	0,2553
4	0,1429	0,1586	0,1707	0,1802	0,1878	0,1939	0,1988	0,2027
5	0,0695	0,0922	0,1099	0,1240	0,1353	0,1447	0,1524	0,1587
6		0,0303	0,0539	0,0727	0,0880	0,1005	0,1109	0,1197
7				0,0240	0,0433	0,0593	0,0725	0,0837
8						0,0196	0,0359	0,0496
9								0,0163

i	*m*							
	19	20	21	22	23	24	25	26
1	0,4808	0,4734	0,4643	0,4590	0,4542	0,4493	0,4450	0,4407
2	0,3232	0,3211	0,3185	0,3156	0,3126	0,3098	0,3069	0,3043
3	0,2561	0,2565	0,2578	0,2571	0,2563	0,2554	0,2543	0,2533
4	0,2059	0,2085	0,2119	0,2131	0,2139	0,2145	0,2148	0,2151
5	0,1641	0,1686	0,1736	0,1764	0,1787	0,1807	0,1822	0,1836
6	0,1271	0,1334	0,1399	0,1443	0,1480	0,1512	0,1539	0,1563
7	0,0932	0,1013	0,1092	0,1150	0,1201	0,1245	0,1283	0,1316

Приложение 6

Критические значения критерия *W*, используемого для проверки экспериментальных данных на нормальность, для *m* = *3* ... 50

m	P_D				
	0,99	0,98	0,95	0,90	0.50
3	0,753	0,756	0,767	0,789	0,959
4	0,687	0,707	0,748	0,792	0,935
5	0,686	0,715	0,762	0,806	0,927
6	0,713	0,743	0,788	0,826	0,927
7	0,730	0,760	0,803	0,838	0,928
8	0,749	0,778	0,818	0,851	0,932
9	0,764	0,791	0,829	0,859	0,935
10	0,781	0,806	0,842	0,869	0,938
11	0,792	0,817	0,850	0,876	0,940
12	0,805	0,828	0,859	0,883	0,943
13	0,814	0,837	0,866	0,889	0,945
14	0,825	0,846	0,874	0,895	0,947
15	0,835	0,855	0,881	0,901	0,950
16	0,844	0,863	0,887	0,906	0,952
17	0,851	0,869	0,892	0,910	0,954
18	0,858	0,874	0,897	0,914	0,956
19	0,863	0,879	0,901	0,917	0,957
20	0,868	0,884	0,905	0,920	0,959
21	0,873	0,888	0,908	0,923	0,960
22	0,878	0,892	0,911	0,926	0,961
23	0,881	0,895	0,914	0,928	0,962
24	0,884	0,898	0,916	0,930	0,963
25	0,888	0,901	0,918	0,931	0,964
26	0,891	0,904	0,920	0,933	0,965
27	0,894	0,906	0,923	0,935	0,965
28	0,896	0,908	0,924	0,936	0,966
29	0,898	0,910	0,926	0,937	0,966
30	0,900	0,912	0,927	0,939	0,967
31	0,902	0,914	0,929	0,940	0,967
32	0,904	0,915	0,930	0,941	0,968
33	0,906	0,917	0,931	0,942	0,968
34	0,908	0,919	0,933	0,943	0,969
35	0,910	0,920	0,934	0,944	0,969
36	0,912	0,922	0,935	0,945	0,970
37	0,914	0,924	0,936	0,946	0,970
38	0,916	0,925	0,938	0,947	0,971

Окончание приложения 6

m	P_D				
	0,99	0,98	0,95	0,90	0.50
39	0,917	0,927	0,939	0,948	0,971
40	0,919	0,928	0,940	0,949	0,972
41	0,920	0,929	0,941	0,950	0,972
42	0,922	0,930	0,942	0,951	0,972
43	0,923	0,932	0,943	0,951	0,973
44	0,924	0,933	0,944	0,952	0,973
45	0,926	0,934	0,945	0,953	0,973
46	0,927	0,935	0,945	0,953	0,974
47	0,928	0,936	0,946	0,954	0,974
48	0,929	0,937	0,947	0,954	0,974
49	0,929	0,937	0,947	0,955	0,974
50	0,930	0,938	0,947	0,955	0,974

Приложение 7

Таблица случайных чисел

5489	5583	3156	0835	1988	3912	0938	7460	0869	4420
3522	0935	7877	5665	7020	9555	7379	7124	7878	5544
7555	7579	2550	2487	9477	0864	2349	1012	8250	2633
5759	3554	5080	9074	7001	6249	3224	6368	9102	2672
6303	6895	3371	3196	7231	2918	7380	0438	7547	2644
7351	5634	5323	2623	7803	8374	2191	0464	0696	9529
7068	7803	8832	5119	6350	0120	5026	3684	5657	0304
3613	1428	1796	8447	0503	5654	3254	7336	9536	1944
5143	4534	2105	0368	7890	2473	4240	8652	9435	1422
9815	5144	7649	8638	6137	8070	5345	4865	2456	5708
5780	1277	6316	1013	2867	9938	3930	3203	5696	1769
1187	0951	5991	5245	5700	5564	7352	0891	6249	6568
.1184	2179	4554	9083	2254	2435	2965	5154	1209	7069
291G	2972	9885	0275	0144	8034	8122	3213	7666	0230
5524	1341	9860	6565	6981	9842	0171	2284	2707	3008
0146	5291	2354	5694	0377	5336	6460	9585	3415	2358
4920	2826	5238	5402	7937	1993	4332	2327	6875	5230
7978	1947	6380	3425	7267	7285	1130	7722	0164	8573
7453	0653	3645	7497	5969	8682	4191	2976	0361	9334
1473	6938	4899	5348	1641	3652	0852	5296	4538	4456

8162	8797	8000	4707	1880	9660	8446	1883	9768	0881
5645	4219	0807	3301	4279	4168	4305	9937	3120	5547
2042	1192	1175	8851	6432	4635	5757	6656	1660	5389
5470	7702	6958	9080	5925	8519	0127	9233	2452	7341
4045	1730	6005	1704	0345	3275	4738	4862	2556	8333
5880	1257	6163	4439	7276	6353	6912	0731	9033	5294
9083	4260	5277	4998	4298	5204	3965	4028	8936	4148
1762	8713	1189	1090	8989	7273	3213	1935	9321	4820
2023	2589	1740	0424	8924	0005	1636	1636	7237	1227
7965	3855	4765	0703	1678	0841	7543	0308	9732	1289
7690	0480	8098	9629	4819	7219	7241	5128	3853	1921
9292	0426 .	9573	4903	5916	6576	8368	3270	6641	0033
0867	1656	7016	4220	2533	6345	8227	1904	5138	2537
0505	2127	8255	5276	2233	3956	4118	8199	6380	6340
6295	9795	1112	5761	2575	6837	3336	9322	7403	8345
6323	2615	3410	3365	1117	2417	3176	2434	5240	5455
8672	8536	2966	5773	5412	8114	0930	4697	6919	4569
1422	5507	7596	0670	3013	1354	3886	3268	9469	2584
2653	1472	5113	5739	1469	9545	9331	5303	9914	6394
0438	4376	3328	8645	8327	0110	4549	7955	5275	2890
2851	2157	0057	7085	1129	0460	6821	8323	2572	8962
7962	2753	3077	8718	7418	8004	1425	3706	8822	1494
3837	4098	0220	1217	4732	0150	1637	1097	1040	7372
8542	4126	9274	2251	0607	4301	8730	7690	6235	3477
0439	0765	8039	9484	2577	7859	1976	0623	1418	6685
6687	1943	4307	0579	8171	8224	8641	7034	3595	3875
6242	5582	5872	3197	4919	2792	5991	4058	9769	1918
6859	9606	0522	4993	0345	8958	1289	8825	6941	7685
0590	1932	6013	3623	1973	4112	1795	8465	2110	8045
3482	0478	0221	6738	7323	5643	4767	0106	2272	9862

СОДЕРЖАНИЕ

Ирина Валентиновна Абакумова,
доцент кафедры конструирования и технологии одежды АмГУ, канд.техн.наук

Методы и средства мсследования технологических процессов:
Учебное пособие

Изд-во АмГУ. Подписано к печати 02.12.10. Формат 60x84/16. Усл. печ. л. 6,51. Тираж 100. Заказ 179.