Л.С. Ушаков
С.А. Рябчук
Ю.Е. Котылев

АКТИВНЫЙ ФАКТОРНЫЙ ЭКСПЕРИМЕНТ

ГТУ ОРЕЛ
ниверситет

МИНИСТЕРСТВО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
ОРЛОВСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ
Факультет новых технологий и автоматизации производства

Л.С. Ушаков, С.А. Рябчук, Ю.Е. Котылев

АКТИВНЫЙ ФАКТОРНЫЙ ЭКСПЕРИМЕНТ
Математическое планирование, организация и статистический анализ результатов

Рекомендовано Учёным советом ОрелГТУ
в качестве учебного пособия

Орел 2002

УДК 519.658.5.012.+519.24/27
У 93

Рецензенты:
доктор технических наук, профессор
В.П. Росляков
доктор технических наук, профессор Самарского института
инженеров железнодорожного транспорта
О.П. Мулюкин

Ушаков Л.С., Рябчук С.А., Котылев Ю.Е.
Активный факторный эксперимент. Математическое планирование, организация и статистический анализ результатов: Уч. пособие. – ОрелГТУ, 2002 – 39 с.

Рекомендуется студентам, обучающимся по специальности 071100 «Динамика и прочность машин», для изучения дисциплин ОПД.Ф.08 «Организация и управление», СД.Ф.11 «Экспериментальная механика», ФТД.00.4 «Организация и планирование НИР и ОКР», а также для студентов и аспирантов машиностроительных и других специальностей.

В пособии рассматриваются постановка опытов, вопросы планирования испытаний и статистической обработки полученных данных для полного факторного эксперимента, при котором все уровни одного фактора комбинируются со всеми уровнями остальных факторов. Уровни варьирований чаще всего находятся в граничных точках интервала варьирования. Опыты рекомендуется ставить согласно матрице планирования.

В основу анализа результатов эксперимента положен регрессионный анализ и статистические методы оценки адекватности теоретических решений.

Учебное пособие предназначено для изучения современных методов научных исследований с целью повышения качества подготовки высококвалифицированных инженеров и оказания помощи аспирантам различных специальностей в планировании и анализе результатов физического, технологического или социологического эксперимента.

УДК 519.658.5.012 +519.24/27

Содержание

Введение

Планирование эксперимента достигло в последние годы значительных успехов и получило широкое распространение. Большой практический опыт, отечественный и зарубежный, свидетельствует об огромных экономических возможностях, заложенных в рационально спланированном активном эксперименте. Однако материалы по методам планирования эксперимента и обработки результатов, имеющиеся в литературе, представлены разрозненно, носят несистематический характер. Наряду с этим, по тем или иным причинам в каждодневной практике ряда отечественных исследователей планирование эксперимента и обработка результатов наблюдений являются лишь демонстрацией возможностей. В сложившейся ситуации авторы сочли целесообразным написать это пособие.

Цель учебного пособия – систематически и доступно изложить теорию математического планирования активного факторного эксперимента и статистической обработки его результатов, адресуя это изложение, главным образом, студентам.

1. Объекты и методы исследований

Задачи инженерных исследований могут решаться *аналитически и экспериментально*. Эти два подхода взаимно дополняют друг друга.[1]

Аналитический путь требует предварительного изучения объекта исследований, создания теории протекающих процессов, схематизации объекта, процесса и описания его в виде функциональных уравнений (дифференциальных). Этот путь длинный для сложных объектов (процессов), но им нельзя пренебрегать.

Сложный объект исследований, характеризующийся большим числом состояний, в котором он может находиться, описывается большим количеством взаимосвязанных переменных, часть из которых не поддаётся контролю. Объект может быть *управляемым*, если любое состояние его может задаваться и стабилизироваться исследователем, *неуправляемым*, если это условие не выполняется, или *частично управляемым*.

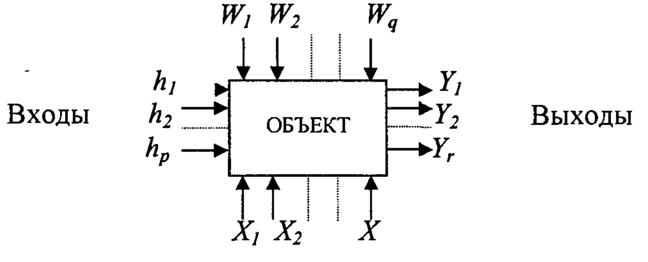

Рисунок 1. Модель (функциональная схема) объекта исследований.

Состояние объекта характеризуется функцией $Y = f(x_1, x_2, x_n, h_1, h_2, h_p, w_1, w_2, w_q)$, называемой функцией отклика, в которой значения независимых величин являются факторами, а параметры y_1, y_2, y_r, которые *изучаются* (или оптимизируются) – *выходами*, или *откликами*. Среди факторов различают следующие группы.

1. *Контролируемые неуправляемые факторы* $\underline{H^T}$:

$$\underline{H^T} = \| \, h_1, h_2, ..., h_p \, \| \, . \qquad (1)$$

Они контролируются, но не поддаются целенаправленному изменению исследователем (паспортные характеристики оборудования, частота, гранулометрический состав, и др.).

2. *Контролируемые управляемые* факторы $\underline{X}^T$ могут изменяться исследователем:

$$\underline{X}^T = \| X_1, X_2, ..., X_n \|. \qquad (2)$$

3. *Контролируемые возмущающие* факторы $\underline{W}^T$:

$$\underline{W}^T = \| W_1, W_2, ..., W_q \| \qquad (3)$$

являются возмущающими воздействиями на объект (например, изменение условий среды, наличие или отсутствие прослоек, примесей и др., а также все факторы, не включённые в программу исследований). Часто информация об их характере, точке приложения, интенсивности отсутствует, а иногда носит случайный характер.

Управляемые факторы должны быть независимыми величинами (некоррелируемыми). *Контролируемые* факторы (управляемые и неуправляемые) могут быть оценены *количественно* или *качественно*, если такой оценки нет. Точки факторов $\underline{H}^T$ и $\underline{X}^T$ находятся в *факторном пространстве*. Часть факторного пространства, доступная для исследования, составляет *область действия*.

Объекты могут быть *статическими* и *динамическими*. *Статическим* объект считается, если его состояние в любой момент времени не зависит от предыдущего состояния, а определяется комбинацией уровней действующих факторов. *Динамическим* – если такая зависимость имеется.

Поведение объекта исследования описывается математической зависимостью отклика от факторов:

$$y = \varphi\,(\underline{H}, \underline{X}, \underline{W}). \qquad (4)$$

Зависимость $y = \varphi\,(\underline{H} \cdot \underline{X})$ называется *функцией отклика*, зависимость $y^I = \delta\,(\underline{W})$ называют *шумом* на выходе, если изменение этих факторов носит стационарный характер, или *дрейфом*, если они изменяются во времени. Часто шум и дрейф наблюдаются одновременно.

Экспериментальные исследования могут проводиться в виде *пассивного эксперимента*, являющегося, по существу, наблюдением, при котором регистрируются произвольно изменяющиеся показатели состояния объекта. Исследователь ставит опыты произвольным образом, интуитивно. *Пассивный эксперимент* является единственно возможным при изучении действующего промышленного объекта. Однако он не позволяет получить достаточной информации об объекте вследствие недостаточного количества параллельных опытов; взаимной корреляции факторов; невозможности сделать раздельную оценку их влияния на отклик; узких интервалов варьирования факторов (так как эксперимент ведётся в рабочем режиме). Поэтому основное внимание следует уделять *активному эксперименту*, при котором программируется изменение воздействия всех управляемых факторов на объект.

Активный эксперимент связан с выбором специальных входных сигналов (факторов), при этом контролируется вход и выход исследуемой системы.

В отличие от пассивного в активном эксперименте:

- имеется возможность проведения параллельных экспериментов, оценки рассеивания наблюдения, проверки адекватности гипотезы;
- планируются изменения уровней управляемых факторов;
- реализуется так называемый рациональный план эксперимента (математическое планирование).

Наряду с этим, в активном эксперименте неуправляемые факторы не учитываются, а планирование по управляемым факторам требует жёсткой стабилизации неуправляемых, что трудно выполнить в промышленных условиях.

В целях сокращения времени и средств на проведение эксперимента, повышения производительности труда исследователя любой эксперимент должен быть спланирован. От исследователя требуется представление об элементах математической статистики и планировании эксперимента.

Планирование эксперимента – это процедура выбора числа и условий проведения опытов, необходимых и достаточных для решения поставленной задачи с достаточной точностью. Суть планирования эксперимента заключается в постановке опытов по некоторой заранее составленной схеме, обладающей какими-то оптимальными свойствами (получение большего объема информации от объекта при меньших затратах).

Простейший метод проведения эксперимента – это испытания при различных сочетаниях факторов (*способ перебора*); однако уже при трёхфакторном эксперименте для получения полной картины необходимо проведение большого числа испытаний. Поэтому часто при решении экспериментальных задач используют следующий способ: изменяя только один фактор (остальные остаются постоянными), находят частный экстремум по этому фактору; затем последовательно определяют следующие частные экстремумы путём изменения второго, третьего и т.д. факторов. Процедура повторяется до тех пор, пока не будет достигнут оптимум по всем факторам. Такая схема значительно уменьшает число опытов, но она не оптимальна и может привести к получению локальных экстремумов.

Разработанные [2-6,7,8] методы математического планирования эксперимента позволяют быстро и экономно приблизиться к поставленной цели эксперимента и упростить статистическую обработку результатов измерений. Известны и широко используются, например, *матричное* планирование эксперимента, планирование *греко-латинскими квадратами*, *квадратами Юдина*, *эволюционное* планирование, планирование способом *3-х монет*. В отличие от традиционных пассивных статистических методов математическое планирование эксперимента является активной процедурой, которая определяет довольно жёсткую схему проведения испытаний и анализ полученных данных. Использование методов математического планирования достигает наибольшего эффекта тогда, когда задача поставлена корректно, а исходные данные выбраны рационально, что зависит, прежде всего, от квалификации и опыта экспериментатора.

Методы планирования эксперимента в настоящее время быстро развиваются, чему способствует возможность широкого использования ЭВМ.

2. Активный факторный эксперимент

Рациональное планирование эксперимента позволяет варьировать все факторы сразу в отличие от традиционного подхода, в котором исследователь изучает действие каждого фактора в отдельности. В эксперименте можно выделить три важных этапа :

- осмысление и подготовка к эксперименту;
- математическое планирование;
- проведение эксперимента и анализ его результатов.

Осмысление и подготовка эксперимента требует высокого уровня априорной информации и предусматривает комплекс мероприятий:

1. Краткое описание объекта и процесса.
2. Изучение условий и результатов, достигнутых при исследовании аналогичных объектов (процессов).
3. Формулировка цели и задач исследований. Если задач несколько, их ранжируют по степени важности.
4. Уточнение перечня факторов. Если при подготовке исследований «забыт» один или несколько существенных факторов, то результаты эксперимента могут оказаться бесполезными или привести к неправильным выводам. Поэтому рекомендуется составить список всех факторов, которые могут влиять на выход процесса. Поскольку таких факторов может быть очень много, то это приводит к неоправданному увеличению объёма и усложнению экспериментальных работ. Возникает необходимость отобрать наиболее важные определяющие факторы. Способы отсеивания слабо влияющих факторов основаны на априорном отборе посредством их ранжирования по степени влияния на выход. Список рекомендуется оформить в виде табл. 1.

Таблица 1

Список существенных факторов в проводимом эксперименте

Номер фактора	Наименование фактора	Обозначение	Размерность	Уровни варьирования факторов			Точность	Шаг варьирования
				Верхний	Нулевой	Нижний		
X_1 X_2 X_3 ... X_n				$X_{i\,max}$	$X_{i\,0}$	$X_{i\,min}$		ΔX_i

5. Выделение факторов, включаемых в реальный эксперимент.

6. Определение возможности фиксации факторов на выбранном уровне и сохранения заданного значения в ходе эксперимента.

7. Выбор условного нулевого уровня фактора и точек, симметричных относительно нулевого уровня. Он может находиться внутри области наилучших условий, на границе или вне её. При наличии опыта эксплуатации объекта за нулевой уровень принимается значение факторов, при которых получились наилучшие отклики.

8. Выбор шага интервала варьирования уровней ΔX_i.

Выбору шага интервала варьирования факторов ΔX_i следует уделять особое внимание. С одной стороны, ΔX_i должно быть достаточно большим, чтобы изменение выходного параметра было больше значения погрешности эксперимента. С другой стороны, ΔX_i должен быть достаточно мал, чтобы полученная информация более строго давала представление о поведении отклика в окрестности нулевой точки. На практике ΔX_i часто принимают равным удвоенной среднеквадратической ошибке определения (измерения) данного параметра $2S_x$, а иногда – равным 0,1 диапазона варьирования.

9. После определения интервалов необходимо установить число уровней (значений) для каждого фактора. Оно должно быть не менее 2, максимальное число ограничено числом опытов. Задача экспериментатора – выдать минимально возможное число уровней для достижения цели. Затем натуральные значения факторов переводят в кодированные. Так, в случае планирования эксперимента на двух уровнях верхний уровень обозначается +1, нижний – 1 (или просто + и –), на трёх уровнях вводится ещё кодированное значение «нулевой» точки – 0 (то есть +1, 0, –1 или +, 0, –).

10. Определение взаимодействия факторов и оценка критических значений факторов (или их комбинаций), приводящих к остановке работы объекта (процесса).

11. Составление списка откликов и точности их определения.

12. Обдумывание желаемых результатов и их оценка, какой из результатов будет считаться «Отл.», «Хор.», «Удов.» или «Неудов.». Информацию по п. 11 и п. 12 целесообразно свести в табл. 2.

Таблица 2

Список откликов в проводимом эксперименте.

Номер опыта	Наименование отклика	Обозначение	Размерность	Область определения	Точность измерений	Примечание

13.Определение необходимого числа опытов. Для установления зависимости между исследуемыми переменными параметрами (откликами) число опытов должно составлять:

$$N = U^{K_\phi}, \qquad (5)$$

где K_ϕ – число факторов,

U – число уровней факторов.

2.1. Математическое планирование активного факторного эксперимента

Математическое планирование эксперимента предусматривает:

1. Определение необходимого числа наблюдений.
2. Обдумывание и описание плана проведения эксперимента и метода получения достоверных данных.
3. Оценку возможности проведения параллельных измерений.
4. Составление матрицы планирования K-факторного эксперимента.
5. Выбор желаемой стратегии проведения экспериментов (например, по 1 в день и т.д.).
6. Выбор математической модели для описания эксперимента вида:

$$Y = f\,(x, f, h...).$$

Сущность планирования эксперимента состоит в том, что опыты ставят по определённой схеме, одновременно варьируя все независимые переменные [2, 3, 6, 8, 9, 10].

План эксперимента желательно составить так, чтобы получить *максимум информации при минимальных затратах средств и времени.* При постановке любого исследования перед экспериментатором возникает проблема выбора: с одной стороны, для получения достоверных результатов необходимо провести большое число опытов, с другой – экономия времени требует проведения минимально возможного количества испытаний. Решить эту проблему можно, основываясь на гипотезе о нормальном законе распределения результатов измерений, разброс которых обусловлен неизбежной погрешностью выбранного экспериментального метода и неоднородностью свойств изучаемого объекта [9, 11].

Для унификации плана проведения эксперимента введём кодированные значения факторов Z_i в скользящей системе координат, начало которой совпадает с нулевым уровнем факторов. Выберем масштабы переменных таким образом, чтобы интервалы варьирования факторов были равны единице. Тогда исходному уровню фактора X_0 будет соответствовать кодированное значение Z_i=0, а верхний и нижний уровни кодированных значений фактора, полученные сложением и вычитанием нулевого уровня и шага, будут соответственно равны +1 и –1.

Переход от натуральных значений X_i факторов к безразмерным (кодированным) Z_i осуществляется по формуле:

$$Z_i = \frac{2(X_i - X_{i\max})}{X_{i\max} - X_{i\min}} + 1, \qquad (6)$$

где $X_{i\,max}$, $X_{i\,min}$ – граничные значения интервала;

X_i – натуральные значения i-го фактора внутри интервала варьирования.

2.2. План проведения 2-факторного 2-уровневого эксперимента

План проведения опытов в кодированном масштабе записывается в виде *матрицы* планирования эксперимента, зависящей только от числа факторов и уровней каждого фактора. Для двух факторов, на двух уровнях каждый, матрица планирования представлена в табл. 3.

Таблица 3

Матрица планирования 2-факторного 2-уровневого эксперимента

Номер вариантов (№)	Значение (вспомогательное) исходного уровня	Планирование уровней факторов		Расчёт	Выход
•	Z_0	Z_1	Z_2	Z_1Z_2	Y
1	+1	–1	–1	+1	y_1
2	+1	+1	–1	–1	y_2
3	+1	–1	+1	–1	y_3
4	+1	+1	+1	+1	y_4

Первая строка соответствует опыту 1, в котором оба фактора находятся на нижнем уровне; вторая – опыту 2, когда первый фактор Z_1 находится на верхнем уровне, а второй Z_2 – на нижнем, и т.д. В расчётном столбце 5 (Расчет) приведены значения произведения Z_1Z_2, которые вводятся в матрицу для последующего вычисления коэффициента регрессии b_{12}; в последнем столбце 6 (Выход) приведены средние значения результатов измерений; во втором столбце – значения исходного уровня, вводимые формально для расчёта коэффициента b_0. Матрица составлена с учётом всех возможных комбинаций значений факторов в данном эксперименте.

2.3. План проведения 3-факторного 2-уровневого эксперимента

Матрица планирования для трёх факторов, варьируемых на двух уровнях, получается из матрицы для двух факторов, повторением последней при значении Z_3 на нижнем и верхнем уровнях (табл. 4).

Таблица 4

Матрица планирования 3-факторного 2-уровневого эксперимента

Номер варианта №	Значение исходного уровня	Планирование уровней факторов			Расчёт				Выход					
	Z_0	Z_1	Z_2	Z_3	Z_1Z_2	Z_1Z_3	Z_2Z_3	$Z_1Z_2Z_3$	y^I	y^{II}	y^{III}	...	y^n	y_n
1	+1	–1	–1	–1	+1	+1	+1	–1						
2	+1	+1	–1	–1	–1	–1	+1	+1						
3	+1	-1	+1	–1	–1	+1	–1	+1						
4	+1	+1	+1	–1	+1	–1	–1	–1						
5	+1	–1	–1	+1	+1	–1	–1	+1						
6	+1	+1	–1	+1	–1	+1	–1	–1						
7	+1	–1	+1	+1	–1	–1	+1	–1						
8	+1	+1	+1	+1	+1	+1	+1	+1						
9	+1	0	0	0	–	–	–	–						

Аналогично составляются матрицы для четырёх, пяти и более факторов; при этом кроме столбцов планирования обычно вводятся значения произведений Z_1Z_i, $Z_1Z_2Z_i$, и т.д. Матрица планирования, записанная в кодированных обозначениях, всегда имеет один и тот же вид для любой задачи с одинаковым числом факторов [12, 13, 14].

2.4. Организация и проведение активного факторного эксперимента

По завершении планирования эксперимента в соответствии с методикой его проведения собрать опытную установку, проверить и отрегулировать (приборы) и датчики, подготовить материалы, составить журнал наблюдений. В нём записывают цель исследований, параметры, подлежащие оптимизации, параметры, характеризующие процесс, факторы и таблица их уровней, интервал варьирования, план эксперимента (матрица планирования). В плане (матрице) лучше проставить кроме кодовых значений факторов их натуральные значения. Здесь же в плане (матрице) нужно записать дату проведения эксперимента и фамилию проводящего эксперимент.

В ходе проведения экспериментов и анализа его результатов проводится:

- измерение и обработка данных;

- вычисление статистических данных для проверки гипотез (дисперсионный анализ) и оценка корреляции параметров;
- интерпретация результатов эксперимента (регрессионный анализ), получение эмпирической математической модели и определение коэффициентов уравнения регрессии.

В журнале наблюдений фиксируются результаты измерений и статистического анализа результатов эксперимента; для оценки корреляционной зависимости между откликами подсчитываются коэффициенты парной корреляции; определяются коэффициенты уравнения регрессии и их значимость. Для облегчения проведения статистической обработки и регрессионного анализа удобно использовать таблицы, приведённые в тексте.

Эксперименты обычно ставятся небольшими сериями по заранее согласованному плану. После каждой небольшой серии опытов производится обработка результатов наблюдений и принимается строго обоснованное решение о том, что делать дальше. При последовательном проведении эксперимент выполняется поэтапно, с тем чтобы результаты каждого этапа анализировать и принимать решение о целесообразности проведения или о направлении дальнейших исследований. В итоге статистической обработки результатов всех этапов эксперимента формулируются выводы и соотношения, продумываются доказательства и подготавливаются иллюстрации.

3. Оценка и анализ результатов испытаний

Полученные при проведении эксперимента значения величин $y_1, y_2, ..., y_n$ заносятся в табл. 5. Для осуществления всех статистических проверок и определения возможности описания факторной зависимости линейными моделями найдём средние арифметические значения измеренного параметра и результатов всех опытов $\bar{y}$, дисперсию S_y^2, среднеквадратическое отклонение S_y, коэффициент вариации $k_{вар}$, коэффициент корреляции и доверительный интервал во всех группах массива результатов каждой серии опытов [9, 15, 16]. Среднее арифметическое значение параметра из n наблюдений (измерений) определяется по формуле:

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \qquad (7)$$

где n – количество наблюдений при одинаковых условиях проведения опыта;

y_i – численное значение i-го наблюдения (i=1, 2, 3, ..., n).

Таблица 5.

Результаты измерений и оценка количественного разброса параметров эксперимента.

<table>
<tr><td rowspan="3">№ опытов п/п</td><td colspan="12">Выход фактора 1.....n</td></tr>
<tr><td colspan="5">Результаты наблюдений</td><td rowspan="2">$\bar{y}_i = \frac{\sum_1^n y_i}{n}$</td><td rowspan="2">$\Delta y_i=(\bar{y}_i - y_i)$</td><td rowspan="2">$\Delta y_i^2$</td><td rowspan="2">$S^2(y)$</td><td rowspan="2">$S(y)$</td><td rowspan="2">$k_{вар}$ %</td><td rowspan="2">y</td></tr>
<tr><td>1</td><td>2</td><td>3</td><td>...</td><td>n</td></tr>
<tr><td>1</td><td>y_{11}</td><td>y_{12}</td><td>y_{13}</td><td>......</td><td>y_{1n}</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>2</td><td>y_{21}</td><td>y_{22}</td><td>y_{23}</td><td>......</td><td>y_{2n}</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>⋮</td><td>⋮</td><td>⋮</td><td>⋮</td><td>⋮</td><td>⋮</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>N</td><td>y_{N1}</td><td>y_{N2}</td><td>y_{N3}</td><td>......</td><td>y_{Nn}</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
</table>

Погрешность отдельных измерений составляет:

$$\Delta y_i = (\bar{y} - y_i). \qquad (8)$$

Проверить, нет ли промахов в измерениях, можно, рассчитав значение выборочной дисперсии результатов измерений:

$$S_{\bar{y}}^2 = \frac{\sum_{i=1}^{n} (\Delta y_i)^2}{n-1}. \qquad (9)$$

Поскольку вычисление значения выборочного среднеквадратического отклонения S_x (погрешность результатов серии измерений) по формулам, приведённым в различных источниках, достаточно трудоёмко, то при сравнительно небольшом числе испытаний (n<30) без больших погрешностей значение S_y можно рассчитать с помощью формулы:

$$S_{\bar{y}} = \frac{1}{\sqrt{n-1}} \sum_{i=1}^{n} (\bar{y} - y_i). \qquad (10)$$

Для этого необходимо провести предварительные эксперименты с меньшим числом измерений n, чтобы оценить степень разброса результатов эксперимента.

Обработка осциллограмм проводится методом наложения абсолютного значения отклонения луча осциллографа на тарировочный график измеряемого параметра [11, 12, 13, 14]. При их обработке каждый цикл по характерным точкам разбивается на фазы. Расстояние от нулевой линии каждого параметра до соответствующей кривой на осциллограмме измеряется с точностью до 0,5 мм. Результаты эксперимента рекомендуется представлять в виде таблиц (табл. 5). Табличная форма представления опытных данных позволяет в компактном виде записать большое количество данных и облегчает статистический и регрессионный анализ. При составлении таблиц необходимо обратить внимание на значащие цифры, так как они несут информацию о точности измеряемой величины.

Необходимое число повторных опытов n и достоверное значение измеряемой величины определяются на основе методов математической статистики [18, 19]. Результаты отдельных измерений носят случайный характер, поэтому степень точности полученных экспериментальных данных при ограниченном числе измерений n может быть оценена путём отыскания погрешностей при установлении средних значений параметров для построения графических зависимостей.

В использованных методах математической статистики, согласно принципу наименьших квадратов, точность искомой величины оценивается допустимой ошибкой [19, 20]

$$k_{доп} = t_j \frac{S_y}{\bar{y}\sqrt{n}}, \qquad (11)$$

где t_j – функция распределения Стьюдента, зависящая от условий эксперимента и задаваемой надёжности α (Приложение 1).

Требуемая точность измерений устанавливается в соответствии с поставленными задачами эксперимента и не должна превышать точности прибора. Показатель точности $k_{доп}$ для общего машиностроения может быть принят в пределах 5…10% ($k_{доп} \leq 0{,}1 \ldots 0{,}05$).

Приведённые формулы можно преобразовать к виду, удобному для решения задачи об установлении наименьшего количества измерений, которые обеспечивают получение среднего арифметического $\bar{x}$ с заданной точностью $k_{доп}$ при определённой вероятности α. При заданной величине ошибки необходимое число повторных опытов определяется выражением:

$$n = (t_j \frac{k_{вар}}{k_{доп}})^2, \tag{12}$$

где $k_{вар}$ – коэффициент вариации, $k_{вар} = \frac{S_y}{\bar{y}} 100\%$. (13)

Эта формула даёт возможность определить необходимое количество наблюдений, зависящее от требуемой вероятности α и числа предварительных испытаний. При малой выборке ($n \leq 30$) в зависимости от принятой доверительной вероятности $P_д$ и числа членов ряда n по таблице выбирается значение коэффициента Стьюдента и определяется погрешность результата измерения.

За доверительный интервал (погрешность) результата измерения y_i принимается интервал, равный:

$$\bar{y} \pm t_j s_{\bar{y}}, \tag{14}$$

в который попадают истинные значения y_i с заданной надёжностью α.

Для получения надёжности правильных выходов, равной 0,95, что вполне достаточно для исследовательских работ по рекомендации проф. Л. И. Барона [19], нормированное отклонение средней величины от интегральной средней t_i принимается равным 1,96. Тогда можно определить число измерений, необходимое для достижения заданной точности искомой величины при одинаковых условиях проведения эксперимента. То есть, для каждых условий проведения опыта необходимо получать не менее n осциллограмм.

Рассмотрим на числовом примере применение вышеприведенных формул и влияние числа измерений на точность результата.

Проведем обработку результатов измерений диаметра штока гидроцилиндра для значений надежности $\alpha = 0,95$ и $\alpha = 0,99$. Измерения проведены мерительным инструментом с ценой деления 0,01 мм.

Значения измеренных диаметров штока.

n	d_i, мм	n	d_i, мм
1	14, 85	6	14, 81
2	14, 80	7	14, 80
3	14, 84	8	14, 85
4	14, 81	9	14, 84
5	14, 79	10	14, 80

Рассмотрим пять первых измерений из этой таблицы. Найдем среднее арифметическое значение диаметра из 5 измерений:

$$\bar{d} = \frac{1}{n}\sum_{i=1}^{n} d_i = \frac{1}{5}(14{,}85 + 14{,}80 + 14{,}84 + 14{,}81 + 14{,}79) = 14{,}818 мм.$$

Вычислим разности $\bar{d} - d_i$ и квадраты этих разностей, результаты вычислений представим таблицей. Значение среднего квадрата погрешности (дисперсии) найдем по формуле (9).

Оценка количественного разброса результатов измерений.

n	d_i, мм	$\bar{d} - d_i$ мм	$(\bar{d} - d_i)^2$ мм^2	$S_{\bar{d}}^2$, мм^2	$S_{\bar{d}}$, мм	$k_{вар}$ %
1	14, 85	– 0, 032	0, 001024			
2	14, 80	+ 0,018	0, 000324			
3	14, 84	– 0,022	0, 000484	$2{,}43 \cdot 10^{-3}$	$4{,}93 \cdot 10^{-2}$	0,333
4	14, 81	+ 0,008	0, 000064			
5	14, 74	+ 0,028	0, 000784			
	сумма	0	0,009736			

Для надежности $\alpha = 0,95$ и $n = 5$ из табл. приложения 1 находим значение коэффициента Стьюдента t_α=2,78 и вычисляем абсолютную погрешность результата измерений по формуле:

$$\Delta d_{\bar{\alpha}} = t_\alpha S_{\bar{d}} = \pm 2{,}78 \cdot 4{,}93 \cdot 10^{-2} = 0{,}137 мм.$$

Тогда доверительный интервал результата измерения диаметра штока с учетом формулы (14) можно представить в виде:

$$d_i = \bar{d} \pm \Delta d_{\bar{d}},$$

$$d_i = (14,818 \pm 0,137) \text{ мм},$$

или $$14,681 \text{ мм} \le d_i \le 14,955 \text{ мм}. \qquad (14а)$$

Найдем погрешность и доверительный интервал для тех же пяти измерений при другом значении надежности $\alpha = 0,9$. Из табл. приложения 1 находим для $n = 5$ и $\alpha = 0,99$ значение коэффициента $t_\alpha = 4,60$, тогда

$$\Delta d_{\bar{d}} = 4,60 \cdot 4,93 \cdot 10^{-2} = 2,267 \cdot 10^{-1} \text{ мм},$$

или $d_i = (14,818 \pm 0,227)$ мм. (14б)

Сравнив результаты (14а) и (14б), мы видим, что границы доверительного интервала при увеличении надежности от $\alpha = 0,95$ до $\alpha = 0,99$ возросли.

Найдем теперь погрешность результата всей серии из десяти измерений. В этом случае мы получим результат, приведенный в таблице.

Оценка количественного разброса десяти измерений.

n	d_i мм	$\bar{d} - d_i$ мм	$(\bar{d} - d_i)^2$ мм	$S_{\bar{d}}^2$	$S_{\bar{d}}$	$k_{вар}$ %
1	14,85	–0,031	0,000961	$5,433 \cdot 10^{-4}$ мм	$2,33 \cdot 10^{-2}$ мм	0,157
2	14,80	0,019	0,000361			
3	14,84	–0,021	0,000441			
4	14,81	0,009	0,000081			
5	14,79	0,029	0,000841			
6	14,81	0,009	0,000081			
7	14,80	0,019	0,000361			
8	14,85	–0,031	0,000961			
9	14,84	–0,021	0,000441			
10	14,80	0,019	0,000361			
	Сумма		0,00489			

Среднее арифметическое значение диаметра штока из 10 измерений равно

$$\bar{d} = \frac{1}{n}\sum_{1}^{n} d_i = 14,819 \text{ мм}$$

При $\alpha = 0,95$ и $i = 10$ из табл. приложения 1 находим значение $t_\alpha = 2,26$.

Абсолютная погрешность всей серии из десяти измерений диаметра штока составит

$$\Delta d_{\bar{d}} = 2{,}26 \cdot 2{,}33 \cdot 10^{-2} = \pm 0{,}053 мм,$$

а доверительный интервал результата измерений в этом случае –

$$d_i = (14{,}819 \pm 0{,}053) \text{ мм}$$

Мы видим, что погрешность результата десяти измерений диаметра штока стала более чем в два раза меньше погрешностей пяти измерений.

3.1. Регрессионный анализ результатов испытаний

Суть регрессионного анализа сводится к установлению уравнения регрессии, т.е. вида кривой, характеризующей зависимость между случайными величинами x и y, оценке тесноты связей между ними, достоверности и адекватности результатов измерений [14, 20, 21, 22]. Даже для вполне выраженной формы корреляционной связи вследствие статистического характера связи исследуемого явления одно значение x может иметь несколько значений y. Если на корреляционном поле осреднить точки, т.е. для каждого значения x_i определить $\bar{y}_i$ и соединить их, то можно получить ломаную линию, называемую *экспериментальной* регрессионной зависимостью (линией). Наличие ломаной линии объясняется погрешностями измерений, недостаточным количеством измерений, физической сущностью исследуемого явления и др. Если на корреляционном поле провести плавную линию между $\bar{y}_i$, которая равноудалена от них, то получится *новая теоретическая* зависимость.

Различают однофакторные (парные) и многофакторные регрессионные зависимости. Парная регрессия в исследуемых процессах может быть реализована одним из следующих видов уравнений связи:

- линейным (зависимость в виде прямой, не проходящей через начало координат):

$$y = ax+b; \qquad (15)$$

- параболическим (зависимость в виде параболы второй степени):

$$y = ax^2+bx+c; \qquad (16)$$

- показательным:

$$y = ab^x; \qquad (17)$$

- степенным:

$$y = ax^b; \qquad (18)$$

• логарифмическим:

$$y = a+b\cdot lgx. \qquad (19)$$

Двухфакторное поле можно аппроксимировать плоскостью, параболоидом второго порядка, гиперболоидом и др. Для переменных факторов связь может быть установлена с помощью n-мерного пространства уравнениями второго порядка:

$$y = b_0 + \sum_1^n b_i x_i + \sum_1^n b_{ij} x_i x_j + \sum_1^n b_{ii} x_i^2 ,$$

где y – функция цели (отклика) многофакторных переменных;

x_i – независимые факторы;

b_i – коэффициент регрессии, характеризующий влияние фактора x_i на отклик;

b_{ij} и b_{ii} – коэффициент, характеризующий двойное влияние факторов x_i и x_j на отклик.

Задача заключается в определении тесноты и степени надёжности связи между коррелируемыми параметрами, т.е. в определении степени влияния искажающих факторов. Они сводятся к определению коэффициента корреляции, или корреляционного отношения, как показателям тесноты связи. Критерием близости корреляционной зависимости между x и y к линейной зависимости является коэффициент парной корреляции, определяемый по формуле [21]

$$r = \frac{\overline{xy} - \overline{x} \cdot \overline{y}}{S_x \cdot S_y}, \qquad (20)$$

где $\overline{xy}$ – среднее значение произведения $\overline{x \cdot y}$;

$\overline{x}$, $\overline{y}$ – средние значения соответствующих признаков;

S_x, S_y – средние квадратические отклонения, найденные по признаку x и признаку y.

Значение коэффициента корреляции всегда меньше единицы. При тесной функциональной линейной зависимости $r_{x,y} = \pm 1$, т.е. каждому значению x соответствует только одно значение y, причём знак его определяется знаком при коэффициенте линейной функции. Таким образом, если коэффициент $r_{x,y}$ имеет значения

$$0 \le r_{x,y} < 1,$$

то величины x и y находятся в такой линейной зависимости, что при увеличении (уменьшении) одной из них, другая имеет тенденцию к увеличению (уменьшению).

Характер зависимости меняется, если коэффициент $r_{x,y}$ имеет значения

$$-1 < r_{x,y} < 0,$$

в этом случае при увеличении (уменьшении) одной из величин другая уменьшается (увеличивается). При этом большему значению величины $r_{x,y}$ соответствует большая степень связи между исследуемыми величинами, и тем ближе к линии регрессии располагаются экспериментальные точки. При $r = 0$ линейная корреляционная связь между x и y отсутствует, но может существовать нелинейная регрессия. При высокой значимости коэффициента корреляции (т.е. при наличии тесной связи между параметрами) любой из двух анализируемых параметров можно исключить из рассмотрения. Исключать следует тот параметр, который менее ясен, труднее измерить. Математическую модель строят для оставшихся (меньшего числа) отзывов.

Для других видов корреляционных связей теснота корреляционной зависимости определяется корреляционным отношением:

$$r = \frac{S \cdot \bar{y}_x}{S_y}, \tag{21}$$

где $S \cdot \overline{y_x}$ – значение стандарта условных средних $\overline{y_x}$.

Обычно считают тесноту связи удовлетворительной при $r \geq 0{,}5$; хорошей при $r = 0{,}8...0{,}85$. Проверка значимости коэффициентов корреляции проводится сравнением коэффициента Стьюдента, взятого из таблицы [9, 14, 15, 22], с вычисленным по формуле:

$$t_j = \frac{r\sqrt{n-r}}{\sqrt{1-r^2}}. \tag{22}$$

3.2. Математическая модель исследуемого процесса

На практике часто возникает потребность в установлении связи между y и многими факторами x_1, $x_2...x_n$ на основе многофакторной регрессии. Теоретическую модель множественной регрессии можно получить методами математического планирования, т.е. активным экспериментом. При этом предполагается, что существует некоторая аналитическая связь между фактором и откликом процесса и требуется выбрать минимальное число условий проведения опытов, позволяющих найти область оптимальных значений параметров. Другими словами, необходимо найти приближённую зависимость выходного параметра от факторов, т.е. построить математическую модель процесса в виде:

$$y = f(x_1, x_2, \ldots, x_n)\,. \qquad (23)$$

Моделей много и разных. Они характеризуют стохастический процесс изучаемого явления. Главное требование к модели – возможность с её помощью предсказывать результат и выбирать направление дальнейших исследований. Это значит, что предсказанное с помощью модели значение отклика не должно отличаться от фактического больше заранее заданной величины. Такая модель называется *адекватной.*

Математическую модель процесса лучше аппроксимировать полиномами. Например, для двух факторов полином

0 степени $y=b_0;$

I степени $y=b_0+b_1|z_1|+b_2|z_2|;$

II степени $y=b_0+b_1|z_1|+b_2|z_2|+b_{12}|z_1z_2|+b_{11}|z_1^2|+b_{22}|z_2^2|;$

III степени и т.д. ....

Для трех факторов полином

0 степени $y=b_0;$

I степени $y=b_0+b_1|z_1|+b_2|z_2|+b_3|z_3|;$

II степени $y=b_0+b_1|z_1|+b_2|z_2|+b_3|z_3|+b_{12}|z_1z_2|+b_{13}|z_1z_3|+$
$+b_{23}|z_2z_3|+b_{123}|z_1\,z_2z_3|+b_{11}|z_1^2|+b_{22}|z_2^2|+b_{33}|z_3^2|;$

III степени и т.д. ....

В ходе эксперимента получаем статистический ряд значений откликов, необходимых для определения численных значений коэффициентов полинома. Чем больше коэффициентов надо найти, тем больше опытов должно быть проведено. Мы же стремимся их сократить. Для этого следует выбрать одну модель, например, II (или I) степени, определить численные значения коэффициентов, оценить их значимость, понизить степень полинома, отказавшись от незначимых коэффициентов или повысить степени (например, с I на II), если модель неадекватна.

Для большинства реальных задач вид математической модели заранее не известен. Очевидно, точность подобной аппроксимации определяется порядком степенного ряда и диапазоном изменения переменных. Так как отклики изучаются в сравнительно узком интервале варьирования факторов, то без большой погрешности можно отбросить члены высших порядков.

Рассмотрим двухфакторный эксперимент, для которого уравнение регрессии представляет собой полином II степени без членов высших порядков:

$$y=b_0+b_1|z_1|+b_2|z_2|+b_{12}|z_1z_2|, \qquad (24)$$

где $|z_1|$ и $|z_2|$ – кодированные значения факторов;

b_0 – свободный член, равный отклику при $z_1=z_2=0$;

b_1 и b_2 – коэффициенты регрессии, показывающие степень влияния соответствующих факторов на выход;

b_{12} – коэффициент, указывающий на наличие взаимодействия факторов x_1 и x_2.

Если требуется провести трёхфакторный эксперимент, то уравнение регрессии можно представить в виде квазилинейной функции трёх факторов:

$$y=b_0+b_1|z_1|+b_2|z_2|+b_3|z_3|+b_{12}|z_1z_2|+b_{13}|z_1z_3|+b_{23}|z_2z_3|+$$
$$+b_{123}|z_1\,z_2z_3|. \qquad (25)$$

Используя ортогональность матрицы эксперимента, вычисляем коэффициенты уравнения регрессии. Для двухфакторного эксперимента они вычисляются по формулам в соответствии с моделью:

$$b_0=\frac{\sum_1^n \bar{y}_i}{k}; \qquad b_j=\frac{\sum_1^n |z_{ji}|\bar{y}_i}{k}, \qquad (26)$$

где k – число коэффициентов уравнения регрессии.

При подсчёте коэффициентов b_0, b_1 используют знаки соответствующего векторного столбца матрицы Z_1, для b_2 – Z_2 и т.д.

Для двух факторов формулы для подсчёта коэффициентов уравнения регрессии имеют вид:

$$b_0=\frac{(+1)\bar{y}_1+(+1)\bar{y}_2+(+1)\bar{y}_3+\bar{y}_4(+1)}{4},$$
$$b_1=\frac{(-1)\bar{y}_1+(+1)\bar{y}_2+(-1)\bar{y}_3+\bar{y}_4(+1)}{4},$$
$$b_2=\frac{(-1)\bar{y}_1+(-1)\bar{y}_2+(+1)\bar{y}_3+\bar{y}_4(+1)}{4}, \qquad (27)$$
$$b_{12}=\frac{(+1)\bar{y}_1+(-1)\bar{y}_2+(-1)\bar{y}_3+\bar{y}_4(+1)}{4}.$$

Аналогично вычисляются коэффициенты регрессии и для большего числа факторов, например, для трёх.

$$b_0 = \frac{+\bar{y}_1 + \bar{y}_2 + \bar{y}_3 + \bar{y}_4 + \bar{y}_5 + \bar{y}_6 + \bar{y}_7 + \bar{y}_8}{8},$$

$$b_1 = \frac{-\bar{y}_1 + \bar{y}_2 - \bar{y}_3 + \bar{y}_4 - \bar{y}_5 + \bar{y}_6 - \bar{y}_7 + \bar{y}_8}{8},$$

$$b_2 = \frac{-\bar{y}_1 - \bar{y}_2 + \bar{y}_3 + \bar{y}_4 - \bar{y}_5 - \bar{y}_6 + \bar{y}_7 + \bar{y}_8}{8},$$

$$b_3 = \frac{-\bar{y}_1 - \bar{y}_2 - \bar{y}_3 - \bar{y}_4 + \bar{y}_5 + \bar{y}_6 + \bar{y}_7 + \bar{y}_8}{8},$$

$$b_{12} = \frac{+\bar{y}_1 - \bar{y}_2 - \bar{y}_3 + \bar{y}_4 + \bar{y}_5 - \bar{y}_6 - \bar{y}_7 + \bar{y}_8}{8},$$

$$b_{13} = \frac{+\bar{y}_1 - \bar{y}_2 + \bar{y}_3 - \bar{y}_4 - \bar{y}_5 + \bar{y}_6 - \bar{y}_7 + \bar{y}_8}{8},$$

$$b_{23} = \frac{+\bar{y}_1 + \bar{y}_2 - \bar{y}_3 - \bar{y}_4 - \bar{y}_5 - \bar{y}_6 + \bar{y}_7 + \bar{y}_8}{8},$$

$$b_{123} = \frac{-\bar{y}_1 + \bar{y}_2 + \bar{y}_3 - \bar{y}_4 + \bar{y}_5 - \bar{y}_6 - \bar{y}_7 + \bar{y}_8}{8}. \qquad (28)$$

Все расчеты и числовые значения коэффициентов регрессии и парной корреляции целесообразно свести в таблицы.

Таблица 6

Расчёт коэффициента корреляции уравнений регрессии.

№ измерений п/п	x_i	y_i	$x_i y_i$	x_i^2	y_i^2	$x_i - \bar{x}$	$y_i - \bar{y}$	$(x_i - \bar{x})(y_i - \bar{y})$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})^2$
1 2 3 . . . n	x_{11} x_{12} x_{13} . . . x_{1n}	y_1 y_2 y_3 . . . y_n								
Сумма	$\sum_1^n x_{1i}$	$\sum_1^n y_{1i}$	$\sum_1^n x_{1i}$	$\sum_1^n x$	$\sum_1^n y_i^2$					
Среднее значение	$\bar{x}_1$	$\bar{y}_1$	$\overline{x_1 y_i}$	$\bar{x}_1^2$	$\bar{y}_i^2$					

Таблица 7

Числовые значения коэффициентов регрессии и парной корреляции.

Условия проведения эксперимента		Значения коэффициентов регрессии								$\bar{b} = \frac{\sum b_i}{Z}$	$S_{(b_i)} = \frac{\sum(\bar{b} - b_i)}{\sqrt{Z-1}}$	r_2
		b_0	b_1	b_2	b_3	b_{12}	b_{13}	b_{23}	b_{123}			
А	I											
	II											
В	I											
	II											
и т.д.												

Коэффициент регрессии может оказаться незначимым, когда соответствующий фактор не влияет на функцию цели (отклик), так как условный нулевой уровень по данному фактору расположен в оптимальной области или если выбран слишком малый шаг варьирования данного фактора. Поэтому следует провести оценку значимости коэффициентов регрессии при членах высших порядков. Расчеты целесообразно свести в табл. 8.

Для оценки статистической значимости коэффициентов регрессии необходимо найти выборочную дисперсию $S^2(b_i)$ или ошибку $S(b_i)$.

Для вычисления погрешности коэффициентов регрессии, воспользуемся формулой, согласно которой дисперсия для каждой строки определяется выражением [22, 24]:

$$S^2(y_i) = \frac{(\bar{y}_N - y_i{}')^2 + (\bar{y}_N - y_i{}'')^2 + (\bar{y}_N - y_i{}''')^2}{2};$$

$$S^2(y_N) = \frac{(\bar{y}_N - y_N{}')^2 + (\bar{y}_N - y_N{}'')^2 + (\bar{y}_N - y_N{}''')^2}{2}. \quad (29)$$

После этого определяется дисперсия среднего значения

$$S^2(\bar{y}) = \frac{\sum_1^n S^2(y_i)}{N \cdot k_\phi}, \tag{30}$$

где $k_ф$ – число факторов,

N – число опытов,

и дисперсия коэффициентов регрессии

$$S^2(b_i) = \frac{S^2(\bar{y})}{k}, \tag{31}$$

где k– число коэффициентов уравнения регрессии.

Ошибка коэффициентов регрессии составит:

$$S(b_i) = \sqrt{S^2(b_i)}. \tag{32}$$

Число *степеней свободы*, с которым определены коэффициенты регрессии, равно:

$$f_1 = k(k_\phi - 1). \tag{33}$$

Если в эксперименте каждый вариант имел только одну повторность, то дисперсия среднего значения $S^2(\bar{y})$ принимается равной дисперсии метода измерений (найденной, например, из предварительного эксперимента) и тогда:

$$S^2(b_i) = \frac{S^2(\bar{y})}{N}. \tag{34}$$

Таким образом, ошибка коэффициента регрессии $S(b_i)$ в $\sqrt{N}$ раз меньше погрешности метода, что является одним из преимуществ многофакторного эксперимента.

Коэффициент регрессии считается статистически значимым, если его абсолютная величина больше доверительного интервала:

$$|b_i| > S(b_i) \cdot t_j, \tag{35}$$

где $S(b_i) \cdot t_j$ – ошибка коэффициента регрессии b_i;

t_j – коэффициент Стьюдента для заданной вероятности α=0,05 и числа измерений n, который определяется из приложения 1.

Таблица 8

Расчёт значимости коэффициентов регрессии

Условия проведения экспери-мента		Среднеста-тистическое отклонение $S_{(b_j)} = \frac{S(\bar{y}_i)}{\sqrt{Z}}$	Величина полуинтервала разброса коэффициентов $b_j = S_{(b_j)} \cdot t_{j,\alpha}$	Границы до-верительного интервала $[b_j]>b_j$ $[b_j]<b_j$	Значи-мость
А	I				
	II				
В	I				
	II				
и т.д.					

3.3. Оценка адекватности математической модели (проверка статнстических гипотез)

В результате эксперимента получаем статистический ряд однофакторных или многофакторных измерений, которые в последующем подвергаются обработке и анализу, подбираются эмпирические формулы и устанавливается их достоверность. Перед подбором эмпирических формул необходимо убедиться в пригодности гипотезы исследований, а также теоретических данных на адекватность.

В практике оценки адекватности применяют различные статистические критерии согласия [6, 9, 18, 22]. Адекватность математической модели малых выборок чаще всего проверяют с помощью критерия Фишера [9, 12, 13, 17, 18, 23]:

$$F_p = \frac{S^2_{ад}}{S^2_{ср}} \le F_T, \tag{36}$$

где F_P – опытный критерий Фишера,

F_T – теоретическое значение критерия,

$S^2_{ад}$ – дисперсия адекватности,

$S^2_{ср}$ – средняя дисперсия всего эксперимента.

$$S_{ad}^2 = \frac{\sum_{i=1}^{n}(\overline{y}_{эi} - y_{mi})^2}{f_1}, \qquad (37)$$

$$S_{cp}^2 = \frac{\sum_{1}^{n}\sum_{1}^{n}(\overline{y}_{эi} - y_{mi})^2}{f_2}, \qquad (38)$$

где y_{mi} – теоретическое значение параметра, предсказанное полученными уравнениями регрессии;

$\overline{y}_{эi}$ – среднее экспериментальное значение параметра в i-ой точке;

f_1, f_2 – число степеней свободы, с которой определяется диспер сия адекватности и дисперсия опыта:

$$f_1 = (n - k), \qquad (39)$$

$$f_2 = N(n - k); \qquad (40)$$

N – число серий (повторностей) в эксперименте.

Опытное значение критерия Фишера F_p сравнивается с табличным значением [9,25] F_m, зависящим от уровня достоверности α (обычно 0,05) и значений f_1 и f_2 (приложение 2). Если неравенство $F_p \le F_m$ выполняется, то с доверительной вероятностью принятая линейная модель считается адекватной, при $F_p > F_m$ модель отвергается.

Пусть, например, исследуемый процесс описывается теоретическим уравнением y=80x и для его подтверждения проведен эксперимент. В каждой их пяти серий (N=5) выполнено по семь измерений (n=7). Результаты эксперимента приведены в табл. 9. Установим адекватность теоретического выражения по этим данным.

Таблица 9

Оценка адекватности математической модели исследуемого процесса

n	x_i	Измеренные значения $y_{iэ}$ в серии					средние значения $\overline{y}_{iэ} = \frac{\sum_1^n y_{iэ}}{m}$	y_{iT}	$y_{iT} - \overline{y}_{iэ}$	$(y_{iT} - \overline{y}_{iэ})^2$	S_{cpi}^2
		$y_{1э}$	$y_{2э}$	$y_{3э}$	$y_{4э}$	$y_{5э}$					
1	0,2	12	17	15	14	16	14,8	16	1,2	1,44	4,4
2	0,3	23	21	24	25	23	23,2	24	0,8	0,64	2,4
3	0,4	30	34	31	35	35	33,0	32	1,0	1,00	3,8
4	0,5	38	43	40	39	42	40,4	40	0,4	0,16	3,6
5	0,6	52	47	48	49	48	47,2	48	0,8	0,64	16,4
6	0,7	59	58	55	54	53	55,8	56	0,2	0,04	5,4
7	0,8	62	66	62	61	63	62,8	64	1,8	1,44	4,4
Итого										5,36	40,4

С этой целью по формуле (37) определяется $S^2_{ад}$:

$$S^2_{ад} = \frac{5{,}36}{(7-1)} = 0{,}89 .$$

Здесь значение k=1, т.к. в теоретическом уравнении один значащий член. Дисперсия $S^2_{ср}$ вначале вычисляется построчно для всех строк (табл. 9), затем вычисляется средняя дисперсия всего эксперимента:

$$S^2_{ср} = \frac{\sum_1^n S^2_{срi}}{n(n-1)} = \frac{40{,}4}{30} = 1{,}34 .$$

После этого подсчитывается экспериментальное значение критерия Фишера:

$$F_p = \frac{S^2_{ад}}{S^2_{ср}} = \frac{0{,}89}{1{,}34} = 0{,}66 .$$

Теоретическое значение критерия Фишера можно принять по данным приложения 2 при $f_1 = 7-1=6$ и $f_2 = 5(7-1)$:

$$F_T{=}2{,}4.$$

Так как F_p=0,66 < F_T=2,4, модель адекватна, то есть полученная модель с доверительной вероятностью 95% хорошо описывает изучаемый процесс.

При больших выборках для проверки адекватности модели более широко используется критерий согласия Пирсона [9,18]. Значение критерия Пирсона χ^2 вычисляется по формуле:

$$\chi^2 = \sum_1^m \frac{(\bar{y}_{эi} - y_{mi})^2}{y_{mi}}, \tag{41}$$

где $y_{эi}$, y_{mi} – значение параметра в каждой группе серий измерений соответственно по данным эксперимента и теоретической кривой,

$\bar{y}_{эi}$ – среднее значение выборки.

В соответствии с этим критерием гипотеза о законе распределения подтверждается, если соблюдается условие:

$$p(\chi^2_э, q) > \alpha, \tag{42}$$

где α – уровень значимости, принимаемый равным 0,10;

q – число степеней свободы, равное разности числа измерений в одной серии и числа параметров (констант):

$$q = m - S. \tag{43}$$

Значение вероятности закона Пирсона $p(\chi_э^2,\ q)$ определяется по табл. приложения 3. В случае, если модель не адекватна, то повышают степень полинома, описывающего исследуемый процесс.

Пусть, например, произведено N=250 измерений некоторых величин x_i. Экспериментальные данные $y_{эi}$ разбиты на семь групп. По результатам измерений в прямоугольных координатах построена кривая, близкая к закону нормального распределения.

Среднее значение параметров в каждой группе измерений.

$\bar{y}_{эi}$	14,85	14,80	14,84	14,81	14,79	14,81	14,80
y_{mi}	15,65	14,20	14,0	15,6	15,54	13,81	15,70

Проверим адекватность модели, используя критерий Пирсона. Вычислим по формуле (41) критерий согласия $\chi_э^2$:

$$\chi_э^2=0{,}32.$$

Из табл. приложения 3 по значению $\chi_э^2$ и q = 7–2 = 5 находим значение:

$$\chi_T^2 = 0{,}987,$$

$$p(\chi_э^2,\ q)=0{,}987>\alpha=0{,}10.$$

Таким образом, условие формулы (42) удовлетворяется.

Заключение

Учебное пособие является обобщением современного представления теории планирования экспериментов и моделирования сложных технических систем. Пособие отражает литературные данные, материалы научно-исследовательских работ, выполненных в ПНИЛ «Силовые импульсные системы», а также некоторый личный опыт авторов и их коллег.

Пособие восполняет пробел в работах ряда исследователей, где не уделяется достаточного внимания эксперименту. Оно может быть полезно студентам и исследователям, самостоятельно изучающим теорию планирования эксперимента и обработки его результатов. Мы думаем, что приведенных в учебном пособии материалов достаточно, чтобы начать работу, адаптироваться к конкретной ситуации и двигаться в нужном направлении, используя в дальнейшем имеющуюся литературу.

Литература

1. Налимов В.В. Теория эксперимента. М.: Наука,1971.

2. Протодьяконов М.М., Тедер Р.Н. Методика рационального планирования экспериментов. М.: Наука, 1970, – 76 с.

3. Адлер Ю.П. Новые идеи в планировании эксперимента. М.: Наука, 1969, – 331 с.

4. Адлер Ю.П., Грановский Ю.В. Обзор прикладных работ по планированию эксперимента. М.: МГУ, 1972, – 425 с.

5. Адлер Ю.П., Маркова Е.В., Грановский Ю.В. Планирование эксперимента при поиске оптимальных условий. М.: Наука, 1976, – 279 с.

6. Хикс Ч.Р. Основные принципы планирования эксперимента. Перевод с англ. Голиковой Т.Н. и др. – М.: Мир, 1967, – 406 с.

7. Горский В.Г. и др. Планирование промышленных экспериментов. М.: Металлургия, 1978, – 112 с.

8. Адлер Ю.П. Введение в планирование эксперимента. М.: Металлургия, 1968,– 155 с.

9. Леман Э. Проверка статистических гипотез. М.: Наука, 1964.

10. Горбунов В.Ф., Лазуткин А.Г., Ушаков Л.С. Импульсный гидропривод горных машин. – Новосибирск: Наука, 1986. – 197 с.

11. Лазуткин А.Г., Ушаков Л.С., Нордин В.В. Влияние энергетических компонент на эффективность разрушения горных пород ударным инструментом //Изв. вузов. Горный журнал. – 1984. – №7. – С. 62-72.

12. Ушаков Л.С. Научные основы разработки и создания ударно-скалывающих исполнительных органов горнопроходческих машин: Дисс…д.т.н. – М.: МГИ, 1989 – 381 с.

13. Рябчук С.А. Исследование и создание тормозных устройств гидропневмоударных исполнительных органов горных машин.: Дисс…к.т.н.– Караганда.: КарПТИ, 1983. – 230 с.

14. Исследование режимов работы и эффективности тормозной системы /Л.С. Ушаков, Б.С. Кузнецов, С.А. Рябчук и др. //Изв. вузов. Горный журнал. – 1977. – №8 – С. 6-72.

15. Кассандрова О.Н., Лебедев В.В. Обработка результатов наблюдений. –М.: Наука, 1970, – 104 с.

16. Яноши Л. Теория и практика обработки результатов измерений.– М.: Мир, 1968.

17. Экспериментальные исследования разрушения песчаников ударными исполнительными органами /А.Г. Лазуткин, Л.С. Ушаков,

В.В. Нордин и др.; – Караганда: КарПТИ, 1979. – 6 с. – Деп. в ЦНИНучель, 1979, №1471.

18. Длин Л.М. Математическая статистика в технике. – М.: Наука, 1958, – 466 с.

19. Барон Л.И. Горно-технологическое природоведение. – М.: Наука, 1977, – 324 с.

20. Большаков В.Ф. Теория ошибок и наблюдений с основами теории вероятностей. – М.: Наука, 1965, – 184 с.

21. Рыжов П.А. Математическая статистика в горном деле. – М.: Радио, 1965, – 185 с.

22. Шилов П.И. Предельные ошибки результатов измерений и уравновешиваний. – Изв. вузов. «Геодезия и аэрофотосъёмка». 1958., вып. 2.

23. Ушаков Л.С., Котылев Ю.Е., Кравченко В.А. Гидравлические машины ударного действия. – М.: Машиностроение, 2000, 416 с., ил.

24. Зайдель А.Н. Элементарные оценки ошибок измерений. –М.: Наука, 1968, – 97 с.

25. Большев С.В. и др. Таблицы математической статистики. –М.: Статистика, 1967.

Приложение 1

Распределение Стьюдента. Значения $t = t(\alpha; f)$.

α / $f=n-1$	0,90	0,95	0,98	0,99
1	6,31	12,7	31,8	63,7
2	2,92	4,30	6,96	9,92
3	2,35	3,18	4,54	5,84
4	2,13	2,78	3,75	4,60
5	2,02	2,57	3,36	4,03
6	1,94	2,45	3,14	3,71
7	1,89	2,36	3,00	3,50
8	1,86	2,31	2,90	3,36
9	1,83	2,26	2,82	3,25
10	1,81	2,23	2,76	3,17
11	1,796	2,201	2,718	3,106
12	1,78	2,18	2,68	3,05
13	1,771	2,160	2,650	3,012
14	1,76	2,14	2,62	2,98
15	1,753	2,131	2,602	2,947
16	1,746	2, 120	2,583	2,921
18	1,734	2, 103	2,552	2,878
20	1,725	2,086	2,528	2,845
25	1,708	2,060	2,486	2,787
30	1,697	2,042	2,457	2,750
35	1,689	2,030	2,437	2,724
40	1,684	2,021	2,423	2,704
45	1,679	2,014	2,412	2,689
50	1,676	2,008	2,403	2,677
60	1,671	2,000	2,390	2,660
70	1,667	1,995	2,381	2,648
80	1,664	1,990	2,374	2,639
90	1,662	1,987	2,368	2,632
100	1,660	1,984	2,364	2,626
∞	1,645	1,960	2,326	2,576

Приложение 2

Значения критерия Фишера F_T для уровня достоверности α=0,95.

f_2	f_1					
	1	2	3	4	5	6
1	161	200	216	225		234
2	18,5	19,0	19,2	19,2		19,3
3	10,1	9,6	9,3	9,1		9,0
4	7,7	6,9	6,6	6,4	6,3	6,3
5	6,6	5,8	5,4	5,2	5,1	5,0
6	6,0	5,1	4,8	4,5	4,4	4,3
7	5,5	4,7	4,4	4,1	4,0	3,9
8	5,3	4,5	4,1	3,8	3,7	3,6
9	5,1	4,3	3,9	3,6	3,5	3,4
10	5,0	4,1	3,7	3,5	3,3	3,2
11	4,8	4,0	3,6	3,3	3,2	3,1
12	4,8	3,9	3,5	3,	3,1	3,0
13	4,7	3,8	3,4	3,2	3,0	2,9
14	4,6	3,7	3,3	3,1	3,0	2,9
15	4,5	3,7	3,3	3,1	2,9	2,8
16	4,5	3,6	3,2	3,0	2,9	2,7
17	4,5	3,6	3,2	3,0	2,8	2,7
18	4,4	3,6	3,2	2,9	2,8	2,7
19	4,4	3,5	3,1	2,9	2,7	2,6
20	4,4	3,5	3,1	2,9	2,7	2,6
22	4,3	3,4	3,1	2,8	2,7	2,6
24	4,3	3,4	3,0	2,8	2,6	2,5
26	4,2	3,4	3,0	2,7	2,6	2,5
28	4,2	3,3	3,0	2,7	2,6	2,4
30	4,2	3,3	2,9	2,7	2,5	2,4
40	4,1	3,2	2,8	2,6	2,5	2,3

Приложение 3

Критерий Пирсона $p(\chi^2_э, q)$

$\chi^2_э$	Число степеней свободы q							
	1	2	3	4	5	6	7	8
1	0,317	0,606	0,801	0,909	0,962	0,985	0,994	0,998
2	157	367	572	735	849	919	459	981
3	083	223	391	557	700	808	885	934
4	045	135	261	406	549	676	779	857
5	025	082	171	287	415	543	660	757
6	014	049	111	199	306	423	539	647
7	008	030	071	135	220	320	428	536
8	004	018	046	091	156	238	332	433
9	002	011	029	061	109	173	252	343
10	001	006	018	040	075	124	188	265
11	000	004	011	026	051	088	138	201
12		002	007	017	034	062	100	151
13		001	004	011	023	043	072	111
14		000	002	007	014	029	051	081
15			001	004	010	020	036	059
16			001	003	006	013	025	042
17			000	001	004	009	017	030
18				001	002	006	012	021
19				000	001	004	008	014
20					001	002	005	010
21						001	003	007
22						001	002	004
23							001	003
24							001	002
25								001

Учебное пособие

Ушаков Леонид Семенович
Рябчук Семен Александрович
Котылев Юрий Евгеньевич

АКТИВНЫЙ ФАКТОРНЫЙ ЭКСПЕРИМЕНТ
Математическое планирование, организация и статистический анализ результатов

Редактор Е.В. Русанова
Технический редактор А.В. Стебакова

Орловский государственный технический университет
Лицензия ИД № 00670 от 05.01.2000

Подписано к печати 04.04.2002. Формат 60 x 84 1/16.
Печать офсетная. Уч.-изд. л. 2,2. Усл. печ. л. 2,4. Тираж 100 экз.
Заказ №__

Отпечатано с готового оригинал-макета
на полиграфической базе ОрелГТУ,
302030, г. Орел, ул. Московская, 65.