ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ
«ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ - УЧЕБНО-НАУЧНО-ПРОИЗВОДСТВЕННЫЙ КОМПЛЕКС»
ФАКУЛЬТЕТ ПИЩЕВОЙ БИОТЕХНОЛОГИИ
И ТОВАРОВЕДЕНИЯ

Кафедра «Технология хлебопекарного, кондитерского
и макаронного производства»

Н.А. Березина

МОДЕЛИРОВАНИЕ ТЕХНОЛОГИЧЕСКИХ ПРОЦЕССОВ ПРОИЗВОДСТВА ПИЩЕВЫХ ПРОДУКТОВ

Методические указания
по выполнению лабораторных работ

Дисциплина – «Моделирование технологических процессов производства пищевых продуктов»
Направление – 260100.68 «Продукты питания из растительного сырья»

**Допущено ФГБОУ ВПО «Госуниверситет - УНПК»
для использования в учебном процессе в качестве
методических указаний для высшего
профессионального образования**

Орел 2014

Автор: канд. техн. наук, доц. каф. ТХКиМП Н.А. Березина

Рецензент: канд. техн. наук, доц. каф. ТХКиМП О.Л. Ладнова

В методических указаниях рассматриваются методы моделирования технологических процессов производства пищевых продуктов с помощью ЭВМ. При этом, разобраны интерполяционные задачи, а также оптимизационные задачи, решаемые методами активного эксперимента. Материал методических указаний построен так, чтобы все теоретические положения, методики и технические средства тесно увязывались с процессами пищевой и других отраслей промышленности, связанных с хранением и переработкой сырья и получением продуктов питания.

Предназначены студентам, обучающимся по направлению 260100.68 «Продукты питания из растительного сырья», изучающим дисциплину «Моделирование технологических процессов производства пищевых продуктов».

Редактор Г.А. Осипова
Технический редактор А.В. Жукина

Федеральное государственное бюджетное образовательное
учреждение высшего профессионального образования
«Государственный университет - учебно-научно-
производственный комплекс»

Подписано к печати 22.05.2014 г. Формат 60×90 1/16.
Усл. печ. л. 7,5. Тираж 10 экз.
Заказ №________

Отпечатано с готового оригинал-макета
на полиграфической базе ФГБОУ ВПО «Госуниверситет - УНПК»,
302030, г. Орел, ул. Московская, 65.

СОДЕРЖАНИЕ

Введение

Непрерывное увеличение продукции пищевых отраслей промышленности в настоящее время увязывается с повышением качества, биологической ценности и вкусовых достоинств продуктов питания, улучшением их ассортимента. Основными средствами достижения этих целей являются повышение эффективности и качества научных исследований, совершенствование форм связи науки с производством, ускорение внедрения научных достижений в промышленность.

Планируя активный эксперимент, исследователь имеет возможность поддерживать факторы, влияющие на основной показатель производства (производительность, расход энергии, количество продукта и т. п.), на определенных уровнях и менять по заданному закону комбинации этих уровней. При этом могут решаться две группы задач. Первая, основная группа, *оптимизационные задачи*, когда в результате планированного эксперимента определяют значения факторов, совокупность которых приводит к оптимальной величине показателя производства (этот показатель называется параметром оптимизации). Задачи второй группы — *интерполяционные* — являются составной частью оптимизационных задач, но могут иметь и самостоятельное значение. Эти задачи служат для составления математической модели процесса.

К экспериментально-статистическим методам относится также так называемый *пассивный эксперимент*, когда нельзя установить тот или иной фактор на нужном уровне.

Настоящие методические указания посвящены в основном решению оптимизационных задач методами активного эксперимента, но достаточно полно разобраны и интерполяционные задачи как таковые, а также задачи, решаемые методами пассивного эксперимента. Материал методических указаний построен так, что все теоретические положения, методики и технические средства тесно увязывались с процессами пищевой и других отраслей промышленности, связанных с хранением и переработкой продуктов питания.

Лабораторная работа № 1 Однофакторный дисперсионный анализ

Цель работы: 1) ознакомить с методом однофакторного дисперсионного анализа; 2) овладеть практическими навыками анализа технологических процессов методом однофакторного дисперсионного анализа.

В производственной практике часто возникает следующая задача. Аппаратчики, работая по сменам на одном и том же аппарате или агрегате, производят хлебобулочные или кондитерские изделия с различными качественными показателями. Необходимо выяснить, что является причиной появления неудовлетворительных результатов: несовершенная конструкция аппарата или агрегата, не позволяющая добиться хорошей воспроизводимости, или неодинаковая работа аппаратчиков.

Аналогичная задача возникает, например, при выработке одного ассортимента хлебобулочных или кондитерских изделий на нескольких поточных линиях. Требуется установить, однотипны ли получаемые изделия по своим показателям качества, т. е. оценить, существенно ли воздействует фактор индивидуальности оборудования каждой поточной линии на качественные показатели хлебобулочных или кондитерских изделий.

Рассмотренные и аналогичные задачи решают с применением статистических методов дисперсионного анализа (одно- или многофакторный дисперсионный анализ).

Проведение дисперсионного анализа возможно, если результаты наблюдений являются независимыми случайными величинами, подчиняющимися нормальному закону распределения с одинаковыми дисперсиями.

Дисперсионный анализ основан на свойстве аддитивности дисперсии, т. е. на том, что полная дисперсия интересующего показателя равна сумме составляющих ее частных дисперсий.

Задача однофакторного дисперсионного анализа (ОДА) ставится следующим образом. Пусть изучается влияние фактора x на технологический процесс или некоторый показатель качества. В процессе эксперимента фактор поддерживают на u уровнях. На каждом уровне фактора проводится *m* дублирующих (параллельных) опытов. Результаты однофакторного эксперимента из *uxm* наблюдений представляют в виде матрицы наблюдений (табл. 1.1).

Таблица 1.1

Матрица однофакторного эксперимента

Уровень варьирования фактора х	Параллельный опыт					
	1	2	…..	l	…..	m
1	$У_{11}$	$У_{12}$	….	….	….	$У_{1m}$
2	$У_{21}$	$У_{22}$	….	$У_{21}$	….	$У_{2m}$
…..	….	….	….	….	….	
j	$У_{n}$	$У_{2}$	….	$У_{jl}$	….	$У_{m}$
….	….	….	….	….	….	….
u	$У_{u1}$	$У_{u2}$	…	$У_{ul}$	….	$У_{um}$

В табл. 1.1 приняты следующие обозначения: j - порядковый номер уровня варьирования фактора x (j = 1, 2,..., u); l - порядковый номер параллельного опыта в серии на каждом j-ом уровне (l = 1, 2, mj); $У_{jl}$ - значение функции отклика, полученное в 1-ом параллельном опыте на j-ом уровне варьирования фактора.

При расположении наблюдений (см. табл. 1.1) их рассеяние между столбцами обусловливается ошибкой воспроизводимости, а рассеяние между строками - дополнительным действием изучаемого фактора.

Рассеяние отдельных наблюдений относительно общего среднего обусловлено действием как случайных причин, так и влиянием фактора х. Действие фактора случайности проявляется в рассеянии (с оценкой дисперсии S^2_{ε}) наблюдений серий параллельных опытов на каждом уровне x, вокруг среднего арифметического $\overline{y_j}$, своей серии.

Таким образом, имеем равенство

$$S_0^2 = S_\varepsilon^2 + S_x^2 \qquad 1.1$$

где S_0^2 - оценка «общей» дисперсии;

S_ε^2 - оценка остаточной дисперсии;

S_x^2 - оценка дисперсий «между сериями».

При обработке результатов ОДА предварительно определяют:

суммы наблюдений по сериям

$$Y_j = \sum_{l=1}^{m_j} y_{jl} \qquad 1.2$$

сумму квадратов всех М = *um* наблюдений

$$Q_1 = \sum_{j=1}^{u} \sum_{l=1}^{m_j} y_{jl}^2 \qquad 1.3$$

сумму квадратов итогов по сериям, поделенных на число наблюдений в серии

$$Q_2 = \sum_{j=1}^{u} \frac{1}{m_j} Y_j^2 \qquad 1.4$$

квадрат общего итога, поделенный на число всех наблюдений,

$$Q_3 = \frac{1}{M}\left(\sum_{j=1}^{u} Y_j\right)^2 \qquad 1.5$$

Затем определяют «общую» сумму квадратов отклонений отдельных наблюдений y_{jl} от общего среднего

$$S_0 = Q_1 - Q_3 \qquad 1.6$$

Она характеризует рассеяние наблюдений в результате действия обоих факторов, как случайности ε, так и изучаемого *x*.

Сумму квадратов отклонений «внутри серий», т. е. сумму квадратов разностей между отдельными наблюдениями y_{jl} и средним $\overline{y_j}$ соответствующей j-й серии определяют по формуле

$$S_\varepsilon = Q_1 - Q_2 \qquad 1.7$$

Она характеризует остаточное рассеяние случайных погрешностей опытов, т. е. их воспроизводимость.

Сумму квадратов отклонений «между сериями» или рассеивание по уровням фактора *x*, т. е. взвешенную с учетом числа *m* параллельных наблюдений в каждой серии сумму квадратов разностей между средними yj отдельных серий и общим средним по всей совокупности наблюдений определяют по формуле

$$S_x = Q_2 - Q_3 \qquad 1.8$$

Суммы квадратов отклонений (1.6) - (1.8), деленные на соответствующие числа степеней свободы, дают три оценки дисперсий, входящих в выражение (1.1):

оценка общей дисперсии S_0^2 по всем *um* наблюдениям

$$S_0^2 = \frac{S_0}{um-1} \qquad 1.9$$

с числом степеней свободы $f = um-1$;
оценка дисперсии «внутри серий», или оценка остаточной дисперсии S_ε^2, находят как среднее из выборочных дисперсий по каждой серии в отдельности

$$S_\varepsilon^2 = \frac{S_\varepsilon}{u(m-1)} \qquad 1.10$$

с числом степеней свободы $f_\varepsilon = u(m-1)$; оценка дисперсий «между сериями»

$$S_x^2 = \frac{S_x}{u-1} \qquad 1.11$$

с числом степеней свободы $f_x = u-1$.

Числа степеней свободы должны удовлетворять соотношению $f_0 = f_\varepsilon + f_x$, которое используют для проверки.

Выполнение однофакторного дисперсионного анализа заключается в сравнении оценки дисперсии, вызванной изучаемым фактором изменчивости S_x^2, и оценки остаточной дисперсии S_ε^2, имеющей место уже после того, как влияние фактора x было устранено (за счет разложения оценки общей дисперсии S_0^2 на независимые составляющие) и обусловленной исключительно случайными причинами.

Для того чтобы влияние фактора x было признано существенным, необходимо и достаточно, чтобы оценка дисперсии S_x^2 значимо отличалась от S_ε^2. Проверку нуль-гипотезы об однородности этих выборочных дисперсий осуществляют с помощью критерия Фишера:

$$F_p = \frac{S_x^2}{S_\varepsilon^2} \tag{1.12}$$

Влияние фактора x на изучаемый процесс признают существенным, если выполняется условие

$$F_p > F_m \tag{1.13}$$

где F_m - табличное значение критерия Фишера для принятого уровня значимости p и числа степеней свободы $f_1 = f_x = u—1$ и $f_2 = f_\varepsilon = u(m - 1)$ (см. прил. 2).

Если же условие (1.13) не выполняется, то влияние фактора x на изучаемый процесс можно признать несущественным. В этом случае все результаты наблюдений принадлежат одной генеральной совокупности, распределенной нормально.

Следует иметь в виду, что дисперсионный анализ наблюдений эксперимента позволяет оценивать влияние фактора лишь в целом и что выводы, полученные с его помощью, относятся только к данному экспериментальному материалу при данной его систематизации. Так, например, при изменении диапазона варьирования изучаемого фактора или основной (базовой) точки оценка влияния последнего может измениться.

Пример

Основным сырьем макаронного производства является мука из твердой и мягкой стекловидной пшеницы - крупка и полукрупка. К сожалению, из-за недостаточного выращивания зерна пшеницы твердых сортов и, соответственно, недостаточного производства макаронной муки из него, главным образом, из-за экономических факторов допускается применение пшеничной хлебопекарной муки общего помола, а в последнее время пшеничной муки общего назначения для выработки макаронных изделий.

Вместе с тем, на аграрном рынке России в последнее время появилась полукрупка, полученная по сокращенной схеме помола из зерна новых сортов пшеницы Степь 3, Степь 5 и пр., которая так же может быть использованы при производстве макаронных изделий.

В связи с этим возникла необходимость в изучении влияния различных сортов муки на энергоемкость процесса замеса макаронного теста, который является основным технологическим процессом и предопределяет качество готовой продукции.

В качестве объекта исследования использовали макаронную полукрупку, полученную по сокращенной схеме помола зерна из новых сортов пшеницы Степь 3 и Степь 5 в сравнении с макаронной крупкой из твердой пшеницы (контроль 1) и хлебопекарной мукой из мягкой пшеницы высшего сорта (контроль 2) (таблица 1.2).

Таблица 1.2

Обработка результатов однофакторного дисперсионного анализа

Сорт муки	Уровень варьирования фактора j	Значение функции отклика в параллельных опытах		Среднее арифметическое значение функции отклика в параллельных опытах $\overline{y_j}$	Сумма наблюдений по сериям Yj	Квадрат наблюдений в параллельных опытах y_{jl}^2		$\frac{Y_j^2}{m}$
		1	2			1	2	
Контроль 1	1	158,6	159,0	158,8	317,6	25153,96	25281,0	50434,88
Контроль 2	2	123,5	123,7	123,6	247,2	15252,25	15301,69	30553,92
Светлана Степь 3	3	128,1	128,5	128,3	256,6	16409,61	16512,25	32921,78
Линия 3	4	101,9	102,1	102,0	204,0	10383,61	10424,41	20808,0
Линия 10	5	130,9	130,7	130,8	261,6	17134,81	17082,49	34217,28
Линия 37 Степь 5	6	126,2	126,4	126,3	252,6	15926,44	15976,96	31903,38
Линия 13	7	124,6	125,0	124,8	249,6	15525,16	15625,0	31150,08
Линия 21 Урожай 2002 г.	8	132,8	133,0	132,9	265,8	17635,84	17689,0	35324,82
Светлана Степь 3	9	139,6	140,0	139,8	279,6	19488,16	19600,0	39088,08
Линия 3	10	127,0	127,4	127,2	254,4	16129,0	16230,76	32359,68
Линия 10	11	143,0	143,2	143,1	286,2	20449,0	20506,24	40955,22
Линия 37 Степь 5	12	115,0	114,8	114,9	229,8	13225,0	13179,04	26404,02
Линия 13	13	115,0	115,4	115,2	230,4	13225,0	13317,16	26542,08
Линия 21	14	98,6	98,8	98,7	197,4	9721,96	9761,44	19483,38
Сумма					3523,8	225659,8	226487,44	452146,6

В качестве независимой переменной x был принят сорт муки, используемой для замеса теста. В ходе эксперимента фактор x варьировался на u = 14 уровнях. В качестве функции отклика y, характеризующей энергозатраты при замесе теста использовали удельную работу замеса теста (кДж/кг). При каждом уровне варьирования фактора проводили m = 2 параллельных опыта.

Представлены значения удельной работы замеса теста в двух параллельных опытах (m = 2) при всех уровнях (u = 14) варьирования факторов (см. табл. 1.2).

В дальнейшем использованы следующие обозначения: порядковый номер уровня варьирования фактора - j (j = 1, 2,..., u); порядковый номер параллельного опыта в серии на каждом j-ом уровне - l (l = 1, 2, m).

Выполним статистическую обработку результатов эксперимента с помощью однофакторного дисперсионного анализа.

Предварительно по формуле (1.2) рассчитываем сумму наблюдений по сериям Y_j и квадрат наблюдений в параллельных опытах y_{jl} (см. табл. 1.2).

Сумма квадратов всех $M = um$ наблюдений, рассчитанная по формуле (1.3), составляет

$$Q_1 = \sum_{j=1}^{u} \sum_{l=1}^{m_j} y_{jl}^2 = 452147{,}24$$

Сумма квадратов итогов по сериям, поделенных на число наблюдений в серии, определяем по формуле (1.4)

$$Q_2 = \sum_{j=1}^{u} \frac{1}{m_j} Y_j^2 = 452146{,}46$$

значения квадратов сумм наблюдений по сериям, поделенных на число наблюдений в серии, представлены (см. табл. 1.2).

По формуле (1.5) определяем квадрат общего итога, поделенный на число всех наблюдений,

$$Q_3 = \frac{1}{M}\left(\sum_{j=1}^{u} Y_j\right)^2 = \frac{1}{14 \cdot 2}(3523{,}8)^2 = 443470{,}23$$

Далее по формулам (1.6) - (1.8) рассчитываем суммы S_0, S_ε, S_x.

$S_0 = Q_1 - Q_3 = 452147{,}24 - 443470{,}23 = 8677{,}01.$
$S_\varepsilon = Q_1 - Q_2 = 452147{,}24 - 452146{,}6 = 0{,}64.$
$S_x = Q_2 - Q_3 = 452146{,}6 - 443470{,}23 = 8676{,}37.$

Затем по формулам (1.9) - (1.11) определяем оценки дисперсий: общая оценка дисперсии воспроизводимости S_0 по всем um наблюдениям с числом степеней свободы f_0 = um - 1 = 28 -1 = 27; остаточная оценка дисперсии воспроизводимости S_0^2 по всем um наблюдениям

$$S_0^2 = \frac{S_0}{um-1} = \frac{8677{,}01}{28-1} = 321{,}37$$

с числом степеней свободы f_0=um-1=28-1=27

остаточная дисперсия воспроизводимости S_ε^2

$$S_\varepsilon^2 = \frac{S_\varepsilon}{u(m-1)} = \frac{0{,}64}{14(2-1)} = 0{,}045$$

с числом степеней свободы f_ε=u(m-1)=14(2-1)=14

оценка выборочной дисперсии рассеивания «между сериями» S_x^2

$$S_x^2 = \frac{S_x}{u-1} = \frac{8676{,}37}{14-1} = 667{,}41$$

с числом степеней свободы $f_x = u - 1 = 14 - 1 = 13$.

Для того чтобы влияние фактора x было признано существенным, необходимо и достаточно, чтобы оценка дисперсии S_x^2 значимо отличалась от S_ε^2. Проверку нуль-гипотезы об однородности этих выборочных дисперсий проводим с помощью критерия Фишера (1.12)

$$F_p = \frac{S_x^2}{S_\varepsilon^2} = \frac{667{,}41}{0{,}045} = 14831{,}4$$

Табличное значение критерия Фишера (см. прил. 2) для принятого уровня значимости p = 5 % и числа степеней свободы числителя $f_1 = f_x = u - 1 = 14 - 1 = 13$ и знаменателя $f_2 = f = u(m - 1) = 14(2 - 1) = 14$ составляет F_m =244,5.

Проверка условия (1.13) показала, что $F_p > F_m$, следовательно, при доверительной вероятности 95 % влияние фактора x (вид муки) на энергоемкость процесса замеса макаронного теста (удельную работу замеса теста) следует признать существенным.

Задание

В лабораторных условиях изучали влияние вида муки на структурно-механические свойства макаронного теста.

В ходе однофакторного эксперимента независимую переменную x (вид муки) варьировали на трех уровнях (табл.1.3). В качестве функции отклика y приняли предел прочности макаронного теста при растяжении (кПа). При каждом уровне варьирования фактора опыты проводили в пятикратной последовательности.

Выполнить статистическую обработку результатов однофакторного дисперсионного анализа.

Таблица 1.3

Матрица однофакторного эксперимента

Уровень варьирования фактора x	Значение фактора (вид муки из пшеницы)	Параллельный опыт				
		1	2	3	4	5
1	Твердой	47,3	49,0	45,4	46,3	47,9
2	Мягкой стекловидной	52,1	52,0	55,9	47,8	53,4
3	Мягкой	77,2	79,2	78,4	73,0	77,9

Контрольные вопросы:

1. Каким условиям должны удовлетворять результаты наблюдений случайной величины для проведения дисперсионного анализа?
2. Как формируется матрица наблюдений для проведения однофакторного дисперсионного анализа?
3. В чем заключается основная идея однофакторного дисперсионного анализа?
4. Каким образом устанавливают степень влияния контролируемого фактора на изучаемый процесс?
5. Какого типа практические задачи обычно решают методом однофакторного дисперсионного анализа?
6. Влияет ли изменение диапазона варьирования изучаемого фактора на результаты однофакторного дисперсионного анализа?
7. Из каких составляющих складывается оценка «общей» дисперсии случайной величины?

8. Какой статистический критерий используют для оценки влияния факторов на изучаемый технологический процесс?

Лабораторная работа №2 Проведение однофакторного дисперсионного анализа с помощью программы EXCEL

Цель работы: определить, являются ли одинаковыми математические ожидания двух или более совокупностей.

Для дисперсионного анализа в английском языке принято сокращение **ANOVA** (**An**alysis **o**f **Va**riance), что означает «дисперсионный анализ»), которое используется и в некоторых русскоязычных источниках.

Для однофакторного случая результаты расчетов принято представлять в следующем виде (табл. 2.1).

Таблица 2.1

Представление результатов расчета однофакторного дисперсионного анализа

Компоненты дисперсии	Сумма квадратов	Число степеней свободы	Средний квадрат (дисперсия)
Межгрупповая (влияющий фактор)	$\sum_{i=1}^{m}(\overline{X_i}-\overline{X})^2$	m-1	$S_1^2=\sum_{i=1}^{m}(\overline{X_i}-\overline{X})^2$
Внутригрупповая (случайное влияние)	$\sum_{i=1}^{n}\sum_{i=1}^{m}(\overline{X_{ij}}-\overline{X})^2$	m(n-1)	$S_1^2\sum_{i=1}^{n}\sum_{i=1}^{m}(\overline{X_{ij}}-\overline{X})^2/m(n-1)$
Общая	$\sum_{i=1}^{n}\sum_{i=1}^{m}(\overline{X_{ij}}-\overline{X})^2$	m n - 1	$S_1^2\sum_{i=1}^{n}\sum_{i=1}^{m}(\overline{X_{ij}}-\overline{X})^2/(mn-1)$

Здесь внутригрупповая дисперсия характеризует влияние случайной составляющей, а межгрупповая — влияние изучаемого фактора.

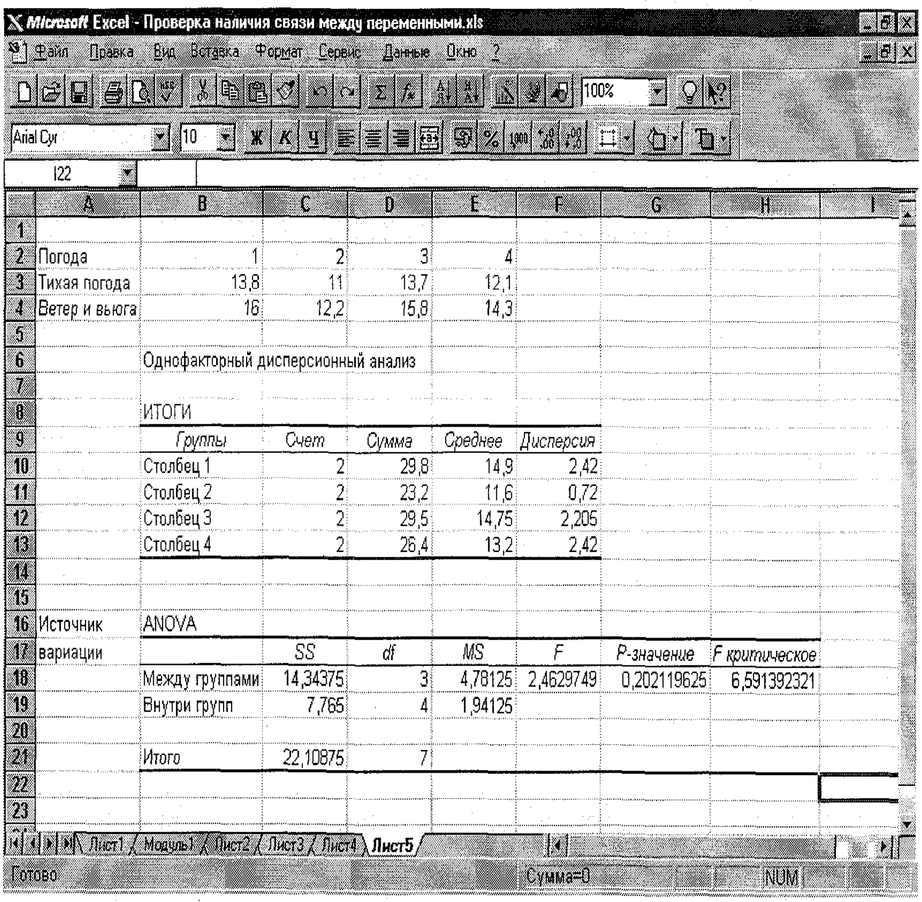
Пример условный, но базируется на реальной задаче, в которой рассматривается влияние погоды на изменение продолжительности систолической остановки сердца при введении бария хлорида. Допустим (табл. 2.2), у нас есть набор данных, которые характеризуют продолжительность реакции лабораторных животных на некоторый препарат при различных погодных условиях.

Таблица 2.2

Исходные данные примера однофакторного анализа

Погода	Экспериментальные животные			
	1	2	3	4
Тихая погода	13,8	11	13,7	12,1
Ветер и вьюга	16	12,2	15,8	14,3

После того как наши данные набраны в электронной таблице (строки 2-4 и столбцы А-Е на рис. 2.1), входим в меню, выбирая последовательно Сервис, а затем Анализ данных, в результате чего появится окно (рис.2.2).



	A	B	C	D	E	F	G	H
1								
2	Погода	1	2	3	4			
3	Тихая погода	13,8	11	13,7	12,1			
4	Ветер и вьюга	16	12,2	15,8	14,3			
5								
6		Однофакторный дисперсионный анализ						
7								
8		ИТОГИ						
9		*Группы*	*Счет*	*Сумма*	*Среднее*	*Дисперсия*		
10		Столбец 1	2	29,8	14,9	2,42		
11		Столбец 2	2	23,2	11,6	0,72		
12		Столбец 3	2	29,5	14,75	2,205		
13		Столбец 4	2	26,4	13,2	2,42		
14								
15								
16	Источник	ANOVA						
17	вариации		*SS*	*df*	*MS*	*F*	*P-значение*	*F критическое*
18		Между группами	14,34375	3	4,78125	2,4629749	0,202119625	6,591392321
19		Внутри групп	7,765	4	1,94125			
20								
21		Итого	22,10875	7				
22								
23								

Рисунок 2.1 Результаты обработки данных однофакторного эксперимента

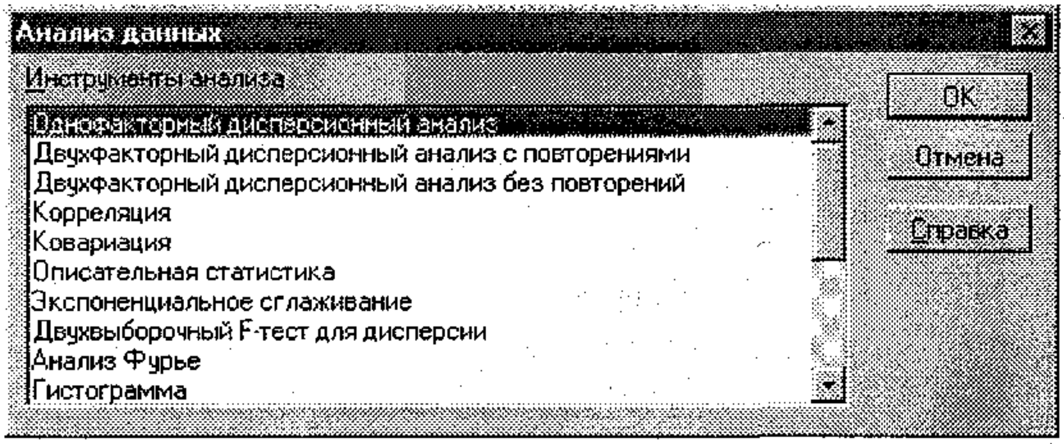


Рисунок 2.2 Окно выбора метода обработки

В этом окне необходимо выбрать Однофакторный дисперсионный анализ, после чего откроется новое окно (рис. 2.3).

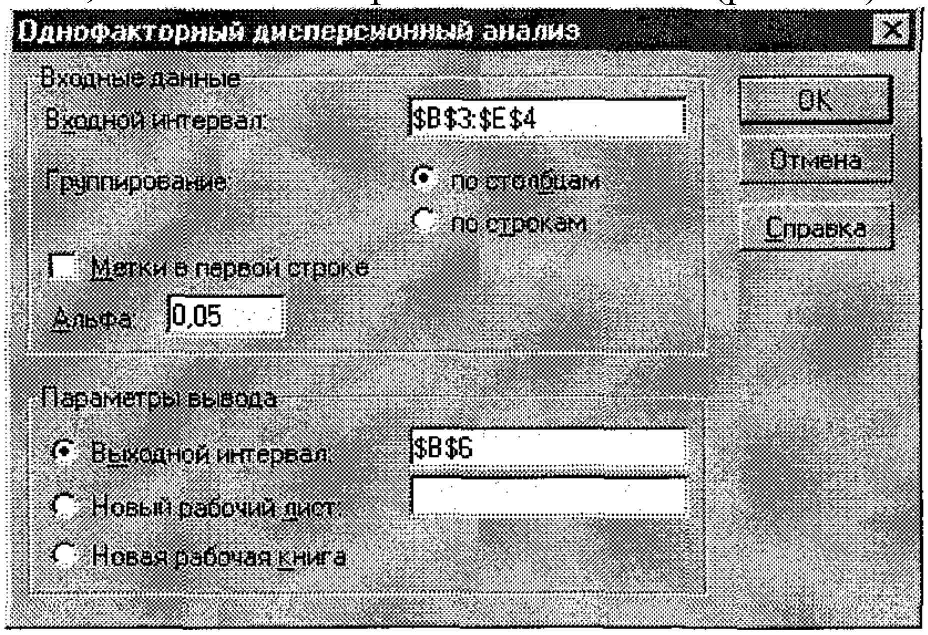


Рисунок 2.3 Окно установки параметров обработки данных для однофакторного дисперсионного анализа

В этом окне необходимо задать исходные данные для дисперсионного анализа.

Входной интервал — необходимо отметить таблицу, в которой размещены исходные данные (левая верхняя и правая нижняя ячейки).

Группирование — необходимо указать, в строках или в столбцах находятся данные, относящиеся к одному уровню фактора (в данной ситуации — в строках).

Альфа — требуемый уровень значимости.

Выходной интервал — вводится ссылка на ячейку, расположенную в левом верхнем углу выходного диапазона (мета, куда вы хотите поместить результат). Размеры выходной области будут рассчитаны автоматически.

Новый рабочий лист — выбирается в том случае, когда вы хотите поместить результаты работы на другой лист; при этом в соответствующем окошке указывается диапазон размещения результатов аналогично предыдущему пункту.

Новая рабочая книга — выбирается, если вы хотите поместить результаты в новую книгу; результаты дисперсионного анализа при этом будут размещаться на первом листе новой книги, начиная с ячейки A1.

Исходные данные и результаты для нашего примера приведены на рис. 2.1.

Здесь:

SS — сумма квадратов;

между группами — межгрупповая сумма квадратов;

внутри групп — внутригрупповая сумма квадратов;

итого — общая (полная) сумма квадратов;

df — число степеней свободы;

MS — средний квадрат (фактически дисперсия);

F — расчетное значение критерия Фишера;

P-значение — расчетное значение минимальной значимости;

F-критическое — критическое значение распределения Фишера.

В нашем примере расчетное значение критерия Фишера (3,026) меньше критического 6.59). Из этого следует, что мы принимаем гипотезу об отсутствии влияния погоды на фармакологические реакции при принятом уровне значимости 0,05.

P-значение в нашем примере равно 0,133. Поскольку оно достаточно велико, нет основания отвергать гипотезу о равенстве дисперсий.

Замечание к дисперсионному анализу в Excel

В функции дисперсионного анализа в Excel проверяется следующая альтернативная гипотеза: дисперсия числителя больше дисперсии знаменателя, то есть односторонняя гипотеза. При этом не учитывается возможность ситуации, когда F-расчетное меньше 1. В таком случае выполняемая в Excel односторонняя проверка с критическим значением распределения Фишера неправомочна. Для корректной проверки необходимо или пересчитать расчетное и критическое значение критериев Фишера и после этого выполнить их сравнение, или же рассчитать нижнюю критериальную границу.

Допустим, что подобная ситуация возникла в нашем примере (то есть расчетное значение критерия меньше 1). Тогда новое значение расчетного значения F-критерия определяется по формуле =1/F16, а соответствующее ему критическое значение — из функции = FPACПОБР(0,05; D17;D16). Если новое расчетное значение больше нового критического, то гипотеза о равенстве дисперсий отклоняется и принимается гипотеза о том, что внутригрупповая дисперсия больше межгрупповой.

Для второго варианта проверки расчетное критическое значение не изменяется, но рассчитываются новые значения для критериальных границ: нижней =FPACПОБР(0,975;D16;D17) и верхней =FPACПОБР(0,025; D16;D17). Если расчетное критическое значение находится внутри этих границ, то гипотеза об отсутствии различия принимается. Если расчетное значение больше верхнего или меньше нижнего, тогда принимается альтернативная гипотеза, которая в данном случае звучит так: межгрупповая и внутригрупповая дисперсия различается статистически значимо.

Однофакторная задача с неравномерным числом испытаний

Достаточно часто возникает ситуация, при которой число опытов для разных значений уровня фактора различно. Это может быть связано, например, с тем, что часть лабораторных животных во время эксперимента погибла или не проявила требуемой реакции и пр.

Рассмотрим пример, использованный в предыдущей задаче, но с меньшим количеством (равным трем) наблюдений для погоды.

Проведем вспомогательные вычисления (таблица 2.3).

Таблица 2.3

Формулы вспомогательных вычислений

Ячейка	Содержание	Комментарий
B5	=ЧСТРОК(B3:B4)	Число уровней варьирования фактора
от B6 до E7	=B3*B3 в B6, остальные размножаются	Квадраты значений наблюдений
от F3	=СУММ(B3:E3) в F3,	Сумма значений наблюдений

до F4	далее размножается	
от G3 до G 4	=ЧИСЛСТОЛБ(ВЗ:ЕЗ) в G3, далее размножается	Количество наблюдений в каждой строке
от НЗ до Н4	=F3*F3/G3 в НЗ, далее размножается	Квадраты сумм, деленные на число наблюдений

После этого мы можем рассчитывать значения, которые непосредственно входят в таблицу дисперсионного анализа. В табл. 2.4 приведены формулы (с соответствием размещения по ячейкам), по которым и выполняется расчет всех необходимых промежуточных значений. Как обычно, однотипные формулы набираются один раз, а затем размножаются.

Таблица 2.4

Формулы для формирования вспомогательной таблицы

	A	B	C	D	E
9	=СУММ(Н3:Н4)-СУММ(F3:F4)*СУММ(F3:F4)/СУММ(G3:G4)	=ЧСТР0К(ВЗ:В4)-1	=B9/C9	=D9/D10	=FРАПОБР(0,05;C9;C10)
10	=B11-B9	=СУММ(G3:G4)-B5	=B10/C10	=D9/D10	
11	=CVMM(F6:F7)-СУММ(F3:F4)*СУММ(F3:F4)/СУММ(G3:G4)				

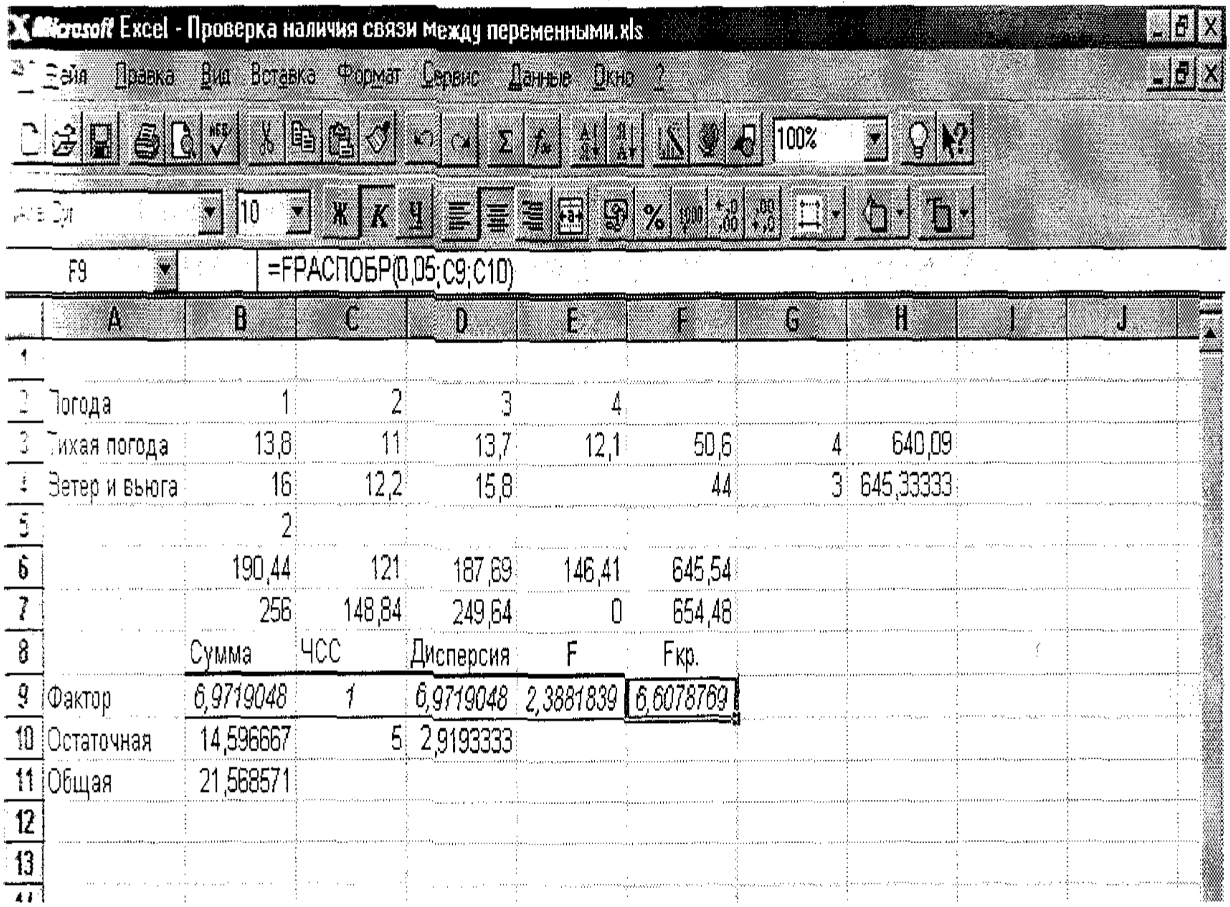
В табл. 2.4, соответствующей по размещению табл. 2.3, приведено описание содержания ячеек.

Таблица 2.4

Описание содержимого ячеек таблицы 2.3

	A	го	C	D	E
9	Сумма квадратов, обусловленная влиянием фактора	Число степеней свободы для дисперсии, обусловленной влиянием фактора	Дисперсия, обусловленная влиянием фактора	Расчетное значение критерия Фишера	Критическое значение распределения Фишера
10	Сумма квадратов, обусловленная влиянием случайной составляющей	Число степеней свободы для дисперсии, обусловленной влиянием случайной составляющей	Дисперсия, обусловленная влиянием случайной составляющей	—	—
11	Общая сумма квадратов				

Результаты работы приведены на рис. 2.4 Из него видно, что, поскольку расчетное значение критерия Фишера меньше критического, следовательно, влияние фактора отсутствует.



	A	B	C	D	E	F	G	H	I	J
1										
2	Погода	1	2	3	4					
3	Тихая погода	13,8	11	13,7	12,1	50,6	4	640,09		
4	Ветер и вьюга	16	12,2	15,8		44	3	645,33333		
5		2								
6		190,44	121	187,69	146,41	645,54				
7		256	148,84	249,64	0	654,48				
8		Сумма	ЧСС	Дисперсия	F	Fкр.				
9	Фактор	6,9719048	1	6,9719048	2,3881839	6,6078769				
10	Остаточная	14,596667	5	2,9193333						
11	Общая	21,568571								
12										
13										

Рисунок 2.4 Результаты обработки данных однофакторного дисперсионного анализа при неравномерном дублировании опытов

Задание. Произвести дисперсионный анализ данных, представленных в работе 1.

Контрольные вопросы:

1. Какие задачи дисперсионного анализа технологических процессов пищевых производства вам известны?
2. В чем сущность однофакторного дисперсионного анализа?
3. Какие условия должны быть для проведения однофакторного анализа?

Лабораторная работа № 3 Многофакторный дисперсионный анализ

Цель работы: ознакомить с методом многофакторного дисперсионного анализа; овладеть практическими навыками анализа технологических процессов методом многофакторного дисперсионного анализа.

Реальные технологические процессы пищевых производств протекают в условиях воздействия двух и более факторов, влияние которых, в ряде случаев, приводит к нестабильности протекания технологического процесса.

Пусть в предыдущем примере с анализом влияния квалификации оператора (фактора x_1) на влажность продукта все показания снимались в различное время суток (в первой, второй и третьей смене - фактор x_2). В этом случае требуется выяснить, обусловливается ли рассеивание полученных значений влажности продукта и их средних значений в группах различием в квалификациях операторов (фактора x_1) или различием между временем суток (фактор x_2).

Рассмотренные и аналогичные задачи могут быть решены также с применением методов дисперсионного анализа.

Задача многофакторного дисперсионного анализа (МДА) формулируется следующим образом (на примере задачи с двумя факторами). Пусть изучается влияние на технологический процесс двух одновременно действующих факторов x_1 и x_2. Представим в табл. 3.1 результаты эксперимента, состоящего из М = u_1u_2m наблюдений y_{jgl}, где j - порядковый номер уровня варьирования фактора x_1 (j = 1, 2,..., u_1); g- порядковый номер уровня варьирования фактора x_2 (g = 1, 2,..., u_2); l - порядковый номер параллельного опыта в серии на каждом jg-ом сочетании уровней двух факторов (l = 1, 2, m_{jg})

При указанном расположении наблюдений их рассеяние в каждой серии относительно среднего той же серии обусловлено действием только случайных причин. Рассеяние самих средних в сериях по всем возможным сочетаниям уровней x_1 и x_2 вокруг общего среднего помимо фактора случайности вызывается влиянием фактора взаимодействия x_1x_2. Кроме этих факторов на рассеяние средних по строкам оказывает влияние только один фактор x_1, а на рассеяние средних по столбцам – только один фактор x_2, так как все уровни другого фактора в каждом их этих случаев усреднены.

Таблица 3.1

Матрица двухфакторного эксперимента

Номер j уровня варьирования фактора x_1	Номер g уровня варьирования фактора x_2					
	1	2	…	g	…	u_2
1	y_{111}	y_{121}	…	y_{1g1}	…	y_{1u21}
	y_{112}	y_{122}	…	y_{1g2}	…	y_{1u22}
	..	..	…	..	…	..
	..	..	…	..	…	..
	y_{11l}	y_{12l}	…	y_{1gl}	…	y_{1u2l}
	..	..	…	..	…	..
	..	..	…	..	…	..
	y_{11m}	y_{12m}	…	y_{1gm}	…	y_{1u2m}
2	y_{211}	y_{221}	…	y_{2g1}	…	y_{2u21}
	y_{212}	y_{222}	…	y_{2g2}	…	y_{2u22}
	..	..	…	..	…	..
	..	..	…	..	…	..
	y_{21l}	y_{22l}	…	y_{2gl}	…	y_{2u2l}
	..	..	…	..	…	..
	..	..	…	..	…	..
	y_{21m}	y_{22m}	…	y_{2gm}	…	y_{2u2m}
			…		…	
j	y_{j11}	y_{j21}	…	y_{jg1}	…	y_{ju21}
	y_{j12}	y_{j22}	…	y_{jg2}	…	y_{ju22}
	..	..	…	..	…	..
	..	..	…	..	…	..
	y_{j1l}	y_{j2l}	…	y_{jgl}	…	y_{ju2l}
	..	..	…	..	…	..
	..	..	…	..	…	..
	y_{j1m}	y_{j2m}	…	y_{jgm}	…	y_{ju2m}
			…		…	
u_1	y_{u111}	y_{u121}	…	y_{u1g1}	…	y_{u1u21}
	y_{u112}	y_{u122}	…	y_{u1g2}	…	y_{u1u22}
	..	..	…	..	…	..
	..	..	…	..	…	..
	y_{u11l}	y_{u12l}	…	y_{u1gl}	…	y_{u1u2l}
	..	..	…	..	…	..
	..	..	…	..	…	..
	y_{u11m}	y_{u12m}	…	y_{u1gm}	…	y_{u1u2m}

Таким образом, имеем равенство

$$S_0^2 = S_\varepsilon^2 + S_{x_1}^2 + S_{x_2}^2 + S_{x_1 x_2}^2 \qquad 3.1$$

где S_0^2 - оценка «общей» дисперсии;

S_ε^2 - оценка остаточной дисперсии;

$S_{x_1}^2$ - оценка дисперсии рассеивания «между строками»;

$S_{x_2}^2$ - оценка дисперсии рассеивания «между столбцами»;

$S_{x_1x_2}^2$ - оценка дисперсии рассеивания «между сериями».

При обработке результатов МДА предварительно определяют: суммы наблюдений отклика Y_j по строкам и $Y_g^{'}$ по столбцам

$$Y_j = \sum_{j=1}^{u_2} \sum_{l=1}^{m} y_{jgl} \qquad Y_g^{'} = \sum_{j=1}^{u_2} \sum_{l=1}^{m} y_{jgl} \qquad 3.2$$

сумму квадратов всех M u_1u_2m = наблюдений

$$Q_1 = \sum_{j=1}^{u_1} \sum_{g=1}^{u_2} \sum_{l=1}^{m} y_{jgl}^2 \qquad 3.3$$

сумму квадратов итогов (сумм) по строкам, деленную на число наблюдений в строке,

$$Q_2 = \frac{1}{u_2 m} \sum_{j=1}^{u_1} Y_j^2 \qquad 3.4$$

сумму квадратов итогов (сумм) по столбцам, деленную на число наблюдений в столбце,

$$Q_3 = \frac{1}{u_1 m} \sum_{g=1}^{u_2} Y_g^{'2} \qquad 3.5$$

сумму квадратов итогов (сумм) по сериям, деленную на число наблюдений в серии,

$$Q_4 = \frac{1}{u} \sum_{j=1}^{u_1} \sum_{g=1}^{u_2} \left(\sum_{l=1}^{m} y_{jgl} \right)^2 \qquad 3.6$$

квадрат общего итога (суммы), деленный на число всех наблюдений,

$$Q_5 = \frac{1}{M}\left(\sum_{j=1}^{u_1} Y_j\right)^2 = \frac{1}{M}\left(\sum_{g=1}^{u_2} Y'_g\right)^2 \qquad 3.7$$

Затем определяют суммы квадратов S_0, S_ε, S_{x_1}, S_{x_2}, $S_{x_1x_2}$.

«Общая» сумма квадратов, характеризующая рассеяние отдельных наблюдений y_{jgl} в общей совокупности за счет влияния всех факторов

$$S_0 = Q_1 - Q_5 \qquad 3.8$$

Сумма квадратов отклонений «внутри серий», характеризующая рассеивание отдельных наблюдений y_{jgl} в сериях только за счет влияния фактора случайности, так как на протяжении серии факторы x_1 и x_2 остаются неизменными

$$S_\varepsilon = Q_1 - Q_4 \qquad 3.9$$

Сумма квадратов отклонений «между строками». Сумма S_{x_1}/u_2m характеризует рассеяние средних по строкам в результате действия фактора случайности, фактора x_1 и фактора взаимодействия

$$S_{x_1} = Q_2 - Q_5 \qquad 3.10$$

Сумма квадратов отклонений «между столбцами». Сумма S_{x_2}/u_1m характеризует рассеяние средних по столбцам в результате действия фактора случайности, фактора x_2 и фактора взаимодействия

$$S_{x_2} = Q_3 - Q_5 \qquad 3.11$$

Сумма квадратов отклонений «между сериями». Сумма $S_{x_1x_2}/m$ характеризует рассеяние средних $\overline{y_{jg}}$ серий в результате действия фактора случайности и фактора взаимодействия.

$$S_{x_1x_2} = Q_4 + Q_5 - Q_2 - Q_3 \qquad 3.12$$

Каждая из указанных сумм квадратов, поделенная на отвечающее ей число степеней свободы f_0, f_ε, f_{x1}, f_{x2}, f_{x1x2}, дает оценку соответствующей дисперсии, входящей в формулу (3.1):

оценка «общей» дисперсии по всем $M=u_1u_2m$ наблюдениям

$$S_0^2 = \frac{S_0}{M-1} \qquad 3.13$$

с числом степеней свободы $f = u_1 u_2 m-1=M-1$;
оценка дисперсии «внутри серий», или оценка остаточной дисперсии

$$S_\varepsilon^2 = \frac{S_\varepsilon}{u_1 u_2 (m-1)} \qquad 3.14$$

с числом степеней свободы $f_\varepsilon = u_1 u_2 (m-1)$;

оценка дисперсий «между сериями»:

$$S_{x_1}^2 = \frac{S_{x_1}}{u_1 - 1} \qquad 3.15$$

с числом степеней свободы $f_{x1} = u_1 - 1$.
оценка дисперсии рассеивания «между столбцами»:

$$S_{x_2}^2 = \frac{S_{x_2}}{u_2 - 1} \qquad 3.16$$

с числом степеней свободы $f_{x2} = u_2 - 1$.
оценка дисперсии рассеивания «между сериями»:

$$S_{x_1 x_2}^2 = \frac{S_{x_1 x_2}}{(u_1 - 1)(u_2 - 1)} \qquad 3.17$$

с числом степеней свободы $f_{x1x2} = (u_1-1)(u_2-1)$.

Правильность подсчета числа степеней свободы проверяют с помощью соотношения $f_0 = f_\varepsilon + f_{\text{x}} + f_{x1} + f_{x2} + f_{x1x2}$.

Анализ существенности влияния факторов $x1$, $x2$ и их взаимодействия $x_1 x_2$ производят на основании критерия Фишера при выбранном уровне значимости p.

Для оценки существенности влияния каждого фактора в отдельности рассчитывают соответствующие значения критериев Фишера:

$$F_p = \frac{S_{x_1}^2}{S_{x_1 x_2}^2} \qquad F_p = \frac{S_{x_2}^2}{S_{x_1 x_2}^2} \qquad 3.18$$

если выполняется условие $F_p > F_m$ то влияние фактора на изучаемый процесс признают существенным.

Для оценки существенности влияния взаимодействия $x_1 x_2$ рассчитывают значение критерия Фишера:

$$F_p = \frac{S^2_{x_1 x_2}}{S^2_{\varepsilon}} \qquad 3.19$$

Влияние взаимодействия x_1x_2 на изучаемый процесс признают существенным, если выполняется условие $F_р > F_т$.

Пример

Смешивание сыпучих и жидких рецептурных компонентов в порошковой технологии помадных конфет является одним из основных технологических процессов. Характеристикой процесса смешивания помадной массы является энергоемкость процесса, т. е. величина удельной работы смешивания. Значение последней, как известно, зависит от рецептурного состава помадной массы, конструкции камеры смешивания, конфигурации месильных органов и их частоты вращения.

Предварительные эксперименты показали, что при периодическом смешивании помадной массы величина затрат механической энергии значительно изменяется при изменении условий дозирования – способа и последовательности подачи жидких и сыпучих рецептурных компонентов в смеситель.

Однако предварительные эксперименты не позволили однозначно сказать, что является причинами значительной нестабильности удельной работы смешивания – неконтролируемые изменения технологических параметров смешивания, случайные ошибки измерений или изменения условий дозирования рецептурных компонентов.

В связи с этим были проведены дополнительные эксперименты, в ходе которых периодическим способом путем смешивания жидких и сыпучих компонентов в универсальной смесительно-формующей машине (УСФУ) готовили помадную массу на основе порошкообразного сахаропаточного полуфабриката.

В качестве переменных, влияющих на энергоемкость процесса смешивания были приняты: способ дозирования сыпучих компонентов (x_1), способ дозирования жидких компонентов (x_2).

В качестве выходной величины использовали удельную работу смешивания y (кДж/кг).

Помадную массу с массовой долей влаги 9 % готовили в УСФУ в течение 10 мин при частоте вращения месильных органов 100 мин$^{-1}$. Каждый из факторов варьировался на двух уровнях: уровень 1 –

мгновенное внесение всего заданного количества твердой (жидкой) фазы; уровень 2 – дозирование всего заданного количества твердой (жидкой) фазы в течение продолжительности смешивания (10 мин). В ходе смешивания рецептурных компонентов измеряли удельную мощность смешивания. По окончании смешивания определяли удельную работу смешивания. При каждом сочетании уровней варьирования факторов было проведено по три параллельных опыта. В табл. 3.2 представлены значения удельной работы смешивания, полученные в параллельных опытах.

Таблица 3.2

Матрица двухфакторного эксперимента

Номер j уровня варьирования фактора x_1	Номер g уровня варьирования фактора x_2	
1	740,2	628,7
	742,0	630,0
	744,9	636,6
2	392,3	310,0
	390,7	310,4
	398,0	311,6

Статистическую обработку экспериментальных данных проведем с помощью метода многофакторного дисперсионного анализа (МДА).

Для этого по формулам (3.2) определяем сумму Yj по строкам для каждого уровня фактора x_1.

Для уровня j = 1

$$Y_1 = \sum_{j=1}^{2} \sum_{l=1}^{3} y_{1gl} = 740{,}2 + 742{,}0 + 744{,}9 + 628{,}7 + 630{,}0 + 636{,}6 = 4122{,}4$$

Для уровня j = 2

$$Y_2 = \sum_{j=1}^{2} \sum_{l=1}^{3} y_{2gl} = 392{,}3 + 390{,}7 + 398{,}0 + 310{,}0 + 310{,}4 + 311{,}6 = 2113$$

Определяем сумму Y'_g по столбцам для каждого уровня фактора x_2.

Для уровня g = 1

$$Y_1' = \sum_{j=1}^{2} \sum_{l=1}^{3} y_{j1l} = 740{,}2 + 742{,}0 + 744{,}9 + 392{,}3 + 390{,}7 + 398{,}0 = 3408{,}1$$

Для уровня $g = 2$

$$Y_2' = \sum_{j=1}^{2} \sum_{l=1}^{3} y_{j2l} = 628{,}7 + 630{,}0 + 636{,}6 + 310{,}0 + 310{,}4 + 311{,}6 = 2827{,}3$$

По формуле (3.3) рассчитываем сумму квадратов всех наблюдений

$$Q_1 = \sum_{j=1}^{2} \sum_{g=1}^{2} \sum_{l=1}^{3} y_{jgl}^2 = 3605251{,}8$$

Сумму квадратов итогов (сумм) по строкам, деленную на число наблюдений в строке, определяем по (3.4)

$$Q_2 = \frac{1}{2 \cdot 3} \sum_{j=1}^{2} Y_j^2 = \frac{1}{2 \cdot 3}(4122{,}4^2 + 2113^2) = 3576491{,}8$$

Расчет суммы квадратов итогов (сумм) по столбцам, деленной на число наблюдений в столбце, проводим по (3.5)

$$Q_3 = \frac{1}{2 \cdot 3} \sum_{g=1}^{2} Y_g'^2 = \frac{1}{2 \cdot 3}(3408{,}1^2 + 2827{,}3^2) = 3268128{,}5$$

По формуле (3.6) вычисляем сумму квадратов итогов (сумм) по сериям, деленную на число наблюдений в серии,

$$Q_4 = \frac{1}{3} \sum_{j=1}^{2} \sum_{g=1}^{2} \left(\sum_{l=1}^{3} y_{jgl} \right)^2 = \frac{1}{3}(2227{,}1^2 + 1181^2 + 1895{,}3^2 + 939^2) = 3605173{,}8$$

Квадрат общего итога (суммы), деленный на число всех наблюдений, определяем по формуле (3.7)

$$Q_5 = \frac{1}{M} \left(\sum_{j=1}^{2} Y_j \right)^2 = \frac{1}{2 \cdot 2 \cdot 3}(4122{,}4^2 + 2113^2) = 3240017{,}7$$

По формулам (3.8 – 3.12) определяем суммы квадратов отклонений

$S_0 = 365234{,}1$ $\qquad$ $S_\varepsilon = 78$

$S_{x_1} = 336474{,}1$ $S_{x_2} = 28110{,}8$ $S_{x_1x_2} = 571{,}2$

Число степеней свободы:

$f_0 = M - 1 = u_1u_2m - 1 = 2 \cdot 2 \cdot 3 - 1 = 11;$

$f_\varepsilon = u_1u_2\,(m - 1) = 2 \cdot 2 \cdot (3 - 1) = 8;$

$f_{x1} = u_1 - 1 = 2 - 1 = 1;$

$f_{x2} = u_2 - 1 = 2 - 1 = 1;$

$f_{x1x2} = (u_1 - 1)(u_2 - 1) = (2 - 1)(2 - 1) = 1$

Оценки дисперсий определяем по формулам (3.13 – 3.17) как отношение каждой суммы квадратов отклонений к соответствующему числу степеней свободы:

$S_0^2 = 33203{,}1$ $S_\varepsilon^2 = 9{,}75$

$S_{x_1}^2 = 336474{,}1$ $S_{x_2}^2 = 28110{,}8$ $S_{x_1x_2} = 571{,}2$

Для оценки существенности влияния факторов x_1, x_2 и их взаимодействия x_1x_2 на выходную величину y по формулам (3.18 – 3.19) определяем расчетные значения критериев Фишера:

$$F_{p1} = \frac{S_{x_1}^2}{S_{x_1x_2}^2} = \frac{336474{,}1}{571{,}2} = 589{,}06$$

$$F_{p2} = \frac{S_{x_2}^2}{S_{x_1x_2}^2} = \frac{28110{,}8}{571{,}2} = 49{,}21$$

$$F_p = \frac{S_{x_1x_2}^2}{S_\varepsilon^2} = \frac{571{,}2}{9{,}75} = 58{,}58$$

Табличное значения критерия Фишера при уровне значимости $p = 0{,}05$ и числе степеней свободы $f_1 = f_{x1} = f_{x2} = 1$ и $f_2 = f_{x1\,x2} = 1$ составляет $F_т = 161{,}0$ (см. прил. 2).

Сравниваем расчетные значения критерия Фишера с табличным: $F_{р1} > F_т$ и $F_{р2} < F_т$, следовательно, влияние фактора x_1 следует признать существенным, а влиянием фактора x_2 можно пренебречь (табл. 3.3).

Таблица 3.3

Результаты многофакторного дисперсионного анализа

Входная переменная	Значение критерия Фишера		Влияние на выходную величину
	расчетное	табличное	
x_1	589,06	161,0	Значимое
x_2	49,21	161,0	Не значимое
x_1x_2	58,58	5,32	Значимое

Вместе с тем реальный процесс смешивания протекает при совместном влиянии факторов x_1 и x_2 (одновременное дозирование сыпучих и жидких компонентов), в связи с чем к рассмотрению следует принять их совместное влияние x_1x_2.

Табличное значения критерия Фишера при уровне значимости $p = 0{,}05$ и числе степеней свободы $f_1 = f_{x1x2} = 1$ и $f_2 = f_\varepsilon = 8$ составляет $F_т = 5{,}32$ (см. прил. 2).

Сравниваем расчетное значения критерия Фишера с табличным: $F_{р12} > F_т$, следовательно, влияние взаимодействия факторов x_1x_2 следует признать существенным.

Таким образом, результаты двухфакторного дисперсионного анализа позволяют сделать вывод о том, что способ дозирования жидких и сыпучих рецептурных компонентов в смеситель при периодическом способе приготовления помадной массы порошковой технологии существенно влияет на затраты удельной работы смешивания. Указанный факт следует учитывать в производственных условиях при приготовлении помадной массы порошковой технологии.

Задание

Необходимо оценить влияние стекловидности зерна пшеницы и механических факторов (комбинации системы помола) на выход муки высоких сортов.

Стекловидный фактор x_1 варьировался на двух уровнях: на первом уровне ($j = 1$) стекловидный фактор имел значение $x_1 = 40$ %, на втором уровне ($j = 1$) – $x_1 = 58$ %. Сортовой помол x_2 был трех разновидностей. Для каждой комбинации уровней варьирования факторов было выполнено по три повторных наблюдения выхода муки y, %, значения которых представлены в табл. 3.4.

Выполнить статистическую обработку результатов многофакторного дисперсионного анализа, оценив влияние каждого фактора на выход муки высоких сортов.

Таблица 3.4

Матрица двухфакторного эксперимента

Номер j уровня варьирования фактора x_1	Номер g уровня варьирования фактора x_2		
	1	2	3
1	24,7	24,8	20,8
	24,5	25,0	20,8
	25,0	25,1	20,8
2	25,9	25,9	21,3
	26,0	26,3	21,4
	26,0	26,3	21,7

Контрольные вопросы

1. Как формируется матрица наблюдений для проведения многофакторного дисперсионного анализа?
2. На какие составляющие раскладывается оценка «общей дисперсии» в двухфакторном дисперсионном анализе?
3. Каким образом производят оценивание существенности влияния факторов изменчивости и их взаимодействия в многофакторном дисперсионном анализе?
4. Как в двухфакторном дисперсионном анализе формируются оценки дисперсий рассеиваний: «общего», «внутри серий», «между строками», «между столбцами», «между сериями»?

Лабораторная работа №4 Проведение многофакторного дисперсионного анализа с помощью программы EXCEL

Цель работы: определить, являются ли одинаковыми математические ожидания двух или более совокупностей.

Двухфакторная задача с равномерным числом наблюдений в ячейке

В Excel есть два варианта двухфакторного регрессионного анализа: с повторениями и без повторений. Двухфакторный дисперсионный анализ без повторений — такая его разновидность, в каждой ячейке которого содержится только одно наблюдение. Таблица исходных данных для такого анализа имеет вид, подобный таблице однофакторного, только вместо номера испытания — номер уровня второго фактора. Таблица результатов похожа на приведенную выше таблицу, только в ней отсутствует строка взаимодействия факторов.

К сожалению, функция двухфакторного дисперсионного анализа с повторениями в Excel работает неправильно и выполнять расчеты по ней невозможно.

Двухфакторный дисперсионный анализ без повторений

(Пример условный). Допустим, что необходимо выяснить, не оказывает ли значимое влияние на изучаемую характеристику личность экспериментатора и питомник, из которого получают животных для лабораторных исследований. A priori предполагается, что не оказывает (табл. 4.1).

Таблица 4.1

Исходные данные для двухфакторного дисперсионного анализа без повторений

Экспериментатор	Питомник 1	Питомник 2	Питомник 3
A	75	67	78
B	68	65	59
C	56	69	65

Поместив данные в таблицу Excel, выбираем в меню ***Сервис***, ***Анализ данных***. Появится окно в котором выбираем Двухфакторный дисперсионный анализ без повторений. После этого появляется окно установки параметров (рис. 4.1).

Рисунок 4.1 Окно установки параметров

Входной интервал — ссылка на левую верхнюю и правую нижнюю ячейки таблицы, в которых находятся данные. Могут быть отмечены мышкой.

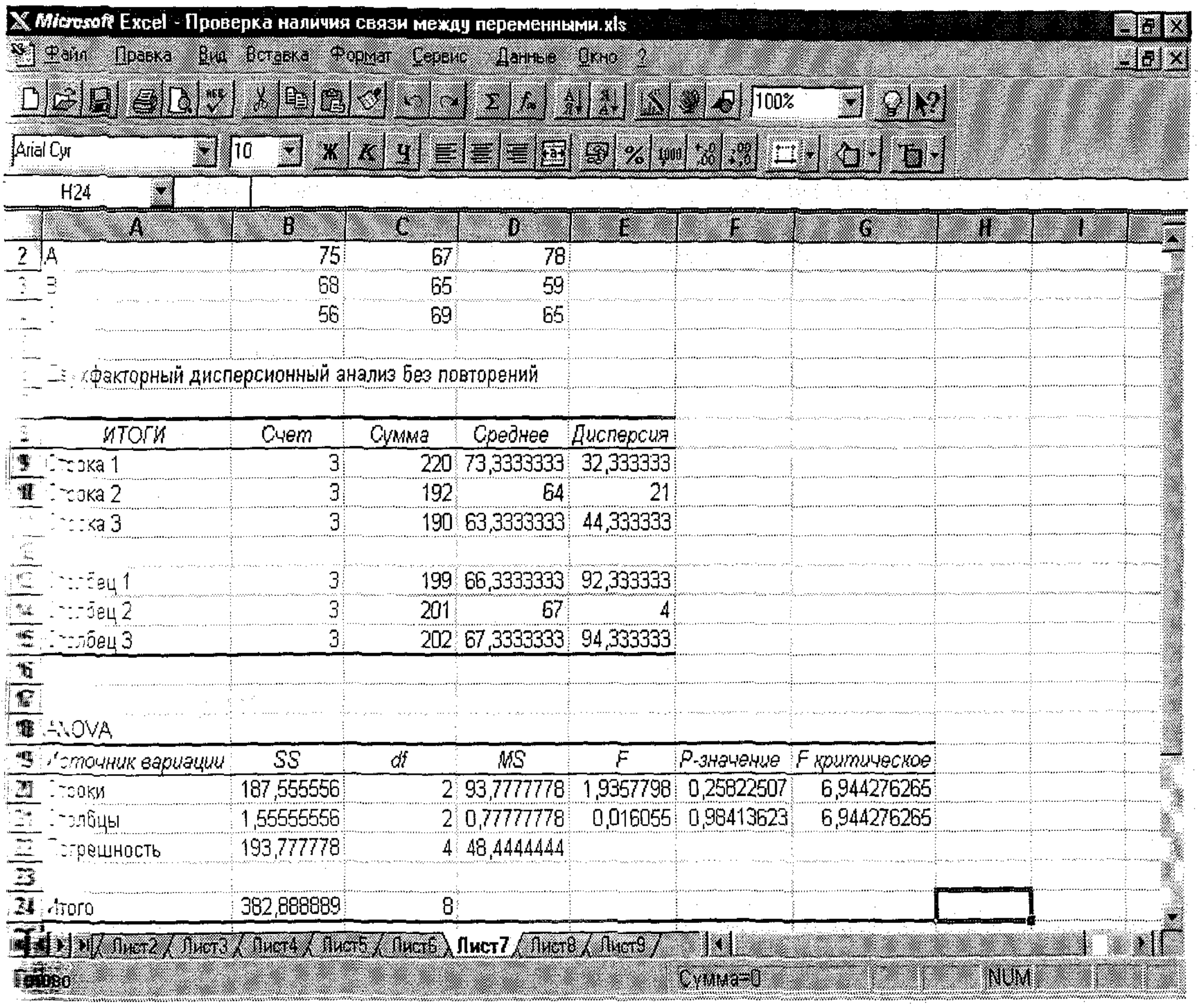
Альфа — требуемый уровень значимости.

Выходной интервал — вводится ссылка на ячейку, расположенную в левом верхнем углу выходного диапазона. Размеры выходной области будут рассчитаны автоматически.

Новый рабочий лист — выбирается в том случае, если вы хотите поместить результаты работы на другой лист (при этом в соответствующем окошке указывается диапазон размещения результатов, аналогичный указанному в предыдущем пункте).

Новая рабочая книга — выбирается в том случае, когда вы хотите поместить результаты в новую книгу (результаты дисперсионного анализа при этом будут размещаться на первом листе новой книги, начиная с ячейки А1).

После установки всех необходимых параметров выбирается ОК. Результаты работы представлены на рис. 4.2



	B	C	D
A	75	67	78
B	68	65	59
[illegible]	56	69	65

Двухфакторный дисперсионный анализ без повторений

ИТОГИ	Счет	Сумма	Среднее	Дисперсия
Строка 1	3	220	73,3333333	32,333333
Строка 2	3	192	64	21
Строка 3	3	190	63,3333333	44,333333
Столбец 1	3	199	66,3333333	92,333333
Столбец 2	3	201	67	4
Столбец 3	3	202	67,3333333	94,333333

ANOVA

Источник вариации	SS	df	MS	F	P-значение	F критическое
Строки	187,555556	2	93,7777778	1,9357798	0,25822507	6,944276265
Столбцы	1,55555556	2	0,77777778	0,016055	0,98413623	6,944276265
Погрешность	193,777778	4	48,4444444			
Итого	382,888889	8				

Рисунок 4.2 Результаты двухфакторного дисперсионного анализа без повторений

Здесь:

SS — сумма квадратов;

строки — сумма квадратов, обусловленная первым фактором (измением по строкам);

столбцы — сумма квадратов, обусловленная вторым фактором (изменением по столбцам);

погрешность — остаточная сумма квадратов, обусловленная случайной ошибкой;

df — число степеней свободы;

MS — средний квадрат (фактически дисперсия);

F — расчетное значение критерия Фишера;

P - значение — расчетное значение минимальной значимости;

F-критическое — критическое значение распределения Фишера.

Для первого фактора (строки) F-расчетное (1,93578) меньше критического 6,944276. Это заставляет принять нулевую гипотезу о равенстве дисперсий, следовательно, первый фактор не влияет на изучаемый показатель. Что касается второго фактора (столбцы), то здесь расчетное значение критерия Фишера меньше 1 (0,016055), и проверка в том виде, который приведен в Excel, некорректна.

Проверка может быть осуществлена двумя способами.

Первый способ. Проверяется односторонний критерий: одна дисперсия больше другой. Для этого находим обратное значение к имеющемуся расчетному значению =1/E21 (получим 62,28571), а затем — новое критическое значение =FРАСПОБР(0,05;C22;C21), в котором степени свободы поменялись местами (получили 19,24673). Поскольку расчетное больше критического, то принимается гипотеза, что дисперсия погрешности больше дисперсии столбцов.

Второй способ. Не изменяя расчетное значение критерия Фишера, рассчитываем новые критические значения для проверки двухсторонней альтернативной гипотезы (дисперсии неравны). Нижнее критическое при этом вычисляется по формуле: =FРАСПОБР(0,975;C21;C22) (результат равен 0,025475), а верхнее — по формуле =FРАСПОБР(0,025;C21;C22) (результат равен 10,64905). Поскольку 0,016055 меньше 0,025475, то гипотеза о неравенстве дисперсий принимается. При этом дисперсия, вызванная случайными факторами, больше.

Двухфакторный дисперсионный анализ с повторениями

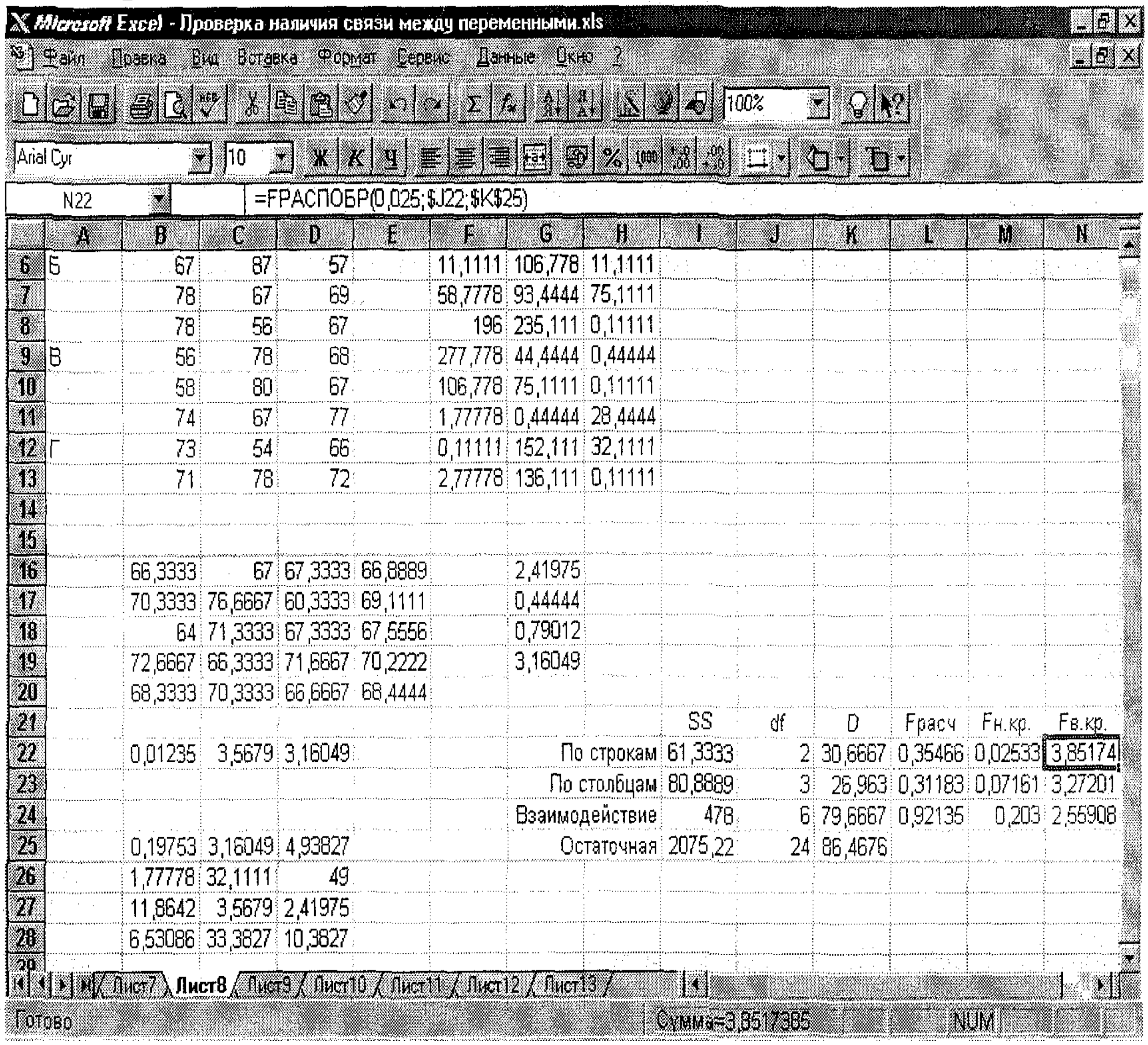
Исходные данные для анализа представлены в табл. 4.2. Каждый опыт (для каждого экспериментатора) был проведен 3 раза.

Таблица 4.2

Экспериментатор	Питомник 1	Питомник 2	Питомник 3
А	75	67	78
	68	65	59
	56	69	65
Б	66	76	55
	67	87	57
	78	67	69

В	78	56	67
	56	78	67
	58	80	67
Г	74	67	77
	73	54	66
	71	78	72

Исходные данные вводятся в Excel и содержатся в ячейках от A1 до D13 (рис. 4.3).



	A	B	C	D	E	F	G	H	I	J	K	L	M	N
6	Б	67	87	57		11,1111	106,778	11,1111						
7		78	67	69		58,7778	93,4444	75,1111						
8		78	56	67		196	235,111	0,11111						
9	В	56	78	68		277,778	44,4444	0,44444						
10		58	80	67		106,778	75,1111	0,11111						
11		74	67	77		1,77778	0,44444	28,4444						
12	Г	73	54	66		0,11111	152,111	32,1111						
13		71	78	72		2,77778	136,111	0,11111						
14														
15														
16		66,3333	67	67,3333	66,8889		2,41975							
17		70,3333	76,6667	60,3333	69,1111		0,44444							
18		64	71,3333	67,3333	67,5556		0,79012							
19		72,6667	66,3333	71,6667	70,2222		3,16049							
20		68,3333	70,3333	66,6667	68,4444									
21									SS	df	D	Fрасч	Fн.кр.	Fв.кр.
22		0,01235	3,5679	3,16049				По строкам	61,3333	2	30,6667	0,35466	0,02533	3,85174
23								По столбцам	80,8889	3	26,963	0,31183	0,07161	3,27201
24								Взаимодействие	478	6	79,6667	0,92135	0,203	2,55908
25		0,19753	3,16049	4,93827				Остаточная	2075,22	24	86,4676			
26		1,77778	32,1111	49										
27		11,8642	3,5679	2,41975										
28		6,53086	33,3827	10,3827										

Рисунок 4.3 Исходные данные и результаты для двухфакторного эксперимента

Выполняем следующие действия.

1. Вычисляем $\overline{X_{ij*}}$ которые размещаем в ячейках от B16 до D20.

2. Вычисляем средние значения по строкам $\overline{X_{i**}}$

3. Вычисляем средние значения по столбцам $\overline{X_{*j*}}$

4. Вычисляем общее среднее $\overline{X}$.

Формулы, использованные для вычислений, представлены в табл. 4.3.

Таблица 4.3

Ячейка	B	C	D	E
16	=СРЗНАЧ(B$2:B$4)	=СРЗНАЧ(C$2:C$4)	= СРЗНАЧ (D$2:D$4)	= СРЗНАЧ ($B16:$D16)
17	=СРЗНАЧ(B$5:B$7)	=СРЗНАЧ(C$5:C$7)	= СРЗНАЧ (D$5:D$7)	= СРЗНАЧ ($B17:$D17)
18	=СРЗНАЧ(B$8:B$10)	=СРЗНАЧ(C$8:C$10)	= СРЗНАЧ (D$8:D$10)	= СРЗНАЧ ($B18:$D18)
19	=СРЗНАЧ(B$11:B$13)	=СРЗНАЧ(C$11:C$13)	= СРЗНАЧ (D$11:D$13)	= СРЗНАЧ ($B19:$D19)
20	=СРЗНАЧ(B$16:B$19)	=СРЗНАЧ(C$16:C$19)	= СРЗНАЧ (D$16:D$19)	= СРЗНАЧ (E16:E19)

Набрать можно не все формулы, а только часть — остальные получим путем размножения перетягиванием. Необходимо набрать формулы в столбец B от 16-й до 20-й строчки, а формулы в столбцах C и D получим построчным перетягиванием ячеек столбца B. В столбце E набираем ячейку E16, а ячейки E17~E19 получаем перетягиванием. Ячейка E20 набирается вручную, причем безразлично, получите ли вы в ней сумму по строкам или по столбцам. Это будет одно и то же значение.

1. Находим квадраты разностей $(\overline{X_{ijk}} - \overline{X_{ij*}})^2$ для каждой ячейки. Массив разностей размещен в ячейках с F2 по H13.

Для этого размещаем формулы в следующих ячейках формулы, приведенные в таблице 4.4.

Таблица 4.4

Ячейка	Содержимое
F2	=(B2-B$16)*(B2-B$16)
F5	=(B5-B$17)*(B5-B$17)
F8	=(B8-B18)*(B8-B18)
F11	=(B11-B$19)*(B11-B$19)

В остальных ячейках формулы размножаются перетягиванием. При этом в столбце F ячейки размножаются сверху вниз, а столбцы G и H получаем построчным перетягиванием соответствующих ячеек столбца F.

1. Находим квадраты построчных разностей $(\overline{X_{i**}} - \overline{X})^2$. Для этого в ячейке G16 помещаем формулу =(E16-$E$20)*(E16-$E$20), которую затем размножаем на ячейки G17-G19 перетягиванием.

2. Находим квадраты разностей по столбцам $(\overline{X_{*j*}} - \overline{X})^2$ Для этого в ячейку B22 помещаем формулу =(B20-$E$20)*(B20-$E$20), которую затем размножаем построчно на ячейки C и D.

3. Находим квадраты разностей, необходимые для расчета дисперсии взаимодействия $(\overline{X_{ij*}} - \overline{X} - \overline{X_{*j*}} - \overline{X_{i**}})^2$ Для этого в ячейку B25 помещаем формулу =(B16+$E$20-B$20-$E16)*(B16+$E$20-B$20-$E16), которую размножаем на все ячейки таблицы до D28 перетягиванием.

4. Формируем таблицу дисперсионного анализа (ячейки от G21 до N25).

Таблица по содержанию в основном соответствует табл. 4.4. Отличие состоит в отсутствии строки для полной суммы квадратов. Дополнительно есть столбцы для расчетного значения критерия Фишера нижнего и верхнего критического значения для проверки гипотезы. Расчетные формулы, которые помещены в ячейки, представлены в табл. 4.5 и 4.6.

Таблица 4.5

Ячейка	I	J
22	=СУММ(А16:019)*ЧИСЛСТОЛБ(В2:D2)*ЧСТР0К(В2:В4)	=(ЧИСЛСТОЛБ(В2:02Ы)
23	=СУММ(В22:D22)*4*ЧСТРОК(В2:В4)	=4-1
24	=ЧСТРОК(В2:В4)*СУММ(В25:D28)	=(4-1)*(ЧИСЛСТОЛБ(В2:02)-1)
25	=СУММ(F2:H13)	=4*ЧИСЛСТ0ЛБ(В2:D2)*(ЧСТРОК(В2:В4)-1)

Таблица 4.6

Ячейка	K	L	M	N
22	=I22/J22	=K22/$K$25	=FРАСПОБР(0,975;$J22;$K$25)	=FРАСПОБР(0,025;$J22;$K$25)
23	=I23/J23	=K23/$K$25	=FРАСПОБР(0,975;$J23;$K$25)	=FРАСПОБР(0,025;$J23;$K$25)
24	=124/J24	=K24/$K$25	=FРАСПОБР(0,975;$J24;$K$25)	=FРАСПОБР(0,025;$J24;$K$25)
	=I25/J25			

Все три расчетных значения критерия Фишера меньше нижней критической границы. Поэтому принимается гипотеза о значимом превышении остаточной дисперсии над дисперсиями строк, столбцов и взаимодействий. Это значит, что в данном примере влияние экспериментатора и питомника значимо меньше влияния неизвестных нам случайных факторов. Если бы это были результаты реального эксперимента, то из статистического анализа следует вывод:

- Результаты экспериментов фальсифицированы, то есть рассеивание от влияющих факторов слишком мало. Или оно должно быть больше рассеивания случайных факторов (при наличии влияния факторов), или не должно отличаться от них (при отсутствии влияния). В данном случае результаты фальсифицированы.
- Эксперимент проведен некачественно. На результаты опытов влиял фактор (или несколько факторов), который мы не учитывали.

Примечание. Существуют более сложные схемы, учитывающие большее количество факторов, их взаимодействия и возможность проведения повторных опытов, неравномерность числа наблюдений в ячейке и пр. Следует отметить, что с ростом количества компонент расчеты и анализ становятся более громоздкими. При числе факторов больше двух возможно использование (с некоторыми ограничениями) регрессионного анализа.

Непараметрический дисперсионный анализ Фридмана

В том случае, когда закон распределения не является нормальным, используется непараметрический дисперсионный анализ Фридмана.

Нулевая гипотеза. Средние значения всех выборок равны.

Все случайные величины взаимно независимы.

Данные каждой выборки распределены по одному закону распределения. Обратите внимание: закон распределения каждой выборки может отличаться от закона распределения других.

Исходные данные представляются в следующем виде (табл. 4.7).

Таблица 4.7

Номера элементов совокупностей	1	2	…	j	…	n
Номера совокупностей						
1	X_{11}	X_{12}		X_{1j}		X_{1n}
2	X_{21}	X_{22}		X_{2j}		X_{2n}
…			…		…	
i	X_{i1}	X_{i2}		X_{ij}		X_{jn}
…			…		…	
m	X_{m1}	X_{m2}		X_{mj}		X_{mn}

Для этого в каждом столбце значения X заменяют их рангами (другими словами, вместо значений переменных ставится их номер в ряду, упорядоченном по возрастанию). Затем рассчитывается значение критерия:

$$\chi^2 = \frac{12\sum_{j=1}^{n}(\sum_{i=1}^{m}R_{ij})^2}{nm(n+1)} - 3m(n+1) \qquad 4.1$$

где R_{ij} — соответствующие значения рангов.

Если расчетное значение $\chi 2$ будет больше критического, взятого с заданным уровнем значимости и (n-1) степенью свободы, гипотеза о различии между партиями принимается.

При расчетах можно проверить правильность расстановки рангов и расчетов, зная, что имеет место соотношение:

$$\sum_{i=1}^{m}\sum_{i=1}^{n} R_{ij} = \frac{mn(m+1)}{2} \qquad 4.2$$

При малых значениях m и n тип критерий $\chi 2$ дает слишком грубое приближение, и при этом возможно принятие неправильного решения. Поэтому критерий $\chi 2$ применяется в том случае, когда выполняются следующие условия: $m = 3$ и $n > 9$ или $m = 4$ и $n > 4$ или $m > 4$, $n > 9$. Если эти условия не выполняются, то проверка осуществляется по критерию Фридмана.

Пример

Рассмотрим вариант примера, использованного: у нас есть 4 вида погодных условий, для каждого из которых проведено 5 экспериментов. Необходимо выяснить, значимо ли влияет изменение погодных условий на результаты экспериментов. Все исходные данные и результаты работы приведены на рис.4.4.

Сначала необходимо построить ранги по столбцам. Для этого в ячейке B8 набираем вызов функции построения рангов =Rankl(B3; B$3:B$6;1), а затем размножаем ее перетягиванием на ячейки от B8 до F11.

Находим суммы рангов по строкам и помещаем их в столбец G. Для этого в G8 набираем формулу =СУММ(B8:F8), а потом размножаем ее перетягиванием на остальные ячейки столбца. Здесь желательно выполнить проверку расстановки рангов. Для этого в ячейке G12 формируем сумму рангов (=СУММ(G8:Gll)), а в ячейке G13 - проверочное значение, которое определяется по формуле =(ЧСТРОК(B9:B12)* ЧИСЛСТОЛБ(B9:P9)*(ЧСТРОК(B9:B12)+1))/2.

Если эти значения совпадают, то расчеты выполнены правильно.

Далее в столбце H формируем квадраты сумм рангов по строчкам. Для этого в ячейке H8 набираем формулу =G8*G8 и размножаем перетягиванием на остальные ячейки столбца.

Находим сумму квадратов =СУММ(H8:H11) и помещаем ее в ячейке H14.

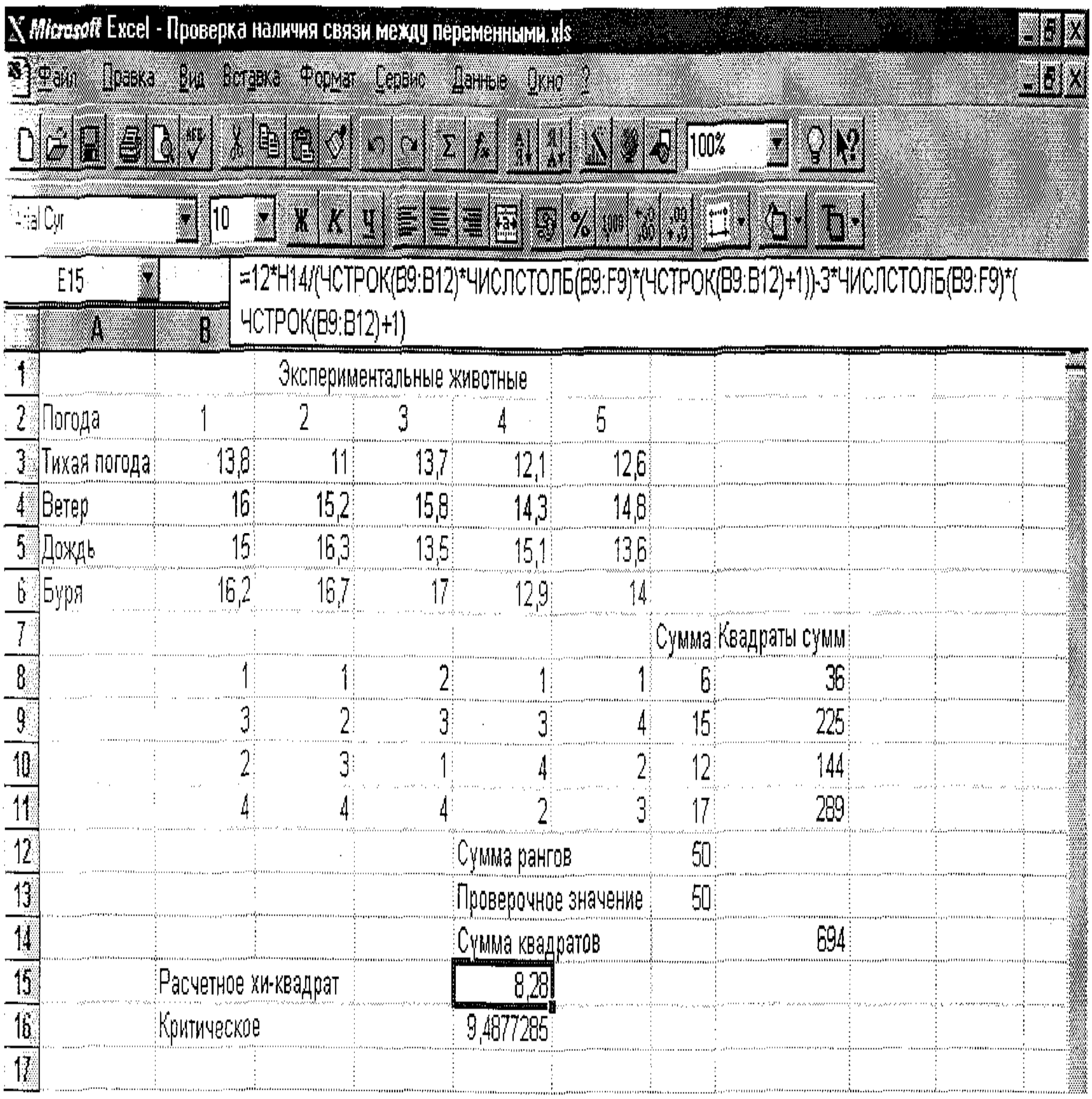
Теперь мы можем рассчитать критериальное значение по формуле:

= 12*H14/(ЧСТРОК(B8:B11)* ЧИСЛСТОЛБ(B8:P8)*(ЧСТРОК(B8:B11) +1))-3*ЧИСЛСТОЛБ(B8:P8)*(ЧСТРОК(B8:B11)+1)

и определить критическое значение критерия $\chi 2$ вызовом функции:

=ХИ2ОБР(0,05;ЧСТРОК(B8:B11-1)).

Поскольку расчетное критериальное (8,28) меньше критического (9,4877), то принимается гипотеза об отсутствии значимого влияния погодных условий на результаты эксперимента.



Microsoft Excel - Проверка наличия связи между переменными.xls

E15 =12*H14/(ЧСТРОК(B9:B12)*ЧИСЛСТОЛБ(B9:F9)*(ЧСТРОК(B9:B12)+1))-3*ЧИСЛСТОЛБ(B9:F9)*(ЧСТРОК(B9:B12)+1)

	A	B	C	D	E	F	G	H
1			Экспериментальные животные					
2	Погода	1	2	3	4	5		
3	Тихая погода	13,8	11	13,7	12,1	12,6		
4	Ветер	16	15,2	15,8	14,3	14,8		
5	Дождь	15	16,3	13,5	15,1	13,6		
6	Буря	16,2	16,7	17	12,9	14		
7							Сумма	Квадраты сумм
8		1	1	2	1	1	6	36
9		3	2	3	3	4	15	225
10		2	3	1	4	2	12	144
11		4	4	4	2	3	17	289
12					Сумма рангов		50	
13					Проверочное значение		50	
14					Сумма квадратов			694
15		Расчетное хи-квадрат			8,28			
16		Критическое			9,4877285			
17								

Рисунок 4.4 – Результат работы программы

Задание. Произвести дисперсионный анализ данных, представленных в работе 3.

Контрольные вопросы

1. В чем заключается основная идея метода дисперсионного анализа?
2. Влияет ли изменение диапазонов варьирования изучаемых факторов на результаты многофакторного дисперсионного анализа?
3. Из каких составляющих складывается оценка «общей» дисперсии случайной величины?
4. Какой статистический критерий используют для оценки влияния факторов на изучаемый технологический процесс?

Лабораторная работа №5 Экспериментальный анализ одномерной и двумерной совокупности случайных величин

Цель работы: по заданной выборке исследовать свойства одномерной и двумерной случайной величины.

Понятие «эксперимент» означает действие, направленное на создание условий в целях осуществления того или иного явления и по возможности наиболее частого, т.е. не осложняемого другими явлениями. Основной целью эксперимента являются выявление свойств исследуемых объектов, проверка справедливости гипотез и на этой основе широкое и глубокое изучение темы научного исследования.

Однофакторный эксперимент предполагает: выделение нужных факторов; стабилизацию мешающих факторов; поочередное варьирование интересующих исследователя факторов.

Порядок выполнения работы

По заданной выборке исследовать свойства одномерной случайной величины:

- построить вариационный ряд;
- построить диаграмму накопленных частот;
- построить гистограмму выборки;
- определить оценки математического ожидания, дисперсии и среднего квадратичного отклонения.

1 Экспериментальный анализ одномерной совокупности

Пусть имеется набор (выборка) экспериментальных данных: x_1. $x_2,...x_N$ Обработку этих данных для получения эмпирических характеристик одномерной случайной величины производят обычно в следующей последовательности.

1 Построение вариационного ряда (ряда распределения)

Вариационный ряд $z_1, z_2,...z_N$ получают из исходных данных путем расположения x_m $(m = 1, 2,..., N)$ в порядке возрастания от x_{min} до x_{max} так, чтобы

$$x_{\min} = z_1 \leq z_2 \leq \ldots \leq z_N = x_{\max} \qquad 5.1$$

Пример. Пусть имеется выборка из 10 наблюдений (N = 10):

$x_1 = 5$, $x_2 = 2$, $x_3 = 4$, $x_4 = 5$, $x_5 = 7$, $x_6 = 3$, $x_7 = 6$, $x_8 = 8$. $x_9 = 3$, $x_{10} = 9$.

Тогда им соответствует вариационный ряд

$z_1 = 2$, $z_2 = 3$, $z_3 = 3$, $z_4 = 4$, $z_5 = 5$, $z_6 = 5$, $z_7 = 6$, $z_8 = 7$, $z_9 = 8$, $z_{10} = 9$.

3.2 Построение гистограммы выборки

Гистограмма $f_N(x)$ является эмпирическим аналогом функции плотности распределения $f(x)$ Этапы построения гистограммы:

1.1. Нахождение предварительного количества интервалов, на которое должна быть разбита ось O_x. Это количество K определяется с помощью оценочной формулы:

$$K = 1 + 3{,}2\ 1\mathrm{g}N \qquad 5.2$$

Найденное значение необходимо округлить до ближайшего целого.

Пример: $K = 1 + 3{,}2\mathrm{Igl}0 = 1 + 3{,}2 = 4{,}2 = 5$

1.2. Определение длины интервала:

$$\Delta x = (x_{\max} - x_{\min}) / K \qquad 5.3$$

Величину Δx можно округлить для удобства вычислений.

Пример: Δx = (9 - 2)/5 = 7/5 = 1,4

1.3. Середину области изменения выборки (центр распределения) $(x_{\max} + x_{\min})/2$ принимают за центр некоторого интервала, после чего определяют границы и окончательное количество указанных интервалов так, чтобы в совокупности они *перекрывали всю область от* $x_{\min}$.*до* $x_{\max}$

Пример: $(x_{\max} + x_{\min})/2$= (9+2)/2 = 5,5 - центр распределения выборки. Для данного примера при выбранном числе интервалов - это середина третьего интервала. Следовательно, слева и справа от этого значения необходимо отложить по 2,5 интервала или 2,5Δx.

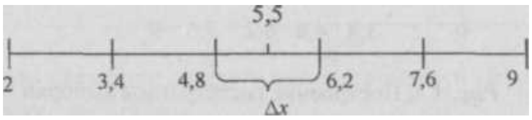


Рисунок 5.1 Разбиение оси Ох на интервалы

1.4. Подсчет количества наблюдений N_m, попавших в каждый интервал: N_m равно числу членов вариационного ряда, для которых справедливо неравенство

$$x_m \le z_1 < x_m + \Delta x \qquad 5.4$$

где x_m и $x_m + \Delta x$ - границы m-го интервала.

Значения z_1, попавшие на границу между (m-1) и m интервалами, относят к m интервалу.

Пример. В первый интервал попало три значения z_1 (N_1 = 3), во второй - одно значение z_1 (N_2 =1), в третий - три значения z_1 (N_3= 3), в четвертый - одно значение z_1 (N_4 = 1), в пятый - два значения z_1 (N_5 = 2)

1.5. Подсчет относительного количества наблюдений N_m / N, попавших в данный интервал.

Пример: N=10; *$N_1 / N.=3/10=0.3$; $N_2 / N=1/10=0.1$; $N_3 / N=3/10=0.3$; $N_4 / N=1/10=0.1$; $N_5 / N=2/10=0.2$.*

1.6. Построение гистограммы, которая представляет собой ступенчатую кривую, значение которой на *m* интервале $(x_m, x_m + \Delta x)$ постоянно и равно N_m / N.

Пример. По результатам предыдущих этапов строим гистограмму для данного примера.

1.7 Определение оценок математического ожидания $\bar{x}$, дисперсии S_x^2 и среднего квадратичного отклонения S_x

$$\bar{x} = (1/N)\sum_{i=1}^{N} x_i \qquad 5.5$$

$$S_x^2 = \left[\sum_{i=1}^{N}\left(x_i - \bar{x}\right)^2\right]/(N-1) \qquad 5.6$$

$$S_x = +\sqrt{S_x^2} \qquad 5.7$$

Пример. Для данной выборки искомые значения определяются следующим образом:

$$\bar{x} = (1/10)\sum_{i=1}^{10} x_i = (2+3+3+4+5+5+6+7+8+9)/10 = 52/10 = 5.2;$$

$$S_x^2 = \left[\sum_{i=1}^{10}\left(x_i - 5.2\right)^2\right]/(10-1) = \left[(2-5.2)^2 + (3-5.2)^2 + (3-5.2)^2 +\right]$$
$$\left[+ (4-5.2)^2 + (5-5.2)^2 + (5-5.2)^2 + (6-5.2)^2 + (7-5.2)^2 +\right]$$
$$\left[+ (8-5.2)^2 + (9-5.2)^2\right]/9 = (1.79 + 4.84 + 4.84 + 1.44 + 0.04 + 0.04 +$$
$$+ 0.64 + 3.24 + 7.84 + 14.44)/9 = 39.15/9 = 4.35;$$

$$S_x = +\sqrt{S_x^2} = \sqrt{4.35} = 2.09/$$

2 Экспериментальный анализ двумерной совокупности

Пусть получена выборка из двумерной совокупности при наблюдении двух случайных величин X и Y.

Таблица 5.1

i	1	2	3	…	i	…	N
X	x_1	x_2	x_3	…	x_i	…	x_N
Y	y_1	y_2	y_3	…	y_i	…	y_N

В этом случае обработку результатов наблюдений можно осуществлять по следующим этапам.

1.1 Построение поля рассеяния

Для этого на плоскости с координатами x, y отмечают экспериментальные точки.

1.2. Составление таблицы двумерного распределения

Таблица составляется следующим образом. Оси O_x и O_y разбивают на отдельные интервалы длиной Δx и Δy. Величины $\Delta x, \Delta y$ количество K_x, K_y (обычно $K_x = K_y = K$, так как число точек общее для X и $У$) и размещение этих интервалов для каждой из переменных находят с помощью правил, изложенных для экспериментального анализа одномерной случайной величины. Соответствующие границы наносят на диаграмму рассеяния и затем подсчитывают количество точек $N_{m1,m2}$ попавших в каждый из образовавшихся прямоугольников (если какая-либо точка расположена на границе, то ее относят к правому или верхнему прямоугольнику).

Следующий шаг - составление таблицы, в которой отмечают величины $N_{m1,m2}$, а также относительные величины $N_{m1,m2}/N$.

Таблица 5.2

Интервалы для y / Интервалы для x	$[y_1, y_1 + \Delta y]$	$[y_2, y_2 + \Delta y]$		$[y_{m2}, y_{m2} + \Delta y]$		$[y_{k2}, y_{k2} + \Delta y]$
$[x_1, x_1 + \Delta x]$			…			
$[x_2, x_2 + \Delta x]$					…	
					…	
$[x_{m1}, x_{m1} + \Delta x]$			…	$N_{m1\,m2}$	…	
				$N_{m1\,m2}/N$		
					…	
$[x_{k1}, x_{k1} + \Delta x]$						

С помощью таблицы двумерного распределения можно получить исходные данные для построения гистограмм, соответствующих каждой из двух одномерных случайных величин X и Y. Для этого достаточно просуммировать значения таблицы либо по каждому столбцу (при построении гистограммы для $У$), либо по каждой строке (при построении гистограммы для X).

1.3 Вычисление оценки коэффициента корреляции

Вычисление коэффициента корреляции производят по формуле

$$\rho_{xy} = \frac{1}{(N-1)S_x S_y} \sum_{i=1}^{N} (x_i - \bar{x})(y_i - \bar{y}) \qquad 5.8$$

где $\bar{x}, \bar{y}$ и S_x, S_y находят с помощью формул приведенных при определении оценок математического ожидания $\bar{x}$, дисперсии S_x^2 и среднего квадратичного отклонения S_x.

Между значением и знаком коэффициента корреляции, с одной стороны, и видом диаграммы рассеяния, с другой, существует определенная связь. Если начало координат перенести в точку ($\bar{x}, \bar{y}$), то:

при $p > 0$ точки на диаграмме рассеяния группируются в основном в I и III квадрантах;

при $p < 0$ - во II и IV квадрантах;

при $p = 0$ точки беспорядочно разбросаны во всех четырех квадрантах;

при $p = \pm 1$ точки группируются на прямых (находящихся в 1 и III или во II и IV квадрантах).

Задание на лабораторную работу

Исходные данные (выборка экспериментальных данных):

1. 17, 15, 17, 11, 13, 14, 11,13, 14, 17.
2. 10, 8, 9, 4, 9, 9, 1, 6, 6, 10.
3. 5, 8, 2, 7, 4, 1, 1, 9, 7, 10.
4. 7, 4, 1, 1, 9, 7, 10, 8, 2, 5.
5. 12, 11, 12,15, 19, 18, 11,12,19, 17.
6. 12, 19, 17, 15, 18, 17,18,18, И, 11.
7. 23, 26, 23, 20, 26, 20, 24, 25, 25, 21.
8. 9, 9, 1, 6, 6, 10, 5, 8, 2, 7.
9. 37, 31, 35, 32, 33, 35, 33, 32, 32, 37.
10.32, 33, 35, 33, 32, 32, 37, 35, 31, 33.
11.35, 32, 33, 35, 33, 32, 32, 37, 35, 31.
12.6, 10, 5, 8, 2, 7, 4, 1, 1, 9.
13.19, 18,11,12,19, 17, 15,18, 17,18.
14.29, 28, 21, 22, 29, 27, 25, 28, 27, 18.
15.20, 24, 21, 26, 24, 27, 23, 24, 20, 25.
16.22, 23, 28, 25, 29, 29, 23, 23, 27, 22.

Контрольные вопросы:

1. В каких случаях целесообразно использование однофакторного эксперимента?
2. Какие задачи могут решаться в результате обработки результатов исследований однофакторного эксперимента?
3. Какие могут быть виды связей между исследуемыми параметрами однофакторного эксперимента?
4. В каком виде представляют результаты экспериментальных данных однофакторного эксперимента?
5. каким образом проводят графическую обработку результатов исследований однофакторного эксперимента с помощью равномерных и функциональных шкал?
6. Каким образом определяется характер уравнения связи между экспериментальными данными в результате графической обработки однофакторного эксперимента?

Лабораторная работа №6 Обработка результатов однофакторных экспериментов методом наименьших квадратов

Цель работы: изучить методы нахождения коэффициентов уравнения регрессии для аппроксимации опытных данных линейной зависимостью и показательной функцией.

Целью математической обработки результатов эксперимента является не нахождение истинного характера зависимости между имеющимися переменными, а преставления результатов наблюдений в виде простой формулы, дающей возможность производить интерполирование.

Определение зависимостей между исследуемыми переменными следует начинать с графического изображения результатов эксперимента. Графики, прежде всего строят в равномерных шкалах, по возможности с равными масштабами.

Первый этап обработки данных – изучение характера полученных кривых. Если характер связи между исследуемыми величинами известен, то первый этап обработки данных сводится к проверке совпадения полученных данных с заданной кривой.

Если предварительные сведения о характере уравнения отсутствуют, то первым этапом обработки данных является нахождение определенной кривой, совпадающей с опытными точками. Эта задача решается методом подбора. Причем необходимо выбирать кривую с наиболее простым уравнением.

Уравнение зависимости между исследуемыми величинами наиболее просто определяется тогда, когда эмпирические точки, достаточно хорошо опадают с прямой линией, т.е. апроксимируются уравнением вида:

$$y=ax+b \qquad 6.1$$

Нелинейные зависимости путем замены переменных и графического изображения в функциональных шкалах так же можно привести к линейным, например, зависимость:

$$y=ax^m \qquad 6.2$$

путем замены переменных

$$y_1=lgy \qquad 6.3$$

$$x_1=lgx$$

$$a_1=lga$$

приводится к виду

$$y_1=mx+a_1 \qquad 6.4$$

Коэффициенты данного уравнения определяются методом наименьших квадратов, т.е. чтобы сумма квадратов отклонений между экспериментальными и вычисленными по уравнению значениям параметра была минимальной

$$R = \sum_{i=1}^{N} \varepsilon_i^2 = R_{\min} \qquad 6.5$$

Чтобы определить наиболее вероятные значения коэффициентов в исходном уравнении необходимо решить систему так называемых нормальных уравнений, которая для уравнения 6.1 имеет вид:

$$\begin{cases} a\sum_{i=1}^{N} x_i^2 + b\sum_{i=1}^{N} x_i - \sum_{i=1}^{N}(x_i y_i) = 0 \\ a\sum_{i=1}^{N} x_i + bN - \sum_{i=1}^{N} y_i = 0 \end{cases}$$

Для уравнения 6.2:

$$\begin{cases} m\sum_{i=1}^{N}(\lg x_i)^2 + a\sum_{i=1}^{N}\lg x_i - \sum_{i=1}^{N}(\lg x_i \lg y_i) = 0 \\ m\sum_{i=1}^{N}\lg x_i + aN - \sum_{i=1}^{N}\lg y_i = 0 \end{cases}$$

Проверку адекватности полученной модели проводят по F-критерию Фишера. Если принятая аппроксимирующая функция не удовлетворяет критерию Фишера, то она должна быть заменена другой.

Определение линейной зависимости результатов эксперимента от продолжительности обработки

1. В результате экспериментальных исследований установлено, что данные эксперимента в зависимости от продолжительности обработки имеют следующие значения (данные берутся из таблицы 6.1):

R_z						
T, мин						

Опыт при каждом значении T проводился n=5, что соответствует уровню доверительной вероятности 0,95.
Усредненная дисперсия всех опытов S^2=0,454.

Таблица 6.1

Зависимость результатов эксперимента от продолжительности обработки

Продолжительность, Т, мин	Последняя цифра шифра									
	0	1	2	3	4	5	6	7	8	9
0	10	15	13	20	5	9,6	8	10	8	15
5	13	20	13,5	26	15	12,6	11,2	12,1	10	175
10	15	24	15	30	25	16	14,1	14	12	20
15	18	30	17	36	35	17	17	16	14	22
20	21	35,5	18	42	45	20,6	20	17,9	16	25
25	24	41	20	28	55	225	23	20	18	275
30	25	45	21	49	65	25	25,8	22	20	30
35	28	50	21,5	56	75	26,5	29	24	22	32,5
40	30	56	23	60	85	29,5	21,7	26	24	35
45	34	60	24	67	90	32,5	35	28	26	32,5

2. В координатах R_z-T наносим экспериментальные точки

Рисунок 6.1 Зависимость результатов эксперимента от продолжительности обработки

Данную зависимость можно аппроксимировать линейным уравнением:

$$R_z = aT + b \qquad 6.6$$

3. Вычисляем коэффициенты данного уравнения. Система нормальных уравнений:

$$\begin{cases} a\sum_{i=1}^{10} T_i^2 + b\sum_{i=1}^{10} T_i - \sum_{i=1}^{10}(R_{zi}T_i) \\ a\sum_{i=1}^{10-} T_i + 10b - \sum_{i=1}^{10} R_{zi} = 0 \end{cases}$$

Вычисления a и b сводим в таблицу

№	T	T^2	R_z	R_z T	Проверка адекватности		
1					R_{zp}	$\lvert R_{zp}-R_z \rvert$	$\lvert R_{zp}-R_z \rvert^2$
2							
3							
4							
5							
6							

7							
8							
9							
10							
$\sum$					-	-	

Затем данная система решается одним из известных способов и определяется коэффициенты a и b.

4. проверка адекватности уравнения регрессии.

По полученному уравнению вычисляем R_{zp} и заносим в таблицу

Дисперсия адекватности S^2ад рассчитывается по формуле:

$$S^2_{АД} = \frac{\sum_{i=1}^{N} \left| R_Z - R_{ZP} \right|^2}{N-1} \qquad 6.7$$

Где N=10, число серий опытов.

Значение F-критерия Фишера

$$F = \frac{S^2_{ад}}{S^2_y} \qquad 6.8$$

По ПРИЛОЖЕНИЮ для f_1=N-1, f_2=(n-1)N

Определяем табличное значение критерия Фишера – F_p.

Делаем вывод об адекватности полученного уравнения.

5. Строим аппроксимирующую кривую.

Определение зависимости результатов эксперимента от продолжительности обработки показательной функцией

1. В результате экспериментальных исследований установлено, что данные эксперимента в зависимости от продолжительности обработки имеют следующие значения (данные берутся из таблицы 6.1 по последней цифре шифра +4):

P_z						
S, мин						

Опыт при каждом значении Т проводился n=5, что соответствует уровню доверительной вероятности 0,95.

Усредненная дисперсия всех опытов S^2=0,56

2. В координатах P_z-S наносим экспериментальные точки

Рисунок 6.2 Зависимость результатов эксперимента от продолжительности обработки

Данную зависимость можно аппроксимировать степенной функцией

$$P=CS^m \qquad 6.9$$

3. Вычисляем коэффициенты данного уравнения. Система нормальных уравнений:

$$\begin{cases} m\sum_{i=1}^{10}\lg S_i^2 + \lg C\sum_{i=1}^{10} l_y S_i - \sum_{i=1}^{10}(\lg P_i \lg S_i) \\ m\sum_{i=1}^{10-}\lg S_i + 10C - \sum_{i=1}^{10}\lg P_i = 0 \end{cases}$$

Вычисления c и m сводим в таблицу

№	lgS	lgS^2	lgP	lgP lgS	Проверка адекватности			
1					P	P_p	$\|P_p-P\|$	$\|P_p-P\|^2$
2								
3								
4								
5								
6								
7								
8								
9								
10								
$\sum$							-	

Затем данная система решается одним из известных методов и определяется коэффициенты *c* и *m*.

4. Проверяем адекватность полученного уравнения регрессии.
5. Строим аппроксимирующую кривую.

Контрольные вопросы:

1. Каким образом представляют математически зависимости результатов обработки однофакторных экспериментов в функциональных шкалах после определения характера уравнения в результате графической обработки?
2. Каким образом осуществляют математическую обработку функциональной зависимости результатов исследований однофакторных экспериментов?
3. Каким образом находят общую формулу для выражения математической зависимости однофакторных экспериментов при влиянии нескольких факторов?
4. Какие ошибки наблюдений при статистической обработке экспериментальных данных для одного фактора Вам известны?
5. Дайте характеристику основных показателей случайных ошибок однофакторных экспериментов: среднего, математического ожидания, дисперсии, плотности нормального распределения?

6. Каким образом, пользуясь экспериментальными данными можно построить эмпирическую кривую распределения однофакторного эксперимента?

Лабораторная работа №7 Обработка результатов однофакторных экспериментов методом средних и методом Брандона

Цель работы: обработка результатов экспериментов методом средних и методом наименьших квадратов, метод последовательного исключения влияния независимых переменных (метод Брандона).

Метод средних

Метод средних позволяет определить аналитическое выражение для кривой регрессии — уравнение регрессии, исходя из требования равенства нулю суммы отклонений экспериментально полученных $y_{Эi}$ и y_{Pi} вычисленных из искомого уравнения

$$\sum_{i=1}^{N}(y_{эi} - y_{pi}) = 0 \qquad 7.1$$

Последовательность операций при расчете по методу средних следующая.

1. Производится выбор вида уравнения регрессии. При этом могут быть использованы все функции, приводимые к линейным, степенные многочлены вида $y=a_0+ + a_1x+a_2x^2 + a_3x^3 + ... + a_nx^n$ и другие функции.

2. Экспериментальные значения у, х разбиваются на группы, число которых равно числу коэффициентов a_i подлежащих определению. Число экспериментальных пар у, х в каждой из групп лучше брать одинаковым. Если общее число экспериментов не кратно числу групп, то некоторые пары у, х могут быть записаны в несколько групп.

3. В каждой из групп подсчитывается сумма отклонений, выражение которой приравнивается нулю. В качестве y_p используется выбранная функция, т. е. $y_{pi} = = \varphi(a_0, a_1, a._n, x_i)$, которая называется также аппроксимирующей. Суммирование производится по числу

пар значений y, x в группе. При раскрытии выражения суммы отклонений пользуются следующими правилами:

$$\sum_{i=1}^{N}(y_i + x_i) = \sum_{i=1}^{N} y_i + \sum x_i \qquad 7.2$$

$$\sum_{i=1}^{N} a x_i = a \sum_{i=1}^{N} x_i \qquad 7.3$$

$$\sum_{i=1}^{N} a_i = Na \qquad 7.4$$

Найденную таким образом систему уравнений решают относительно подлежащих определению коэффициентов аппроксимирующей функции.

Пример. Воспользуемся методом средних для математического описания технологического процесса остывания хлеба. В работе показано, что интенсивность отвода тепла с поверхности хлеба q меняется во времени по сложной зависимости. Чтобы линеаризовать ее, попробуем в качестве y принять не q, и $lg\ q$. Экспериментальные данные для верхней поверхности каравая подового хлеба, лежащего на верхней полке этажерки, приведены на рис. 7.1, они показывают, что, начиная с четвертого измерения, можно ожидать удовлетворительного описания линейной зависимости.

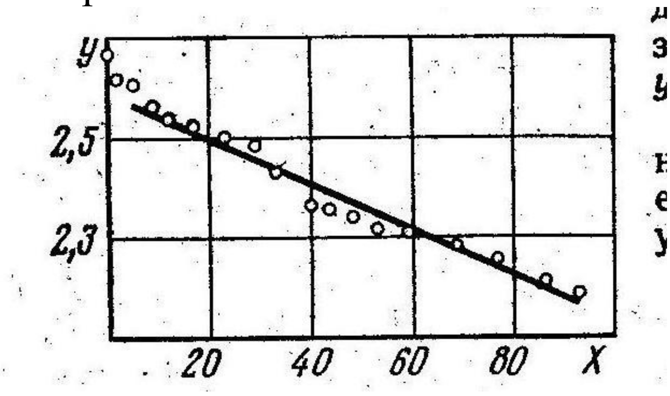


Рисунок 7.1 – К примеру использования метода средних

Этот момент соответствует началу известного в теории теплопроводности регулярного режима охлаждения. Таким образом, Для аппроксимации зависимости y(x), где в качестве x выберем τ —время процесса (в мин) от начала охлаждения, примем линейную функцию

$$y=a_0+ a_1x \tag{7.5}$$

Общее числом экспериментальных точек, начиная с четвертой, N=15, столько же пар значений x и y. Разобьем эти значения на две группы: с 1-го по 8-е и с 8-го по 15-е. Сгруппированные экспериментальные данные представлены в таблице 7.1, здесь же приведены результаты суммирования значений аргумента x и функции y по каждой группе.

Таблица 7.1

Апроксимация по методу средних

k	$x_э$	$y_э$	k	$x_э$	$y_э$
1	9	2,564	8	43	2,356
2	12	2,539	9	48	2,342
3	17	2,527	10	53	2,320
4	23	2,502	11	59	2,316
5	29	2,490	12	69	2,286
6	33	2,434	13	77	2,255
7	40	2,367	14	86	2,214
8	43	2,356	15	93	2,294
	$\sum x=206$	$\sum y=19{,}78$	$\sum x=528$	$\sum y=18{,}28$	

Теперь подсчитаем отклонения в каждой группе и приравняем их нулю, т. е. решим систему уравнений

$$\left.\begin{aligned}\sum_{k=1}^{8}(y_э-a_0-a_1x_э)=0\\ \sum_{k=1}^{15}(y_э-a_0-a_1x_э)=0\end{aligned}\right\}$$

Подставляя в систему подсчитанные суммы из табл. 15, получим

$$\left.\begin{aligned}19{,}78-8a_0-206a_1=0\\ 18{,}28-8a_0-528a_1=0\end{aligned}\right\}$$

откуда a_0=2,59; a_1=-0,465·10^{-2}

Следовательно, уравнение аппроксимирующей функции имеет вид

$\lg q = 2{,}590 — 0{,}456 \cdot 10^{-2}\ \tau$.

Рассмотренный метод дает грубую оценку параметров уравнения регрессии и может применяться для предварительной количественной оценки уравнения связи между величинами у и х.

Метод последовательного исключения влияния независимых переменных (метод Брандона)

Рассмотрим подробнее одну из разновидностей регрессионного анализа – метод последовательного исключения влияния независимых переменных, известный под названием метода Брандона. В его основе лежит предположение .о том, что влияющие факторы слабо связаны между собой, и искомое уравнение связи можно представить в виде

$$y = \bar{y}\, f_1(x_1) f_2(x_2) f_3(x_3) \ldots\ldots\ldots f_n(x_n) \qquad 7.5$$

где $\bar{y}$ — среднее значение выхода (функции).

Этот метод наиболее применим в случае, когда на основании предварительных исследований примерно известен качественный характер влияния каждой из независимых переменных х на исследуемую функцию. Однако он может быть применен и в тех случаях, когда такие сведения отсутствуют. Удобство указанного представления аппроксимирующей функции состоит в том, что каждую из составляющих функций, которая зависит только от одной переменной, можно легко оценить.

Аппроксимация начинается с того, что каждое из экспериментальных значений y_i изображается графически как частное значение функции одного переменного, например x_i. Построение выполняется по экспериментальным данным. Совокупность построенных точек дает возможность оценить правильность выбора вида функции $f_1\ (x_1)$ и если это нужно изменить ее на более подходящую. Далее методом наименьших квадратов или любым другим определяются коэффициенты первой составляющей функции.

Пользуясь функцией $f_1(x_1)$, определенной указанным способом, рассчитывают ее частные значения, соответствующие экспериментальным значениям x_{1i}. Найденные частные значения $f_1(x_{1i})$, исключают из экспериментальных значений y_i при этом получаются частные значения некоторой функции *n-1* переменных, уже не зависящей от x_1.

Пример Воспользуемся опытными данными определения времени прогреваемости вытяжек консервов в капиллярах различного диаметра *d*. Данные приведены для консервов «Бычки в томатном соке» (содержание, жира Ж=6,7%, pH 5,2) и «Сардины в томатном соусе» (Ж=9,3%. pH 5,48) при температуре окончания прогрева 112, 115, 118 и 121,5° С. Принимаем $y=\tau$ (в с); $x_1=d$ (в мм); $x_2=t$ (в °С),, В качестве x_3 выберем жирность Ж (в %), поскольку кислотность консервов по теоретическим соображениям не должна влиять на время прогрева жидкости.

Исходные данные представлены таблице 7.2.

Таблица 7.2

ВРЕМЯ ПРОГРЕВАЕМОСТИ ВЫТЯЖЕК КОНСЕРВОВ

x_2	112		115		118		121,5	
x_1 \ x_3	6,7	9,3	6,7	9,3	6,7	9,3	6,7	9,3
1,8	10,5	11,8	10,2	11,5	10,0	11,3	9,6	11,1
1,9	11,5	13,0	10,8	12,6	10,6	12,3	10,3	12,0
2,0	13,8	13,2	13,1	12,9	12,9	12,6	12,4	12,1
2,1	16,0	14,0	15,3	13,7	15,0	13,3	14,3	13,2
2,2	16,6	15,6	16,3	15,4	16,0	15,0	15,7	14,2
2,3	16,9	16,1	16,6	15,5	16,2	15,3	15,8	14,6

Среднее значение выходного параметра $\bar{y} = \sum_{1}^{48} y 48^{-1} = 13{,}5$

Рассчитывая φ_i как $y_i / \bar{y}$, наносим данные на график (рис. 1.2,а). Видно возрастание $f_i(x_1)$ с ростом x_1. Принимаем линейное приближение и методом наименьших квадратов определяем a_1 = — 0,59; b_1=0,77. Вычисляем значения $f_i(x_i) = 0{,}77x_1 - 0{,}59$,делим φ_1 на $f_i(x_1)$ результат заносим в таблицу 7.3.

Таблица 7.3

ЗНАЧЕНИЯ ФУНКЦИИ φ_2

x_1 \ x_2	112		115		118		121,5	
x_3	6,7	9,3	6,7	9,3	6,7	9,3	6,7	9,3
1,8	0,98	1,10	0,95	1,07	0,93	1,06	0,90	1,04
1,9	0,97	1,10	0,92	1,07	0,90	1,04	0,87	1,02
2,0	1,08	1,03	1,02	1,01	1,01	0,98	0,97	0,94
2,1	1,15	1,01	1,10	0,99	1,08	0,96	1,04	0,95
2,2	1,11	1,05	1,09	1,03	1,07	1,01	0,99	0,95
2,3	1,06	1,01	1,04	0,97	1,01	0,96	0,99	0,91

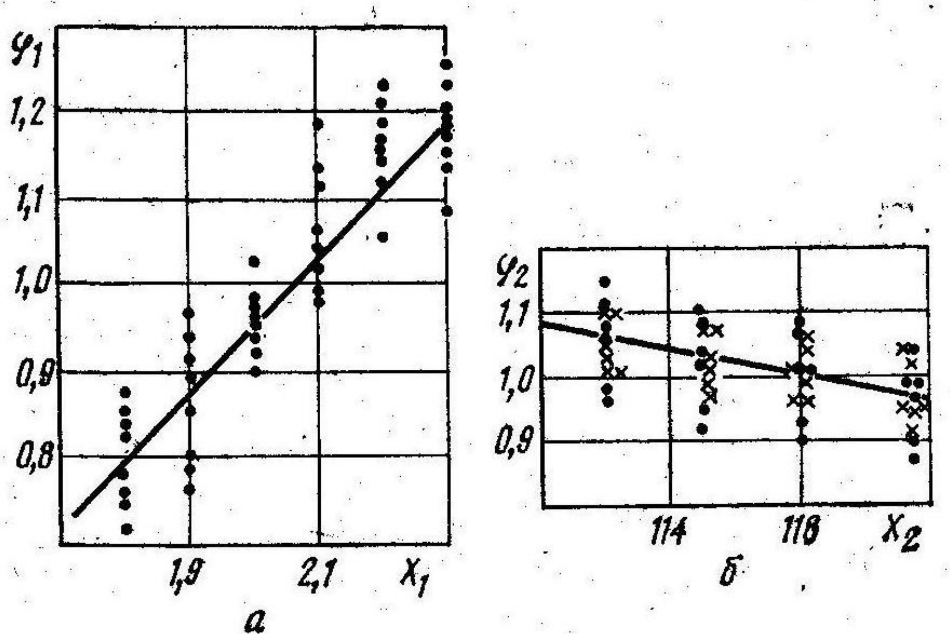


Рисунок 7.2 Пример использования метода Брандона
а – $\varphi_1(x_1)$; б – $\varphi_2(x_2)$;

Нанося значения φ_{2i} на график (рис. 7.2б) обозначим их по-разному для разных значений x_3(точками для x_3= 6,7крестиками для x_3 =9,3), чтобы выяснить вопрос о целесообразности определения φ_3. Из рисунка 7.26 видна, несмотря на большой разброс данных, четкая зависимость $f_2(x_2)$. Принимаем ее линейной: $f(x_2)$— $=a_2+b_2x_2$ и определяем a_2 и b_2:a_2=2,18; b_2= -0,01.

Данные по φ_2 для разных x_3 не разделяются на рис; 1.2, а, поэтому в связи с малым количеством данных по x_3 .и предположительно малым влиянием жирности консервов в таком интервале на прогреваемость вытяжки полагаем φ_2=1.

Таким образом, зависимость времени прогреваемости вытяжки консервов от диаметра капилляра d и температуры t можно выразить уравнением $\tau = 13{,}5 \cdot (0{,}77\ d - 0{,}59)(2{,}18 - 0{,}01\ t)$.

При использовании метода Брандона необходимо иметь в виду, что операции определения составляющих функций и их исключения должны выполняться с достаточно -высокой точностью, чтобы результирующая аппроксимирующая функция удовлетворительно отражала свойства исследуемой зависимости. Существенным является также вопрос о степени влияния отдельных переменных на ход процесса.

Местоположение каждой функции $f_i(x_i)$ в уравнении, а следовательно, и последовательность решения задачи определяются весомостью соответствующей переменной x (чем значительнее ее участие в процессе, тем меньше должен быть порядковый номер. Если же сравнительное влияние переменных x_1, x_2, x_n осталось неизвестным или выявлено неверно и, следовательно, неправильно задана структура уравнения, использование изложенного метода может привести к ложным выводам.

Задание 1 Рассчитать методом средних и методом наименьших квадратов данные эксперимента приведенные в таблице 7.4

Таблица 7.4

Варианты заданий

Вариант	Данные эксперимента			
1	345,4	343,2	340,3	333,1
2	48,3	49,5	50,1	53,0
3	12,1	17,2	18,0	19,2
4	75,8	76,0	77,0	77,2
5	66,4	68,0	69,0	70,2
6	34,5	38,7	44,8	54,9
7	64,1	62,1	61,9	60,3
8	301,0	299	298,5	297,1
9	233,1	233,9	234,1	234,8
10	255,1	256,2	256,3	257,0

Задание 2 Построить уравнение регрессии независимых переменных с помощью метода Брандона на основе данных представленных в таблице 7.5.

Таблица 7.5

Варианты заданий

Вариант		1						2			
x_2		48,3		49,5		50,1		49,5		50,1	
x_3		75,8	76,0	75,8	76,0	75,8	76,0	75,8	76,0	75,8	76,0
x_1	345,4	1,20	1,42	1,73	2,3	2,2	2,1	35,9	26,1	51	26,3
	343,2	2,3	2,37	2,39	2,33	2,37	2,34	35,7	26,0	26,3	26,1
	340,3	2,34	2,6	2,61	2,35	5,67	2,58	35,2	25,8	25	26
	333,1	5,63	5,62	2,8	2,70	2,35	2,67	33,6	23,8	23,4	26
	332	7,71	7,90	2,80	2,9	2,59	2,85	30	22	21	23,8
Вариант		3						4			
x_2		48,3		49,5		50,1		49,5		50,1	
x_3		75,8	76,0	75,8	76,0	75,8	76,0	75,8	76,0	75,8	76,0
x_1	345,4	3,57	3,6	2,9	2,8	2,55	2,6	31,7	22	21	23,8
	343,2	3,3	3,59	2,9	2,7	2,5	2,59	34,8	23,8	23,4	26
	340,3	1,45	1,6	2,7	2,2	2,45	2,47	36,0	25,8	25	26
	333,1	1,2	1,47	2,6	2	2,13	2,4	37	26,0	26,3	26,1
	332	1,13	1,15	2	2	1,7	2,16	38	26,1	51	26,3
Вариант		3						4			
x_2		48,3		49,5		50,1		49,5		50,1	
x_3		48,1	60,0	48,1	60,0	48,1	60,0	48,1	60,0	48,1	60,0
x_1	75,8	0,34	4,2	0,73	2,1	1,7	2,1	0,22	0,23	0,12	0,3
	66,4	0,34	4,2	0,74	2,2	1,8	2,04	0,22	0,23	0,13	0,4
	48,3	0,34	4,3	0,76	2,1	1,8	2,1	0,25	0,24	0,14	0,5
	34,5	0,36	4,3	0,77	2,08	1,8	2,04	0,25	0,24	0,15	0,48
	12,1	0,33	4,5	0,77	2,08	1,9	2,0	0,26	0,21	0,15	0,48

Вариант		5						6			
x_2		48,3		49,5		50,1		49,5		50,1	
x_3		48,1	60,0	48,1	60,0	48,1	60,0	48,1	60,0	48,1	60,0
x_1	75,8	7,71	7,90	6,2	6,6	5,7	4,1	0,22	0,21	0,12	0,3
	66,4	7,3	7,6	5,45	6,47	7,8	3,1	0,22	0,23	0,13	0,4
	48,3	6,20	6,42	5,3	5,6	6,8	3,04	0,25	0,23	0,14	0,48
	34,5	7,63	5,62	5,57	5,59	4,8	2,04	0,25	0,24	0,15	0,48
	12,1	6,34	5,37	4,13	4,15	5,9	2,0	0,26	0,24	0,15	0,5

Вариант	7			8	
x_2	48,3	49,5	50,1	49,5	50,1

x_3		48,1	60,0	48,1	60,0	48,1	60,0	48,1	60,0	48,1	60,0
x_1	75,8	12,1	7,90	12,16	6,6	12,2	4,1	86,1	0,21	85,7	0,3
	66,4	12,34	7,6	12,4	6,47	12,35	3,1	86,0	0,23	84,5	0,4
	48,3	12,58	6,42	12,47	5,6	12,37	3,04	79	0,23	79	0,48
	34,5	12,85	5,62	12,59	5,59	12,59	2,04	75,8	0,24	78	0,48
	12,1	14,67	5,37	12,6	4,15	15,67	2,0	73,8	0,24	78	0,5

Вариант		9				10			
x_2		48,3		49,5		49,5		50,1	
x_3		48,1	60,0	48,1	60,0	48,1	60,0	48,1	60,0
x_1	75,8	5,37	217,3	4,15	244,7	230	12,2	247,8	11,7
	66,4	5,62	233,9	5,59	250,9	230,5	12,35	255,7	11,7
	48,3	6,42	264,1	5,6	251,6	236,3	12,37	261,2	12,13
	34,5	7,6	281,0	6,47	263	255,2	12,59	287	12,45
	12,1	7,90	286	6,6	296	278,0	15,67	287	12,55

В качестве исходных данных в задании можете использовать свои экспериментальные данные.

Контрольные вопросы:

1. Что представляет собой пассивный эксперимент? Недостатки и достоинства пассивного эксперимента.
2. Каким образом проводится пассивный эксперимент для получения математической модели?
3. Какие требования предъявляются к выборке при проведении пассивного эксперимента?
4. Каким образом обрабатывают выборки, полученные в результате пассивных экспериментов?

Лабораторная работа № 8 Составление матрицы планирования эксперимента

Цель работы: приобрести навыки составления схем полного факторного эксперимента и симплекс-решетчатого плана

Планированию эксперимента предшествует анализ результатов предыдущих исследований (как собственных, так и описанных в литературных источниках). На основании этого анализа выбирается нулевой (основной) уровень, или исходный уровень «фона», т. е. такие значения факторов, в области которых значения параметра оптимизации наилучшие..

Значение фактора на нулевом уровне (уровне «фона») обозначается X_{i0} , т. е. X_{10}, X_{20}, .., X_{30}.

Далее, первоначально ставится серия опытов на двух уровнях - верхнем и нижнем. Для каждого фактора эти уровни отличаются от основного уровня (исходного уровня «фона») на одну и ту же величину, обозначаемую как ***интервал варьирования X***. Интервалы варьирования для каждого фактора, как правило, различны, они выбираются произвольно на основании предварительных данных о процессе.

Верхний X^+_i и нижний X^-_i уровни изучаемых факторов образуются в результате прибавления (вычитания) к основному (нулевому) уровню выбранного интервала варьирования:

$$X_i^+ = X_{i0} + \lambda_i \qquad 8.1$$

$$X_i^- = X_{i0} - \lambda_i \qquad 8.2$$

Таким образом, X_{i0} представляет собой среднее арифметическое между выбранными уровнями факторов:

$$X_{i0} = \frac{X_i^+ + X_i^-}{2} \qquad 8.3$$

Для облегчения последующих расчетов коэффициентов регрессии производят преобразование (кодирование) переменных факторов по формуле:

$$x_i^{+(-)} = \frac{X_i^{+(-)} - X_{i0}}{\lambda_i} \qquad 8.4$$

Величина X – значение фактора на верхнем или нижнем уровне в натуральных единицах; x – значение фактора в кодированных единицах.

В соответствии с уравнением изучаемые факторы в кодированных единицах будут равны:

$x_i^+ = +1$; $x_i^- = -1$.

Количество вариантов опытов, которые нужно поставить в исходной серии, зависит от количества исходных факторов. Всего на двух уровнях для n факторов количество вариантов опытов составит 2^n, т. е. для $n = 4$ факторов $2^4 = 16$, для $n = 8$ факторов $2^8 = 256$.

Чтобы приблизительно описать процесс линейным уравнением, необходимым для расчета крутого восхождения и для возможности статистической оценки адекватности этого уравнения, достаточно

поставить от (n + 2) до (n + 6) вариантов опытов. При этом важно, чтобы все варианты были разными и чтобы в них одинаковое число раз встречались верхние и нижние уровни каждого фактора.

Для упрощения расчетов и более точного определения направления движения к оптимуму, опыты ставятся по ортогональным матрицам.

Составление план-матрицы эксперимента составляется следующим образом: например для x_1 уровни чередуются в каждом опыте, для x_2 – через два опыта, для x_3 – через четыре и т.д.

Построенный таким образом план эксперимента называется ортогональным планом второго порядка. Основным преимуществом такого плана является независимая оценка коэффициентов регрессии.

В таблицах 8.1 и 8.2 приведены примеры наиболее часто встречающихся на практике ортогональные матрицы. В этих матрицах факторы обозначены буквой х с соответствующим индексом, верхний уровень (+1) - со знаком плюс, нижний уровень (-1) - со знаком минус. У – параметр оптимизации

Приведенные матрицы являются условиями проведения эксперимента, где строки соответствуют различным опытам, а столбцы значениям факторов. Такие таблицы называют матрицами планирования эксперимента.

Таблица 8.1

Матрица планирования эксперимента 2^2

№ опыта	x_1	x_2	У
1	-1	-1	$У_1$
2	+1	-1	$У_2$
3	-1	+1	$У_3$
4	+1	+1	$У_4$

Таблица 8.2

Матрица планирования эксперимента 2^3

№ опыта	x_1	x_2	x_3	У
1	- 1	- 1	-1	$У_1$
2	+ 1	-1	-1	$У_2$
3	-1	+1	-1	$У_3$
4	+ 1	+ 1	-1	$У_4$
5	- 1	- 1	+ 1	Y_5
6	+ 1	- 1	+ 1	$У_6$
7	-1	+1	+1	$У_7$
8	+ 1	+1	+1	$У_8$

После реализации экспериментов в соответствии с выбранной матрицей получаем некоторые значения параметра оптимизации, представляющие собой среднеарифметические величины. В этом случае уравнение математической модели имеет вид:

$$У = b_0 + b_1 x_1 + b_2 x_2 + \ldots + b_n x_n \qquad 8.4$$

где У – параметр оптимизации процесса;

b – экспериментальные оценки коэффициентов регрессии, показывающие влияние факторов на параметр оптимизации;

x_1, x_2, x_n – значения уровней факторов в кодированных единицах.

При симплексно-решетчатом планировании сумма всех компонент смеси равна 1 (100%) и это приводит к некоторым особенностям планирования, обработки опыта и представления результатов.

Рассмотрим специальную систему координат, в которой автоматически удовлетворяется связь между факторами:

$(0 \leq xi \leq 1)$ 8.5

$x_1 + x_2 + x_3 + \ldots + xq = 1$ 8.6

На рисунке изображена система координат для трехкомпонентной смеси (q = 3). Эта система представляет собой правильный треугольник, в вершинах которого одна из компонент смеси принимает максимально возможное значение xi = 1, на противоположной стороне xi = 0.

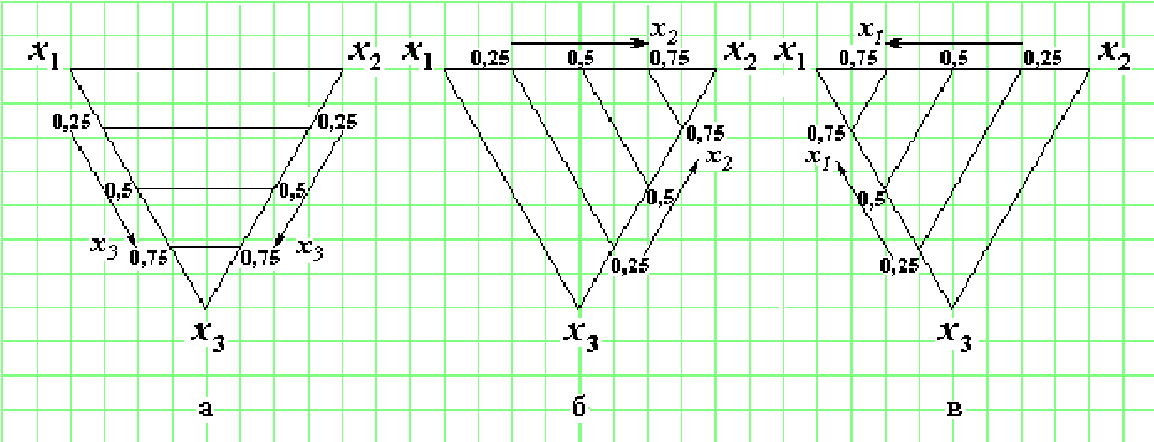


Рисунок 8.1 - Координатные линии для q=3

На рис.8.1(а) изображены координатные линии $x3$ = Const, эти линии параллельны основанию треугольника, на котором $x3$ = 0. Значение координаты $x3$ может быть отмечено на одной из двух боковых сторон. На рис.8.1 (б) изображены координатные линии $x2$ = Const, а на рис. 1(в) – координатные линии $x1$ = Const.

Значение каждой из координат может отмечаться на каких-то двух сторонах треугольника, но обычно значения каждой координаты отмечают на одной из сторон (рис.8.2). Стрелки с указанием направления возрастания координаты можно опустить, т.к. и так ясно, что координата возрастает к вершине (а вот вершины помечать обязательно).

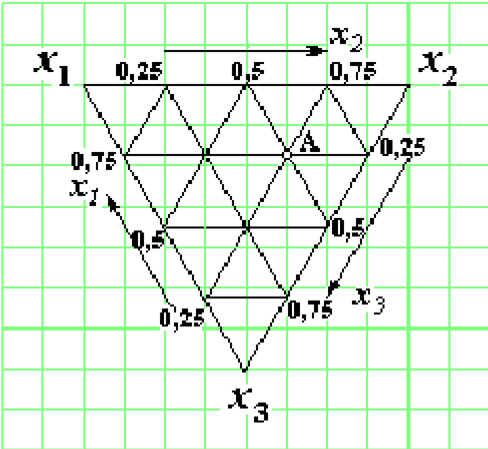


Рисунок 8.2 – Координатная сетка для q = 3

Вершины означают чистые компоненты, стороны – двухкомпонентные смеси (на стороне одна из компонент равна нулю), внутренность треугольника - тройные смеси. Например, точке А на рис. 2 соответствует смесь с компонентами $x1$ = 0,25; $x2$ = 0,5; $x3$ = 0,25.

Расположение опытов в факторном пространстве можно легко проследить на примере трехкомпонентной смеси. Известно, что состав трехкомпонентной смеси может быть представлен на треугольной диаграмме. При этом каждой точке диаграммы соответствует один, вполне определенный состав, а каждому составу соответствует одна точка на диаграмме. В каждой вершине треугольной .диаграммы содержание одного из компонентов составляет 100%.

Для построения линейной модели необходимо поставить три опыта (рис. 8.3, а). Четвертый опыт в центре ставится для проверки адекватности полученной модели. Цифры у точек соответствуют подстрочным индексам откликов. Количество цифр в индексе означает доли компонентов в смеси.

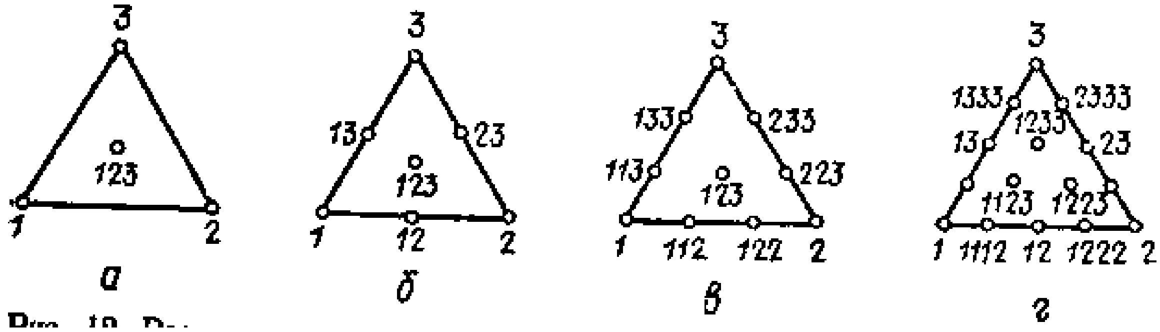


Рисунок 8.3 Расположение опытов для разных моделей при трех компонентах;
д — линейная модель; б — квадратная и специальная кубическая; в — кубическая; г — четвертой степени.

Число повторений каждой цифры равно количеству долей соответствующего компонента. При такой индексации план эксперимента для линейной модели при трех компонентах может быть задан таблицей 8.3.

Таблица 8.3

План эксперимента для линейной модели для трехкомпонентной смеси

Опыт	Содержание компонентов			Отклик
	x_1	x_2	x_3	
1	1	0	0	Y_1
2	0	1	0	Y_2
3	0	0	1	Y_3
4	1/3	1/3	1/3	Y_{123}

Таблица 8.4

План эксперимента для квадратичной модели для трехкомпонентной смеси

Опыт	Содержание компонентов			Отклик
	x_1	x_2	x_3	
1	1	0	0	Y_1
2	0	1	0	Y_2
3	0	0	1	Y_3
4	1/2	1/2	0	Y_{12}
5	1/2	0	1/2	Y_{13}
6	0	1/2	1/2	Y_{23}

Примеры составления матриц планирования эксперимента

Моделирование питательной среды культивирования пекарских дрожжей

Выход биомассы пекарских дрожжей определяется целым рядом факторов, среди которых первостепенное значение имеет концентрация солей азота и фосфора в среде культивирования и уровень биотина, который является стимулятором роста дрожжей.

Рассмотрим зависимость выхода биомассы от концентрации сульфата аммония X_1 диаммонийфосфата X_2 и биотина X_3 в питательной среде при выращивании сахаромицетов.

Принимаем следующую очередность операций:

1. На основании ранее выполненных исследований в качестве основного фона (нулевого уровня) выбираем следующие концентрации изучаемых факторов: X_{10} = 400 мг/л; X_{20} = 30 мг/л; X_{30} - 1 мкг/л. Интервалы варьирования для каждого соответственно равны: λ_1 = 200 мг/л; λ_2 = 20 мг/л; λ_3 = 1 мкг/л.
2. Находим верхний и нижний уровни изучаемых факторов:

$$X_1^+ = 400 + 200 = 600 мг/л$$

$$X_1^- = 400 - 200 = 200 мг/л$$

$$X_2^+ = 30 + 20 = 50 мг/л$$

$$X_2^- = 30 - 20 = 10 мг/л$$

$$X_3^+ = 1 + 1 = 2 мкг/л$$

$$X_3^- = 1 - 1 = 0 мкг/л$$

или в кодированных единицах:

$$x_1^+ = \frac{X_1^+ - X_{10}}{\lambda_1} = \frac{600 - 400}{200} = +1$$

$$x_1^- = \frac{200 - 400}{200} = -1$$

$$x_2^+ = \frac{50 - 30}{20} = +1$$

$$x_2^- = \frac{10 - 30}{20} = -1$$

$$x_3^+ = \frac{2-1}{1} = +1$$

$$x_3^- = \frac{0-1}{1} = -1$$

3. Определяем количество вариантов опытов, необходимое для выявления влияния на процесс не только каждого фактора в отдельности, но и факторов взаимодействия. Это число вариантов: $N = 2^n = 2^3 = 8$.

В таблице 8.5 приведен план опыта в натуральных значениях факторов и кодированных единицах.

Таблица 8.5

Матрица планирования эксперимента 2^3 в натуральных и кодированных значениях

u	Натуральные значения факторов			Факторы в кодированных единицах		
	X_1	X_2	X_3	x_1	x_2	x_3
1	200	10	0	-	-	-
2	600	10	0	+	-	-
3	200	50	0	-	+	-
4	600	50	0	+	+	-
5	200	10	2	-	-	+
6	600	10	2	+	-	-
7	200	50	2	-	+	+
8	600	50	2	+	+	+

В качестве примера приведем расчет состава трехкомпонентной смеси (ржано-пшеничная мука, солод ферментированный, сухой гидролизат картофеля) для приготовления ржано-пшеничного хлеба однофазным способом.

Так, как использование чистых компонентов (гидролизата картофеля и солода ферментированного) в данном случае не представляется разумным, то экспериментирование проводилось не на всей диаграмме, а на триангулированных ее частях (после соответствующей перенормировки компонентов, чтобы выполнялось условие равенства единице суммы концентраций). Область эксперимента представлена на рисунке 8.4.

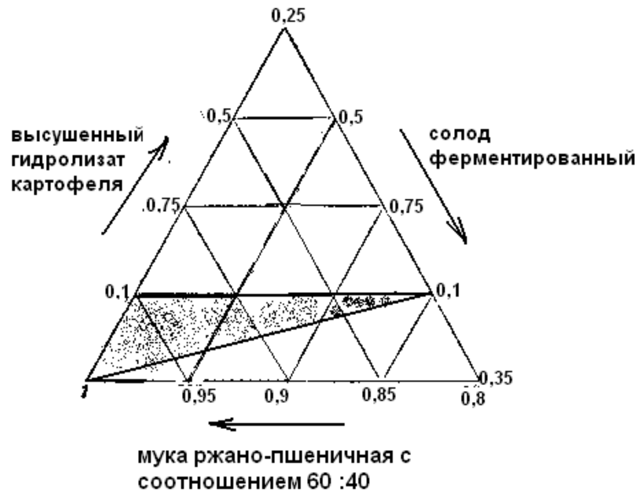


Рисунок 8.4 – Расположение триангулированных компонентов X_1, X_2 и X_3, заштрихованный участок – область экспериментирования

Для этого вводим кодовые переменные, которые сами по себе представляют собой уже не отдельные компоненты, а специально подобранные смеси (таблица 8.6)

Таблица 8.6 – Факторы и составы триангулированной смеси

Фактор (компонент)	Состав смеси (концентрация), %		
	Мука ржано-пшеничная с соотношение 60:40	Солод ферментированный	Гидролизат картофеля
X1	100	0	0
X2	90	10	0
X3	90	0	10

Дозировка компонентов взамен муки подбиралась исходя из предыдущих исследований и литературных данных. Условия эксперимента и результаты откликов для **кубической модели** приведены в таблице 8.7.

Таблица 8.7

Матрица планирования эксперимента в кодированных и натуральных значениях

№	Индекс отклика	Содержание компонентов			Состав смеси		
		X1	X2	X3	Мука ржано-пшеничная (60:40)	Солод ферментированный	Гидролизат картофеля
1	1	1	0	0	100	0	0
2	2	0	1	0	90	10	0
3	3	0	0	1	90	0	10
4	112	0,666	0,333	0	96,7	3,3	0
5	122	0,333	0,666	0	93,3	6,67	0
6	113	0,666	0	0,333	96,7	0	3,3
7	133	0,333	0	0,666	93,3	0	6,67
8	223	0	0,666	0,333	90	6,67	3,33
9	233	0	0,333	0,666	90	3,33	6,67
10	123	0,333	0,333	0,333	93,3	3,33	3,33

Задание 1. Составить план полного факторного эксперимента в кодированной и натуральной форме соответствии с вариантом

Вариант	Интервал варьирования и уровень факторов	Пределы изменения факторов			
		X_1	X_2	X_3	X_4
1	Нулевой уровень	0,04	15	4	55
	Интервал варьирования	0,01	5	1	15
2	Нулевой уровень	2	3		
	Интервал варьирования	1	1		

3	Нулевой уровень	0,7	0,001	0,9	0,6
	Интервал варьирования	0,1	0,0001	0,03	0,25
4	Нулевой уровень	6	0,02	1	3
	Интервал варьирования	0,3	0,01	0,2	0,2
5	Нулевой уровень	678	354	55	
	Интервал варьирования	75	120	12	
6	Нулевой уровень	75	120	12	
	Интервал варьирования	15	35	1	
7	Нулевой уровень	667	0,1	13	
	Интервал варьирования	150	0,01	3	
8	Нулевой уровень	0,1	0,0001	0,03	0,25
	Интервал варьирования	0,025	0,00001	0,01	0,01
9	Нулевой уро-	0,01	5	1	15

	вень				
	Интервал варьирования	0,0001	1	0,2	5
10	Нулевой уровень	150	0,01	3	
	Интервал варьирования	25	0,00025	0,4	
11	Нулевой уровень	75	120	12	
	Интервал варьирования	12	12	0,1	
12	Нулевой уровень	12	76	100	50
	Интервал варьирования	5	10	15	5
13	Нулевой уровень	45	1	67	45
	Интервал варьирования	16	0,05	20	12
14	Нулевой уровень	12	17		
	Интервал варьирования	3	5		
15	Нулевой уровень	35	35	35	

	Интервал варьирования	5	15	20	
16	Нулевой уровень	467	376	12	78
	Интервал варьирования	200	45	1	6
17	Нулевой уровень	65	89	90	
	Интервал варьирования	15	16	45	
18	Нулевой уровень	14	0,5	15	100
	Интервал варьирования	4	0,03	3	12

Задание 2 Осуществить симплексно-решетчатое планирование эксперимента в кодированной и натуральной форме соответствии с вариантом

Вариант и тип модели	Фактор	Состав смеси (концентрация), %		
		X_1	X_2	X_3
1 квадратичная	X_1	100	0	0
	X_2	80	20	0
	X_3	50	0	50
2 квадратичная	X_1	100	0	0
	X_2	90	10	0
	X_3	90	0	10
3 квадратичная	X_1	100	0	0
	X_2	50	50	0
	X_3	50	0	50
4	X_1	100	0	0

линейная	X_2	50	50	0
	X_3	80	0	20
5 кубическая	X_1	100	0	0
	X_2	50	50	0
	X_3	80	0	20
6 кубическая	X_1	80	10	0
	X_2	100	0	0
	X_3	50	25	0
7 квадратичная	X_1	80	10	0
	X_2	100	0	0
	X_3	50	25	0
8 кубическая	X_1	10	90	0
	X_2	35	60	5
	X_3	12	18	70
9 квадратичная	X_1	10	90	0
	X_2	35	60	5
	X_3	12	18	70
10 кубическая	X_1	60	20	20
	X_2	40	10	25
	X_3	50	15	15
11 квадратичная	X_1	60	20	20
	X_2	40	10	25
	X_3	50	15	15
12 кубическая	X_1	100	0	0
	X_2	20	0	0
	X_3	60	0	0
14 квадратичная	X_1	100	0	0
	X_2	20	0	0
	X_3	60	0	0
15 линейная	X_1	100	0	0
	X_2	20	0	0
	X_3	60	0	0
16 кубическая	X_1	1	0	0
	X_2	0,5	0,25	0
	X_3	0,4	0,5	0,1
17 квадратичная	X_1	1	0	0
	X_2	0,75	0	0,25
	X_3	0,5	0,3	0,2

18 кубическая	X_1	25	70	5
	X_2	30	40	30
	X_3	10	90	0

Контрольные вопросы:

1. Что представляет собой интервал варьирования факторов?
2. Каким образом определяется нулевой уровень факторов?
3. Как определить верхний и нижний уровни факторов?
4. Что представляет симплекс-решетчатое планировании эксперимента? В каких случаях оно применяется?

Лабораторная работа №9 Обработка результатов полнофакторных экспериментов

Цель работы: представление результатов опытов в виде математической модели.

Прежде чем строить математическую модель по результатам опытов, необходимо проверить, не было ли влияния на некоторые из них неучтенных факторов. Проверка воспроизводимости представляет собой проверку выполнения предпосылки регрессионного анализа об однородности выборочных дисперсий . Задача состоит в проверке гипотезы о равенстве дисперсий $S^2y_1 = S^2y_2 = S^2y_3 = \ldots = S^2y_8$ при экспериментах во всех многомерных точках (8 точек для трех факторов). Оценки дисперсий находят по формуле:

$$S_{yj}^2 = (m-n)^{-1}\sum_{l=1}^{m}(y_{jl}-\overline{y_j})^2 \qquad 9.1$$

Если каждый опыт проводился трижды, то дисперсии получают по выборкам одинакового объема *m=3*, а число степеней свободы для всех дисперсий одинаково и равно *m— 1=2*.

Для проверки гипотезы об однородности оценок дисперсий следует пользоваться критерием Кохрена *G*, который основан на законе распределения отношения максимальной полученной дисперсии к сумме всех дисперсий. Расчетное значение критерия получают по формуле:

$$G_p = \frac{S^2_{y\max}}{\sum_{i=1}^{N} S^2_{yi}} (f_1 = m-1, f_2 = N) \qquad 9.2$$

Если вычисленное значение критерия *Gp* окажется меньше его табличного значения G_T для выбранного уровня значимости (обычно а=0,05), то гипотеза об однородности дисперсий принимается. Это значит, что различий между дисперсиями не должно быть, а полученное различие случайно.

В этом случае можно рассчитать генеральную дисперсию воспроизводимости:

$$S^2_y = \frac{\sum_{i=1}^{N} S^2_{yi}}{N} = S^2_{воспр} \qquad 9.3$$

имеющую $f_3=N(m — 1)$ степеней свободы.

Если проверка на воспроизводимость дала отрицательный результат, то остается признать невоспроизводимость эксперимента относительно управляемых переменных вследствие флуктуации неуправляемых и неконтролируемых переменных, создающих на выходе большой уровень «шума». При этом следует увеличить число параллельных опытов.

Построение математической модели объекта начинается с определения независимых оценок коэффициентов b_i по формуле:

$$b_i = \sum_{i=1}^{N} \frac{z_i \left| \overline{y_i} \right|}{N} \qquad 9.4$$

После этого необходимо проверить гипотезу о значимости коэффициентов b_i (проверка нуль-гипотезы $b_i=0$). Проверка гипотезы проводится с помощью t-критерия, который в данном случае формируется как:

$$t_{pi} = \frac{|b_i|}{S_{bi}} \qquad 9.5$$

где S_{bi} — среднеквадратическая ошибка определения коэффициента b_i при полном факторном эксперименте, равная:

$$\sqrt{\frac{S_y^2}{N_m}} \qquad 9.6$$

имеющую $f_3=N(m — 1)$ степеней свободы.

Если найденная величина параметра t превышает значение $t_{кр}$, определенное по приложению 4 для числа степеней свободы $f=N(m — 1)$ при заданном уровне значимости, т. е. sgn (t_{Pi} — tкр) = «+», то гипотеза отвергается и коэффициент b_i признается значимым. В противном случае, если sgn (t_{pi} — t_{kp}) = «-», нуль-гипотеза, принимается и коэффициент b_i - считают статистически незначимым (т. е. равным нулю).

Статистическая незначимость коэффициента bi может быть обусловлена следующими причинами:

1. уровень базового режима близок к точке частного экстремума по переменной (или по произведению переменных);

2. шаг варьирования Δx_i выбран малым; .

3. данная переменная (произведение переменных) не имеет функциональной связи с выходным параметром у, т. е. $b_i = 0$ по существу;

4. велика ошибка эксперимента вследствие флуктуации неуправляемых и неконтролируемых переменных.

Поскольку - ортогональное планирование позволяет определять независимые оценки коэффициентов, то если какой-либо из коэффициентов окажется незначимым, он может быть отброшен без пересчета всех остальных. После этого математическая модель объекта составляется в виде уравнения связи выходного параметра у и переменных x_i, включающего только значимые коэффициенты.

Чтобы проверить гипотезу об адекватности результатов эксперимента найденным по уравнению связи, Достаточно оценить отклонение предсказанной уравнением регрессии выходной величины y_{jp} от результатов эксперимента $y_{jэ}$ в точках факторного пространства. Разброс результатов эксперимента относительно линии регрессии в этом случае сравнивается с разбросом точек между собой. Первый характеризуется дисперсией адекватности $S^2{}_{АД}$, оценка которой находится по формуле:

$$S_{ад}^2 = (N-d)^{-1}\sum_{i=1}(\overline{y}_{iэ} - y_{jp})^2 \qquad 9.7$$

где d — число коэффициентов линейного уравнения при числе степеней свободы $f_{АД}=N—n— 1$.

Второй (разброс точек) — дисперсией воспроизводимости $S^2{}_{воспр}$ с числом степеней свободы $t_{воспр} = N(m—1)$. Если $S^2{}_{АД}< S^2{}_{воспр}$, то полученная математическая модель адекватно представляет результаты эксперимента. Если $S^2{}_{АД}> S^2{}_{воспр}$ то нужно выяснить, случайно или не случайно различаются эти дисперсии, и если не случайно, то делается вывод о неадекватности уравнения.

Проверка гипотезы об адекватности проводится с использованием критерия Фишера. Если $F_p<F_{КР}$ для соответствующих степеней свободы $f_{АД} = N - d =N—n—1$ и $f_{воспр}=N(m—1)$ при заданном уровне значимости α, то нуль-гипотеза принимается, т. е. уравнение адекватно при $sgn(F_p—F_{КР})$ =«—». В противном случае гипотеза отвергается и описание признается неадекватным объекту.

Проверка адекватности возможна при $f_{АД}>0$. Бывает, что число вариантов (строчек) варьирования ПФЭ равно числу оцениваемых коэффициентов регрессии ($N=n+1$), следовательно, не остается степеней свободы ($f_{АД} = 0$) для проверки нуль-гипотезы об адекватности экспериментальных данных выбранной форме аппроксимирующего полинома: Если же некоторые коэффициенты регрессии оказались незначимыми, или ими можно пренебречь, или требуется доказать равномерность линейной аппроксимации в заданном интервале варьирования, то число членов проверяемого уравнения в этом случае будет меньше числа вариантов варьирования ($N>1$) и одна или несколько степеней свободы ($f_{АД}>0$) останется для проверки гипотезы об адекватности.

Если гипотеза об адекватности отвергается, необходимо переходить к более сложной форме уравнения либо (если это возможно) проводить эксперимент с меньшим шагом варьирования Δx_i.

Следует отметить, что шаг варьирования выбирается из условия адекватного описания в области планирования. Если при больших шагах варьирования математическая модель неадекватна объекту, то возникают систематические ошибки в определении коэффициентов, для уменьшения которых следует сузить область варьирования. Однако с уменьшением шага варьирования появляется целый ряд новых трудностей: растет отношение помехи к полезному сигналу, что приводит к необходимости увеличения числа дублирующих экспериментов для выделения сигнала на фоне шума; уменьшаются значения коэффициентов b (величины которых непосредственно зависят от шага варьирования (для уравнения с нормированными переменными). Коэффициенты могут стать статистически незначимыми.

Для выбора шага варьирования проводятся предварительные эксперименты. Шаг варьирования можно выбирать равным 0,03:— 0,30 от диапазона варьирования факторов, т.е. область варьирования составляет 6—60% от всей области значений фактора.

Если щаг (интервал) варьирования не превышает 10% от области значений фактора, то говорят обычно об узком интервале варьирования, значения 10-30% соответствуют среднему интервалу, а больше 30% — широкому. Начальная точка варьирования (базовый уровень) выбирается возможно ближе к центру области факторного пространства, в которой ищется математическое описание.

Пример выполнения задания

Матрица планирование эксперимента

№ опыта	X_1	X_2	X_3	Y	
1	14,8	2	0,2	187	198
2	59,8	2	0,2	19	23
3	14,8	6	0,2	57	69
4	14,8	2	0,5	95	116
5	59,8	6	0,2	5	7
6	59,8	2	0,5	9	13
7	14,8	6	0,5	40	52
8	59,8	6	0,5	2	4

1. Проверить однородность дисперсий параллельных опытов. Определить коэффициенты уравнения регрессии (степенная модель) по заданной матрице планирования эксперимента.
2. Проверить значимость коэффициентов уравнения регрессии
3. Проверить адекватность полученной модели
4. Перейти от полученной линейной модели к степенной

1 Проверить однородность дисперсий параллельных опытов. Определить коэффициенты уравнения регрессии (степенная модель) по заданной матрице планирования эксперимента.

Линеаризация

$\ln y = \ln c + \alpha \ln x_1 + \beta \ln x_2 + \gamma \ln x_3$

$b = b_0 + b_1 z_1 + b_2 z_2 + b_3 z_3$

№ опыта	ln X_1	ln X_2	ln X_3	ln Y		$Y_{ср}$	S_i^2
1	2,6946272	0,693147	-1,60944	5,231109	5,288267031	5,259688	0,001634
2	4,0910057	0,693147	-1,60944	2,944439	3,135494216	3,039967	0,018251
3	2,6946272	1,791759	-1,60944	4,043051	4,234106505	4,138579	0,018251
4	2,6946272	0,693147	-0,69315	4,553877	4,753590191	4,653734	0,019943
5	4,0910057	1,791759	-1,60944	1,609438	1,945910149	1,777674	0,056607
6	4,0910057	0,693147	-0,69315	2,197225	2,564949357	2,381087	0,067611
7	2,6946272	1,791759	-0,69315	3,688879	3,951243719	3,820062	0,034418
8	4,0910057	1,791759	-0,69315	0,693147	1,386294361	1,039721	0,240227

Дисперсия воспроизводимости (характеризует среднюю ошибку опытов):

$$S^2\{y\} = \frac{\sum_{i=1}^{n} S_i^2}{n} = \frac{0,45694}{8} = 0,05712.$$

Перейдем к модели в нормированных координатах

№ опыта	Z_0	Z_1	Z_2	Z_3	Z_1*Z_2	Z_1*Z_3	Z_2*Z_3	Z_1*Z_2*Z_3
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
b_i	3,2638	-1,2042	-0,5698	-0,29	-0,081	-0,059	0,026	-0,04581

Уравнение в нормированных координатах:

$$y = 3{,}2638 - 1{,}2042 \cdot Z_1 - 0{,}5698 \cdot Z_2 - 0{,}29 \cdot Z_3 - 0{,}081 \cdot Z_{12} - 0{,}059 \cdot Z_{13} + 0{,}026 \cdot Z_{23} - 0{,}04581 \cdot Z_{123}$$

. Проверка однородности дисперсий (критерий Кохрена):

$$G = \frac{S^2_{\max i}}{\sum_{i=1}^{n} S_i^2} = \frac{0{,}067611}{0{,}45694} = 0{,}523.$$

Для 8 опытов и двух дублей $G_{табл}$=0,6798. Так как $G < G_{табл}$ дисперсии однородны.

№2 Проверить значимость коэффициентов уравнения регрессии

$$S^2\{b\} = \frac{S^2\{y\}}{n \cdot k} = \frac{0{,}05712}{8 \cdot 2} = 0{,}00357.$$

Критерий Стьюдента: t=2,366.

Доверительный интервал:

$$\Delta b = \pm \left| t \cdot \sqrt{S^2\{b\}} \right| = \pm \left| 2{,}366 \cdot \sqrt{0{,}00357} \right| = \pm 0.14\,.$$

Сравнивая с доверительным интервалом, определяем, что значимыми являются коэффициенты b_0, b_1, b_2, b_3.

Окончательный вид модели

$$y = 3{,}2638 - 1{,}2042 \cdot Z_1 - 0{,}5698 \cdot Z_2 - 0{,}29 \cdot Z_3.$$

№3 Проверить адекватность полученной модели

№ опыта	Y_{cp}	Y_i	$(Y_{cp}-Y_i)^2$
1	5,259688	5,327983	0,0046642960
2	3,039967	2,919580	0,0144929069
3	4,138579	4,188374	0,0024795097
4	4,653734	4,747657	0,0088216832
5	1,777674	1,779970	0,0000052724
6	2,381087	2,339254	0,0017499980
7	3,820062	3,608047	0,0449499968
8	1,039721	1,199644	0,0255754634

$$S_{a\partial}^2 = \frac{k\sum_{i=1}^{n}(y_{cp}-y_i)^2}{n-p-1} = \frac{2\cdot 0{,}1027}{8-3-1} = 0{,}05135.$$

F-критерий Фишера

$$F = \frac{S_{a\partial}^2}{S^2\{y\}} = \frac{0{,}05135}{0{,}05712} = 0{,}899.$$

Для 8 опытов и двух дублей $F_{табл}=6{,}04$. Так как $F<F_{табл}$, модель адекватна.

4 Перейти от полученной линейной модели к степенной

$$c = \sum_{i=0}^{p} b_i - 2\sum_{i=1}^{p}\frac{b_i\cdot\ln x_{i\max}}{\ln x_{i\max}-\ln x_{i\min}}.$$

$$c = 2{,}71^{3{,}2638-1{,}2042-0{,}5698-0{,}29-2\frac{-1{,}2042\cdot\ln 59{,}8}{\ln 59{,}8-\ln 14{,}8}-2\frac{-0{,}5698\cdot\ln 6}{\ln 6-\ln 2}-2\frac{-0{,}29\cdot\ln 0{,}5}{\ln 0{,}5-\ln 0{,}2}} = 15455{,}28.$$

$$\alpha_i = \frac{2\cdot b_i}{\ln x_{i\max}-\ln x_{i\min}}.$$

$$\alpha_1 = \frac{2\cdot(-1{,}2042)}{\ln 59{,}8-\ln 14{,}8} = -1{,}725.$$

$$\alpha_2 = \frac{2\cdot(-0{,}5698)}{\ln 6-\ln 2} = -1{,}037.$$

$$\alpha_3 = \frac{2\cdot(-0{,}29)}{\ln 0{,}5-\ln 0{,}2} = -0{,}633.$$

Окончательный вид модели:

$$Y = 15455{,}28X_1^{-1{,}725}X_2^{-1{,}037}X_3^{-0{,}633}.$$

Вариант № 1

1) Определить коэффициенты уравнения регрессии (степенная модель) по заданной матрице планирования эксперимента.

№ опыта	X_1	X_2	X_3	Y
1	14,1	2	0,2	2,2
2	14,1	2	0,5	5,6
3	56,5	6	0,2	1,4
4	56,5	2	0,5	2,3
5	14,1	6	0,2	2,8
6	14,1	6	0,5	5,7
7	56,5	6	0,5	2,7
8	56,5	2	0,2	1,1

2) Проверить значимость коэффициентов уравнения регрессии, если параллельные опыты в нулевой точке плана дали следующие результаты: 3,5; 3,2; 2,9.

3) Проверить адекватность полученной модели.

4) Перейти от полученной линейной модели к степенной.

5) По результатам моделирования (пп. 1 - 4) сделать выводы о характере влияния факторов на исследуемый параметр и о степени пригодности и возможной области применения полученной модели.

Вариант № 2

1) Проверить однородность дисперсий параллельных опытов. Определить коэффициенты уравнения регрессии (степенная модель) по заданной матрице планирования эксперимента.

№ опыта	X_1	X_2	X_3	Y	
1	15,7	2	0,2	74	95
2	61,5	2	0,2	10	17
3	15,7	6	0,2	68	89
4	15,7	2	0,5	38	60
5	61,5	6	0,2	4	8
6	61,5	2	0,5	0,8	2

7	15,7	6	0,5	12	28
8	61,5	6	0,5	1,1	2,3

2) Проверить значимость коэффициентов уравнения регрессии.

3) Проверить адекватность полученной модели.

4) Перейти от полученной линейной модели к степенной.

5) По результатам моделирования (п.п. 1 - 4) сделать выводы о характере влияния факторов на исследуемый параметр и о степени пригодности полученной модели.

Вариант № 3

1) Проверить однородность дисперсий параллельных опытов. Определить коэффициенты уравнения регрессии (степенная модель) по заданной матрице планирования эксперимента.

№ опыта	X_1	X_2	X_3	Y		
1	31,5	2	0,2	12	17	22
2	125	2	0,2	11	15	21
3	31,5	6	0,2	21	28	34
4	31,5	2	0,5	35	40	44
5	125	6	0,2	27	23	18
6	125	2	0,5	33	38	42
7	31,5	6	0,5	42	49	56
8	125	6	0,5	36	44	52

2) Проверить значимость коэффициентов уравнения регрессии.

3) Проверить адекватность полученной модели.

4) Перейти от полученной линейной модели к степенной.

5) По результатам моделирования (пп. 1 - 4) сделать выводы о характере влияния факторов на исследуемый параметр и о степени пригодности применения полученной модели.

Вариант № 4

1) Определить коэффициенты уравнения регрессии (степенная модель) по заданной матрице планирования эксперимента.

№ опыта	X_1	X_2	X_3	Y
1	50	2	0,2	1,440
2	200	2	0,2	1,425
3	50	6	0,5	1,675
4	50	6	0,2	1,505
5	50	2	0,5	1,566
6	200	6	0,2	1,485
7	200	2	0,5	1,545
8	200	6	0,5	1,705

2) Проверить значимость коэффициентов уравнения регрессии, если в нулевой точке плана параллельные опыты дали следующие результаты: 1,505; 1,475; 1,520.

3) Проверить адекватность полученной модели.

4) Перейти от полученной линейной модели к степенной.

5) По результатам моделирования (пп. 1 - 4) сделать выводы о характере влияния факторов на исследуемый параметр и о степени пригодности полученной модели.

Вариант № 5

1) Проверить однородность дисперсий параллельных опытов. Определить коэффициенты уравнения регрессии (степенная модель) по заданной матрице планирования эксперимента.

№ опыта	X_1	X_2	X_3	Y	
1	14,8	2	0,2	187	198
2	59,8	2	0,2	19	23
3	14,8	6	0,2	57	69
4	14,8	2	0,5	95	116
5	59,8	6	0,2	5	7
6	59,8	2	0,5	9	13
7	14,8	6	0,5	40	52
8	59,8	6	0,5	2	4

2) Проверить значимость коэффициентов уравнения регрессии.

3) Проверить адекватность полученной модели.

4) Перейти от полученной линейной модели к степенной.

5) По результатам моделирования (пп. 1 - 4) сделать выводы о характере влияния факторов на исследуемый параметр и о степени пригодности полученной модели.

Вариант № 6

1) Задана полуреплика 2^{4-1} . Определить коэффициенты степенного уравнения регрессии.

№ опыта	X_1	X_2	X_3	X_4	Y
1	0,025	30	30	400	3,4
2	0,010	5	15	400	1,7
3	0,025	5	15	25	2,5
4	0,010	30	15	25	2,7
5	0,010	30	30	25	5,1
6	0,025	5	30	25	4,9
7	0,025	30	15	400	2,2
8	0,010	5	30	400	3,8

2) Проверить значимость коэффициентов уравнения регрессии, если в нулевой точке плана параллельные опыты дали следующие результаты: 3,7; 3,0; 2,5.

3) Проверить адекватность полученной модели.

4) Перейти от полученной линейной модели к степенной.

5) По результатам моделирования (пп. 1 - 4) сделать выводы о характере влияния факторов на исследуемый параметр и о степени пригодности полученной модели.

Вариант № 7

1) Задана полуреплика 2^{4-1} . Определить коэффициенты степенного уравнения регрессии.

№ опыта	X_1	X_2	X_3	X_4	Y
1	0,025	30	30	400	6,3

2	0,010	5	15	400	2,6
3	0,025	5	15	25	3,8
4	0,010	30	15	25	4,0
5	0,010	30	30	25	7,7
6	0,025	5	30	25	7,4
7	0,025	30	15	400	3,3
8	0,010	5	30	400	5,7

2) Проверить значимость коэффициентов уравнения регрессии, если параллельные опыты в центре плана дали следующие результаты: 5,6; 4,5; 3,8.

3) Проверить адекватность полученной модели.

4) Перейти от полученной линейной модели к степенной.

5) По результатам моделирования (пп. 1 - 4) сделать выводы о характере влияния факторов на исследуемый параметр и о степени пригодности и возможной области применения полученной модели.

Вариант № 8

1) Проверить однородность дисперсий параллельных опытов. Определить коэффициенты уравнения регрессии (степенная модель) по заданной матрице планирования эксперимента.

№ опыта	X_1	X_2	X_3	Y	
1	14,1	2	0,2	55	84
2	56,5	2	0,2	65	91
3	14,1	6	0,2	205	280
4	14,1	2	0,5	14	27
5	56,5	6	0,2	218	292
6	56,5	2	0,5	19	32
7	14,1	6	0,5	87	143
8	56,5	6	0,5	95	155

2) Проверить значимость коэффициентов уравнения регрессии.

3) Проверить адекватность полученной модели.

4) Привести полученную модель к степенному виду.

5) По результатам моделирования (пп. 1 - 4) сделать выводы о характере влияния факторов на исследуемый параметр и о степени пригодности полученной модели.

Вариант № 9

1) Определить коэффициенты уравнения регрессии (степенная модель) по заданной матрице планирования эксперимента.

№ опыта	X_1	X_2	X_3	Y
1	2,2	0,033	13	0,037
2	22,0	0,13	115	0,250
3	22,0	0,133	13	0,090
4	2,2	0,13	13	0,062
5	22,0	0,033	13	0,055
6	2,2	0,033	115	0,110
7	22,0	0,033	115	0,155
8	2,2	0,13	115	0,160

2) Проверить значимость коэффициентов уравнения регрессии, если параллельные опыты в центре плана дали следующие результаты: 0,082; 0,116; 0,104.

3) Проверить адекватность полученной модели.

4) Привести полученную модель к степенному виду.

5) По результатам моделирования (пп. 1 - 4) сделать выводы о характере влияния факторов на исследуемый параметр и о степени пригодности полученной модели.

Вариант № 10

1) Определить коэффициенты уравнения регрессии (степенная модель) по заданной матрице планирования эксперимента.

№ опыта	X_1	X_2	X_3	Y
1	14,1	2	0,2	3,5
2	14,1	2	0,5	6,3
3	56,5	6	0,2	1,6

4	56,5	2	0,5	2,3
5	14,1	6	0,2	4,1
6	14,1	6	0,5	6,0
7	56,5	6	0,5	2,3
8	56,5	2	0,2	1,7

2) Проверить значимость коэффициентов полученной модели, если параллельные опыты в центре плана дали следующие результаты: 4,0; 4,3; 4,5.

3) Проверить адекватность полученной модели.

4) Привести полученную модель к степенному виду.

5) По результатам моделирования (пп. 1 - 4) сделать выводы о характере влияния факторов на исследуемый параметр и о степени пригодности полученной модели.

Лабораторная работа №10 Статистическое моделирование полного факторного эксперимента

Цель работы: основной целью является проведение современного эксперимента и разработка математической модели, адекватно описывающей процесс и позволяющей осуществлять управление производством.

1 Подготовка к выполнению работы

При планировании эксперимента исследователь должен:

- обеспечить высокую надежность и четкость интерпретации **ре**зультатов экспериментальных исследований;
- составить последовательную логическую схему построения всего процесса исследования;
- максимально формализовать процесс разработки модели и **со**поставления экспериментальных данных различных опытов одного и того же объекта исследований с целью широкого применения **элек**тронно-вычислительных средств.

Всем требованиям отвечают **статистические методы планирования эксперимента.** Статистические методы планирования активного эксперимента являются одним из эмпирических способов

получения математического описания статики сложных объектов исследования, то есть уравнения связи отклика объекта и независимых управляемых входных переменных (факторов).

При исследовании процессов резания многие зависимости традиционно представляют уравнениями степенного вида. Например, зависимость стойкости режущего инструмента от элементов режима резания часто выражают уравнением:

$$T = cV^{\alpha} S^{\beta} t^{\gamma} \qquad 10.1$$

где v — скорость резания; s — подача; t — глубина резания; c, α, β, γ — постоянные величины.

Уравнение (10.1) в результате логарифмирования линеаризуется:

$$\ln T = \ln c + \alpha \ln V + \beta \ln S + \gamma \ln t \qquad 10.2$$

Уравнение (10.2) можно выразить следующим образом:

$$y = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 \qquad 10.3$$

где, $y = \ln T; x_1, x_2, x_3$ - кодированные значения *v, s, t.*

Кодированное значение x_i фактора:

$$x_i = \frac{2(\ln \bar{x}_i - \ln \bar{x}_{iв})}{\ln \bar{x}_{iв} - \ln \bar{x}_{iн}} + 1 \qquad 10.4$$

где $\bar{x}_i$ - натуральное значение; $\bar{x}_{iв}, \bar{x}_{iн}$ натуральные значения верхнего и нижнего уровней соответственно.

Для оценки коэффициентов уравнения (3) удобно использовать результаты многофакторного эксперимента. При этом результаты опытов обычно представляют полиномом вида:

$$y = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + b_{12} x_1 x_2 + b_{13} x_1 x_3 + b_{23} x_2 x_3 \qquad 10.5$$

или полиномом вида:

$$y = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + b_{12} x_1 x_2 + b_{13} x_1 x_3 + b_{23} x_2 x_3 + b_{11} x_1^2 + b_{22} x_2^2 + b_{33} x_3^2 \qquad 10.6$$

После определения коэффициентов уравнения (10.5) или уравнения (10.6) необходимо проверить гипотезу адекватности линейной части этих полиномов. Если гипотеза подтверждается, т. е. уравнение $y = b_0 + b_1x_1 + b_2x_2 + b_3x_3$ адекватно, то, подставляя в это уравнение значения x_1, x_2, x_3 определяемые соотношением (10.4), получим выражение (10.2). Если гипотеза адекватности уравнения $y = b_0 + b_1x_1 + b_2x_2 + b_3x_3$ не подтверждается, то необходимо проверить гипотезу адекватности уравнения (10.5) или (10.6). В случае, если уравнения (10.5) или (10.6) окажутся адекватными, можно будет предложить другую модель, характеризующую зависимость стойкости инструмента от исследуемых факторов.

2. Порядок выполнения работы

Исходные данные

№ оп	n, об/мин	Шаг S мм	t, мм	P_y, Н		
1	9	2	0.2	260	280	330
2	37	2	0.3	160	160	155
3	9	6	0.4	320	350	300
4	9	2	0.5	810	860	970
5	37	6	0.6	220	200	190
6	37	2	0.7	540	500	560
7	9	6	0.8	870	800	920
8	37	6	0.9	540	580	510

Так как опыты дублировались вычислим среднее значение каждого опыта, а данную модель линеаризуем и заменим X на Z.

Степенная модель: $Y = C * X_1^{\alpha_1} * X_2^{\alpha_2} * X_3^{\alpha_3}$

После линеаризации получили модель вида:

$$\text{lgY=lgC+}\alpha_1\ \text{lgX}_2\text{+}\alpha_3\text{lgX}_3$$

$$\text{lgX}_1\text{+}\alpha_2\text{Y=b}_0\text{+b}_1\text{Z}_1\text{+b}_2\text{Z}_2\text{+b}_3\text{Z}_3\text{+b}_{12}\text{Z}_1\text{Z}_2\text{+b}_{13}\text{Z}_1\text{Z}_3\text{+b}_{23}\text{Z}_2\text{Z}_3\text{+b}_{123}\text{Z}_1\text{Z}_2\text{Z}_3$$

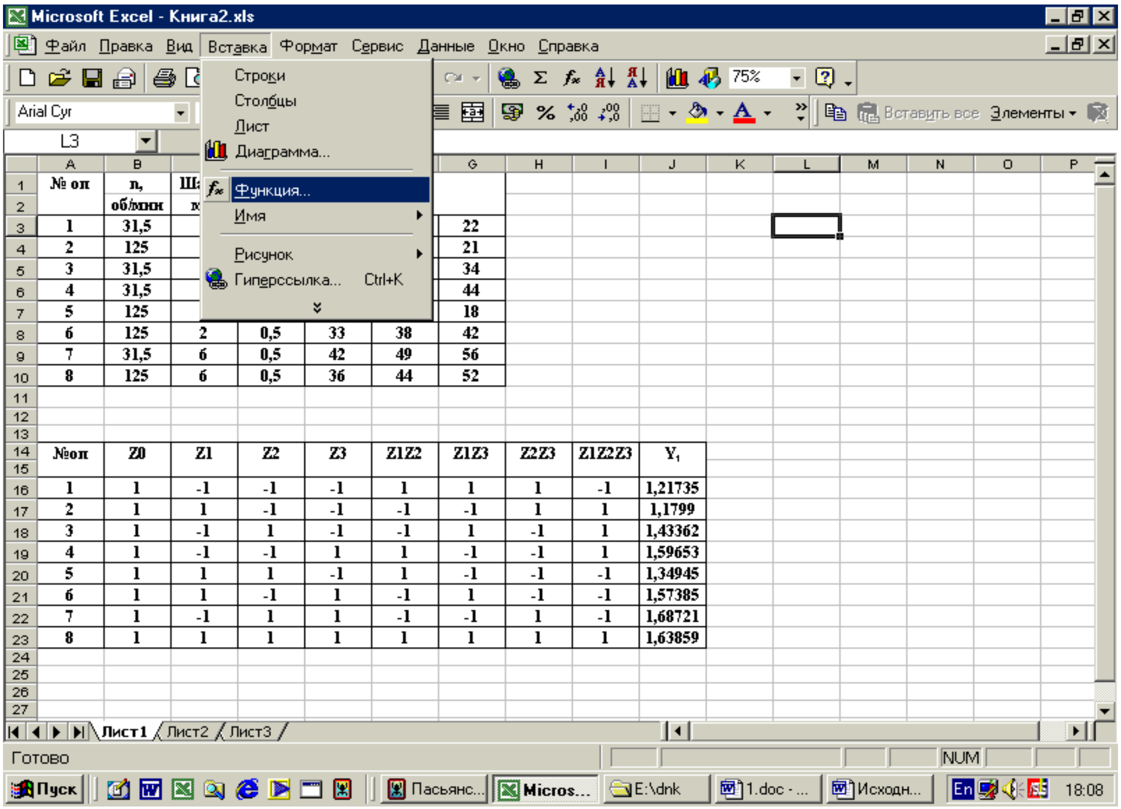
Составим таблицу после линеаризации модели

№оп	Z0	Z1	Z2	Z3	Z1Z2	Z1Z3	Z2Z3	Z1Z2Z3	Y_i
	0	1	2	3	12	13	23	123	2,414973
1	1	-1	-1	-1	1	1	1	-1	2,20412

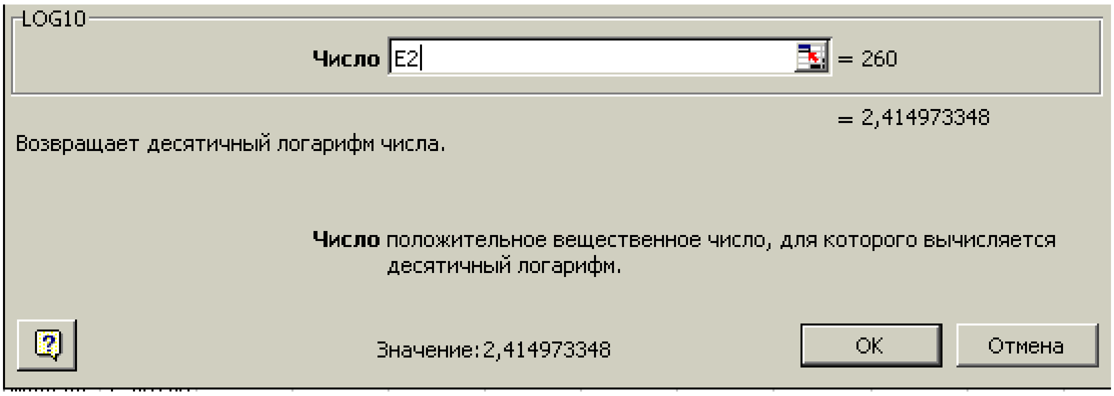
2	1	1	-1	-1	-1	-1	1	1	2,50515
3	1	-1	1	-1	-1	1	-1	1	2,908485
4	1	-1	-1	1	1	-1	-1	1	2,342423
5	1	1	1	-1	1	-1	-1	-1	2,732394
6	1	1	-1	1	-1	1	-1	-1	2,939519
7	1	-1	1	1	-1	-1	1	-1	2,732394
7	1	-1	1	1	-1	-1	1	-1	2,732394

Для того чтобы получить значение Y_i необходимо с помощью десятичного логарифма логарифмировать выходные значения, а затем найти среднее значение в каждом опыте т.е. в электронной таблице Excel выбираем в строке меню «Вставка» в ней выбираем пункт «функция» т.е.

После появления таблицы под названием «Мастер функций». В



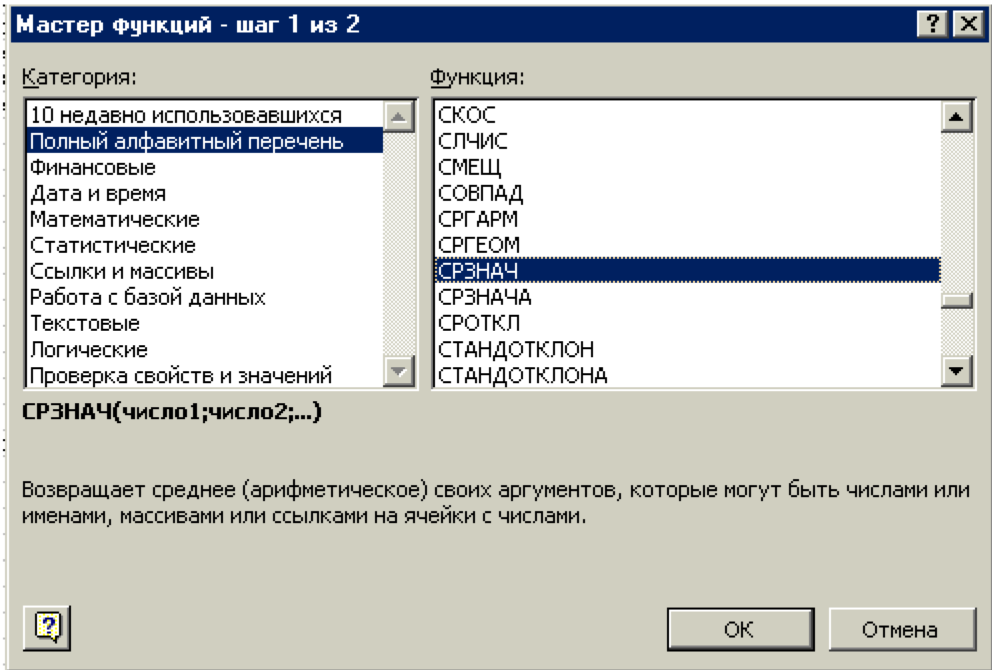
«Мастер функций» выбираем функцию «ln» и нажимаем ОК. Появляется контекстное меню:

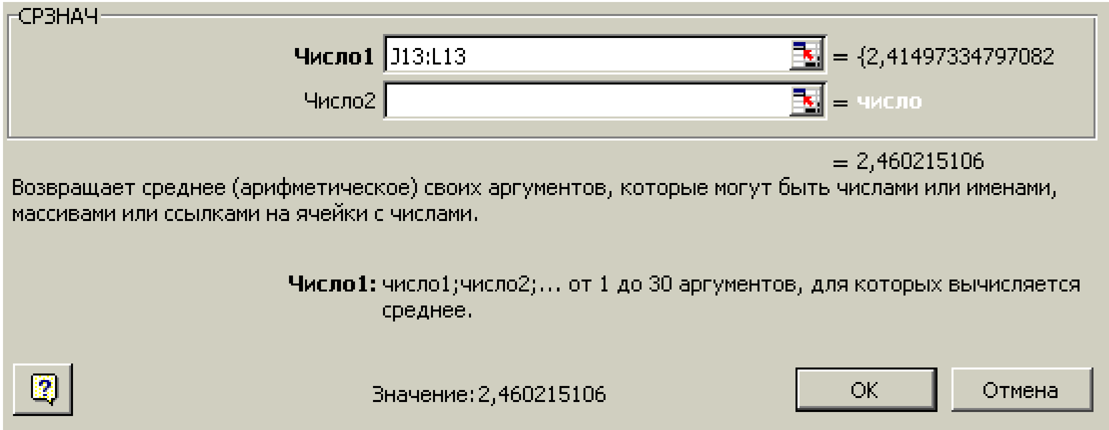


В строку Число заносим данные: E2 и нажимаем кнопку ОК. Получаем результат. Затем копируем его для остальных по аналогии.

Затем находим Y_i как среднее значение т.е.

В электронной таблице Excel выбираем в строке меню «Вставка» в ней выбираем пункт «функция» в которой выбираем СРЗНАЧ и нажимаем ОК т.е.



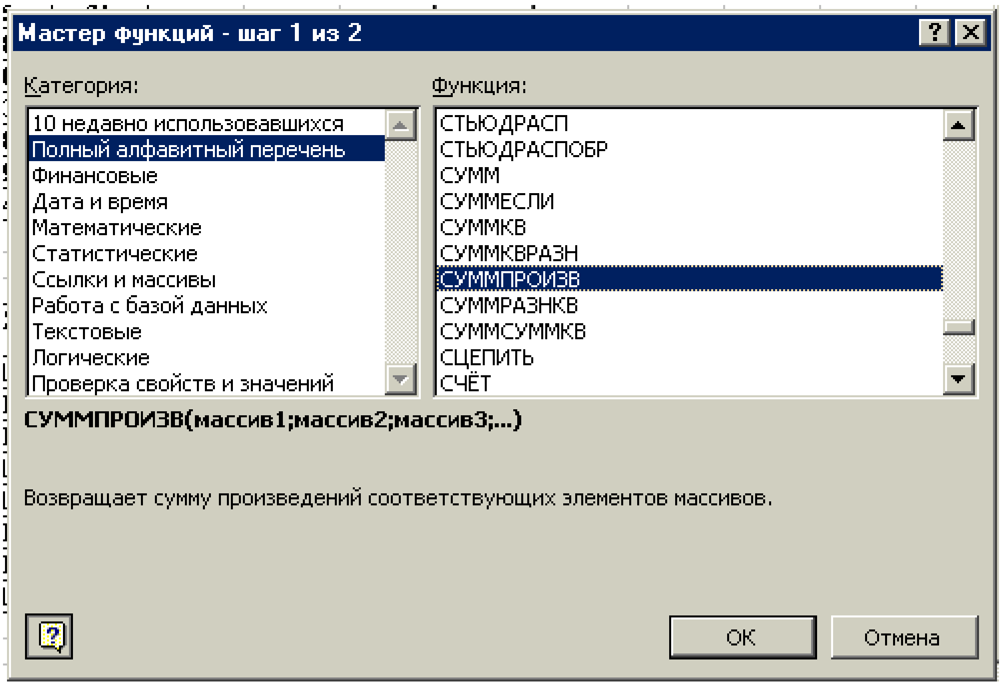


Рассчитываем значения коэффициентов b_0,b_1,b_2, и тд. предварительной модели по формуле:

$$b_i = \frac{1}{n}\sum_{i=1}^{n} Z_{ij} * \overline{Y}_i \qquad 10.7$$

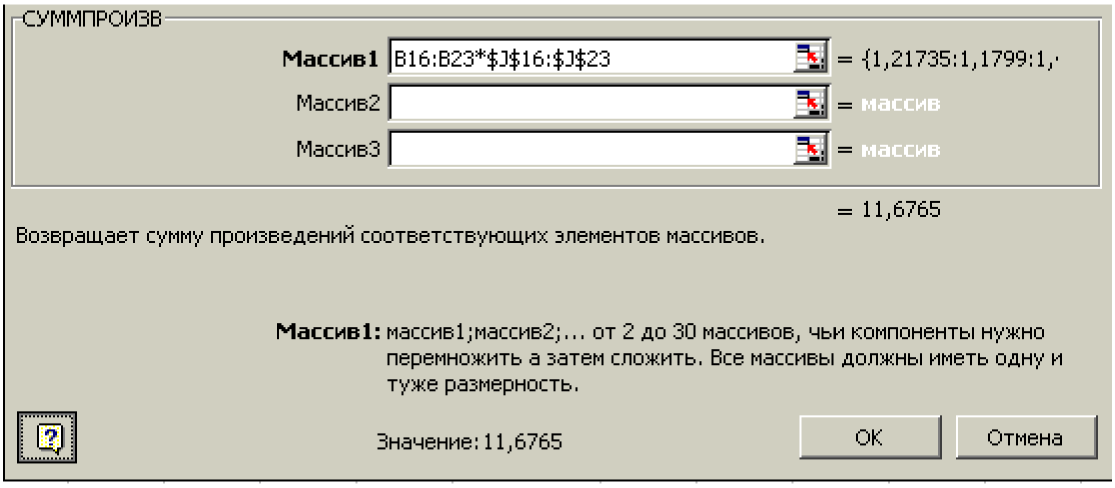
где n – число опытов; n=8

или с помощью «Мастера функций» и функции «Сумма произведений» т.е.



После появления «Мастера функций» и выбрав в нем команду «суммпроиз» нажимаем ОК.

Появляется окно «СУММПРОИЗ» в массив 1 вносим данные и делим их на число опытов т.е. на 8.



Получаем результат: b_0=2,60195. Затем копируем его для $b_1,b_2,b_3,b_1b_2,b_1b_3,b_2b_3,b_1b_2b_3$. получаем.

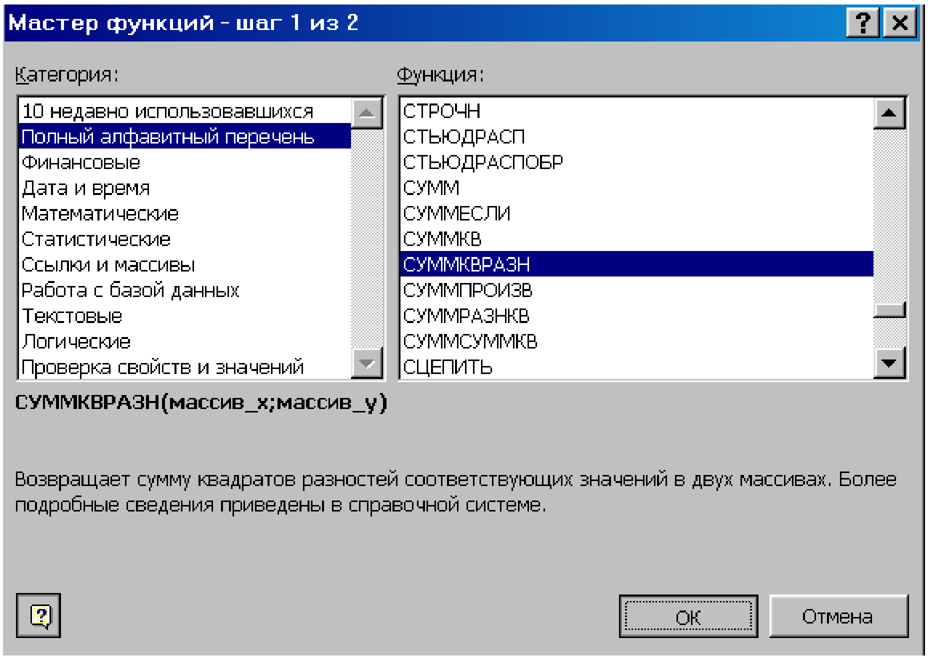
b0	b1	b2	b3	b12	b13	b23	b123
2,60195	-	0,01957	0,23297	0,00938	0,00554	-	-
2	0,10998	5	2	1	2	0,01954	0,00545

Определяем дисперсию каждого опыта по формуле:

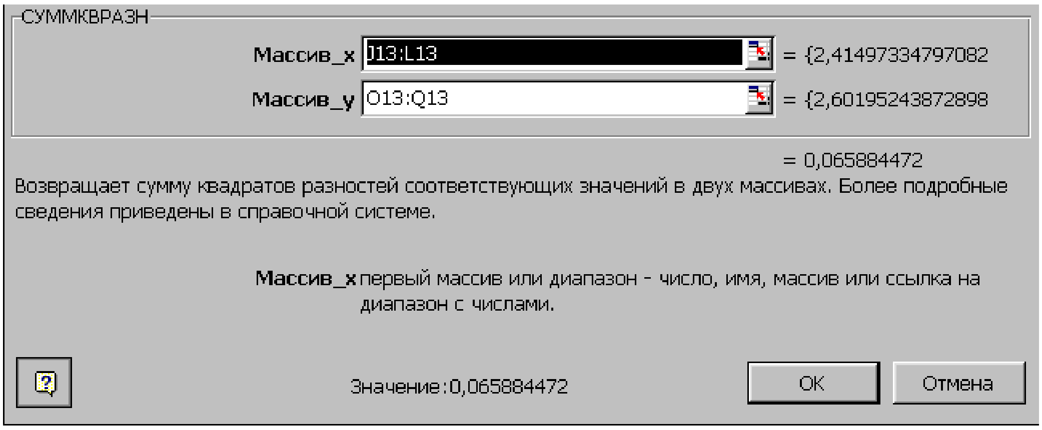
$$S^2 = \frac{1}{k-1}\sum_{i=1}^{k}(Y_i - \overline{Y_i}) \qquad 10.8$$

где к – число дублей; к=3

Выбираем команду «Вставка» в ней выбираем команду «функция», затем функцию «Сумм.кв.знач» и нажимаем ОК.



в строки Массив x и Массив y вносим данные и нажимаем ОК. Полученное число разделим на 2.



И скопируем для остальных опытов эту формулу получим данные которые заносим в таблицу. Из восьми опытов найдем сумму

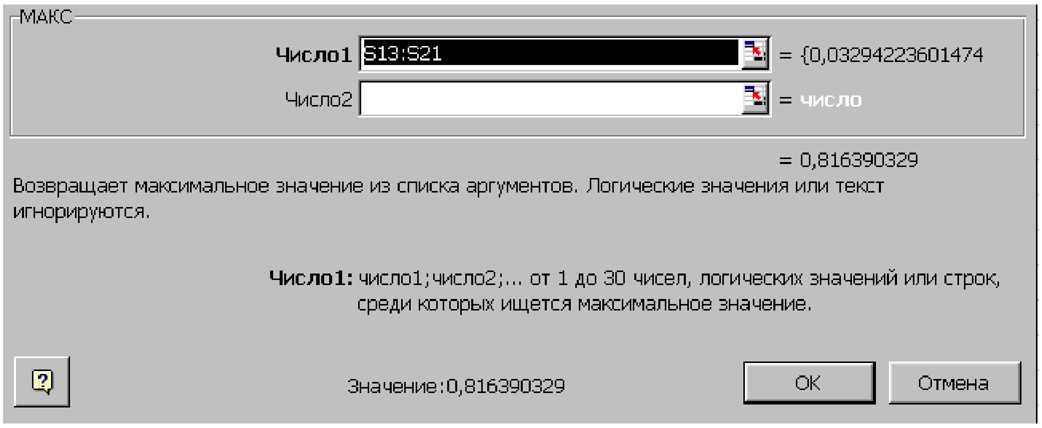
S1	0,032942
S2	0,242986
S3	0,014152
S4	0,176317
S5	0,131184
S6	0,023906
S7	0,16778
S8	0,027122
сумма S	0,81639

Определим однородность дисперсий по формуле:

$$G_P = \frac{S_{i\ \max}^2}{\sum S_i^2} \qquad 10.8$$

Для того, чтобы определить однородность дисперсий необходимо сначала найти $S_{i\ \max}^2$

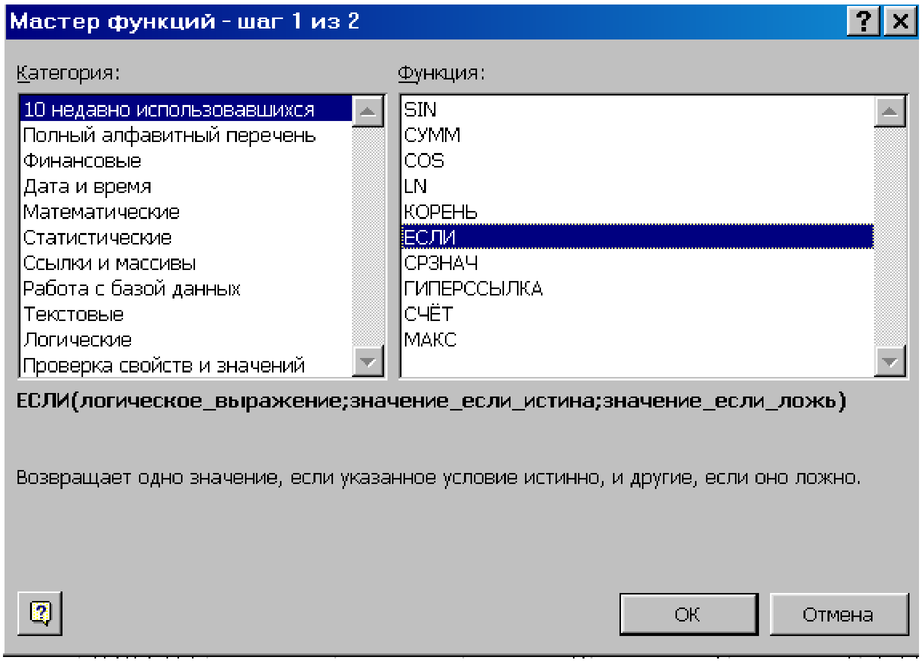
Для этого выбираем команду «Вставка» в ней выбираем команду «функция», затем функцию МАКС. появится контекстное меню вида

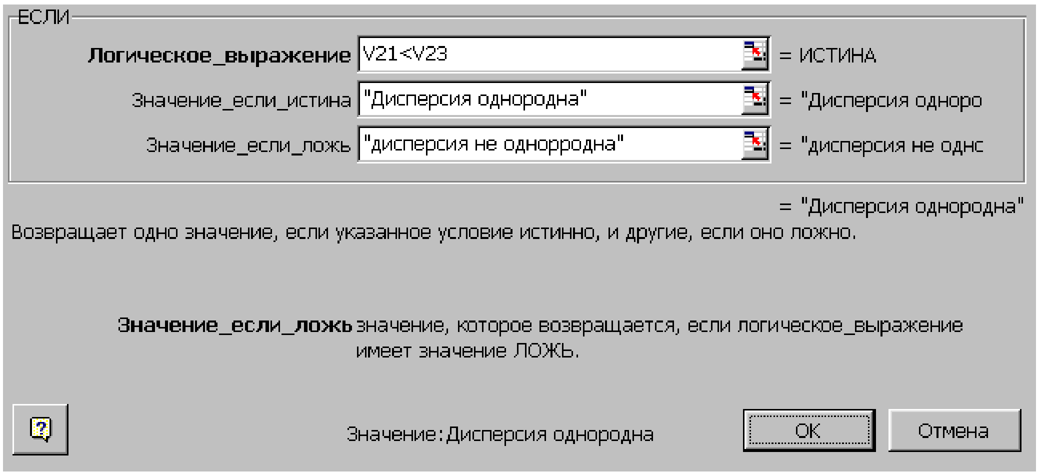


Затем нажимаем ОК. получим результат $S_{i\ \max}^2$=0,242986

Получаем, что G_p=0.297635, а табличное значение $G_{p\ табл}$=0,52

Если $G_p < G_{p\ табл}$, то дисперсия однородна, т.е.





Нажимаем ОК

Проверим значимость коэффициентов уравнения регрессии. Определим дисперсию воспроизводимости по формуле: $S^2\{Y\} = \frac{\sum S_i^2}{n}$

$S^2\{Y\}$=0,102049 т.е

$$S^2\{Y\} \quad =S21/8$$

Определяем дисперсию коэффициентов уравнения регрессии по формуле:

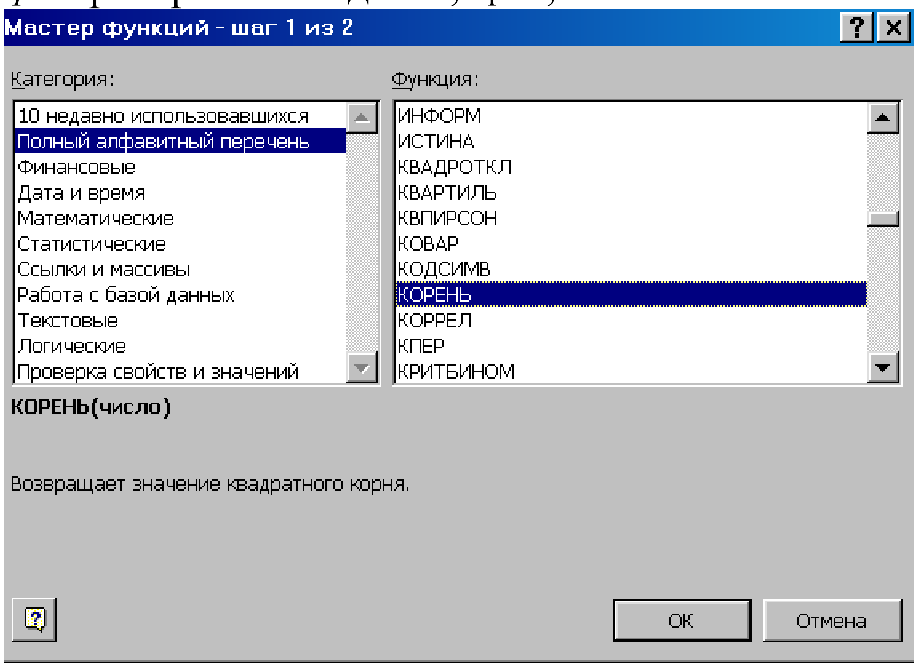
$$S^2\{b\} = \frac{S^2\{Y\}}{n*k} \quad S^2\{b\}=0{,}004252 \text{ т.е.}$$

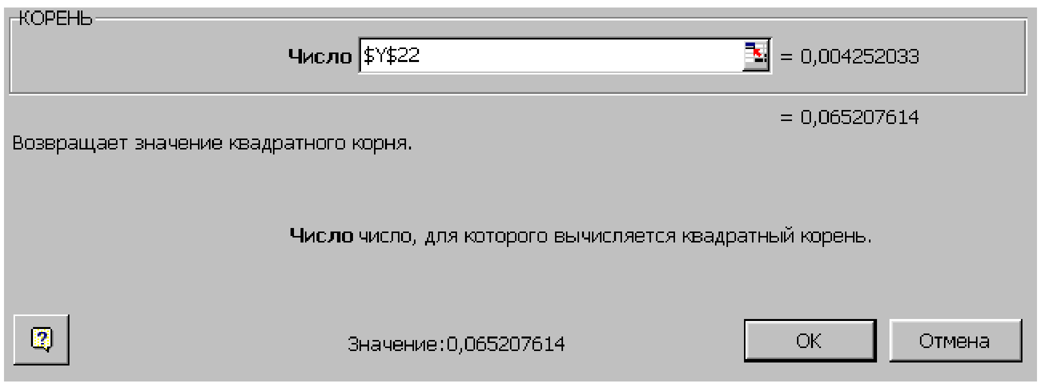
$$S^2\{b\} \quad =Y21/24$$

Определяем доверительный интервал для проверки значимости коэффициентов по формуле:

$$\Delta b = \pm t_T \sqrt{S^2\{b\}} \qquad 10.9$$

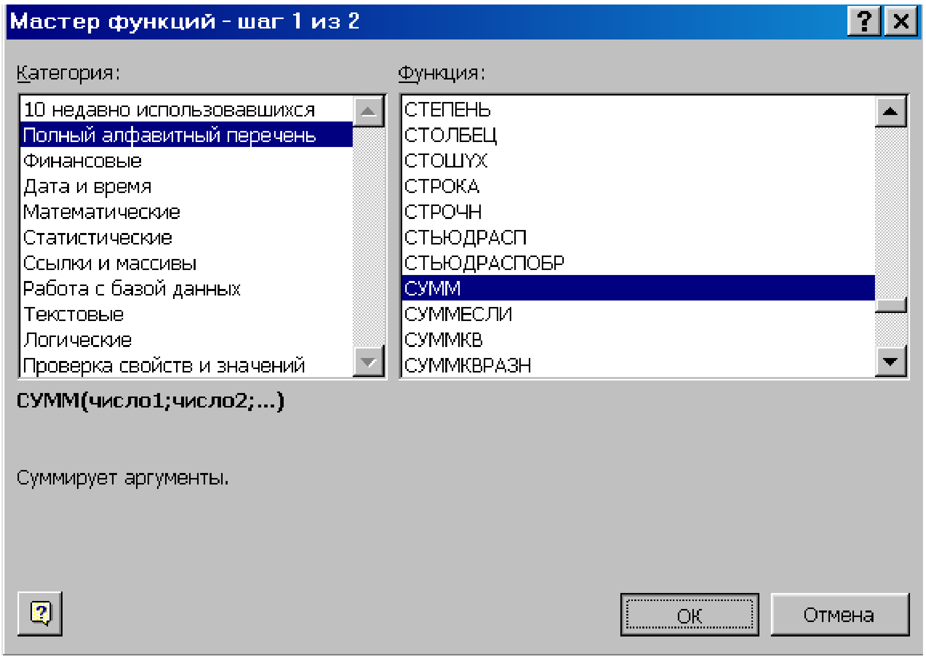
где t_T – критерий Стьюдента; t_T=2,31

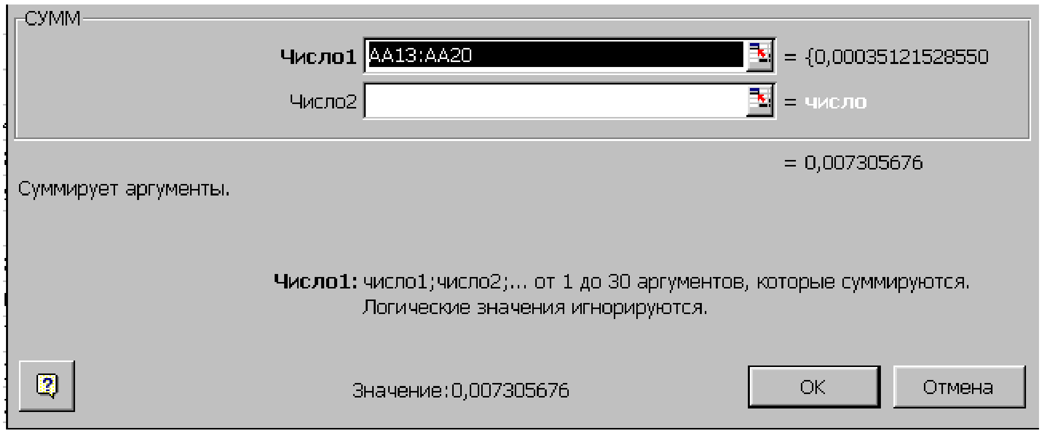




Нажимаем ОК. Затем полученный результат умножим на $Y$23. Получим результат равный: Δb=0,150629589

Затем оцениваем коэффициенты по условию: если $|b_j| > |\Delta b|$, то коэффициент значим. т.е.



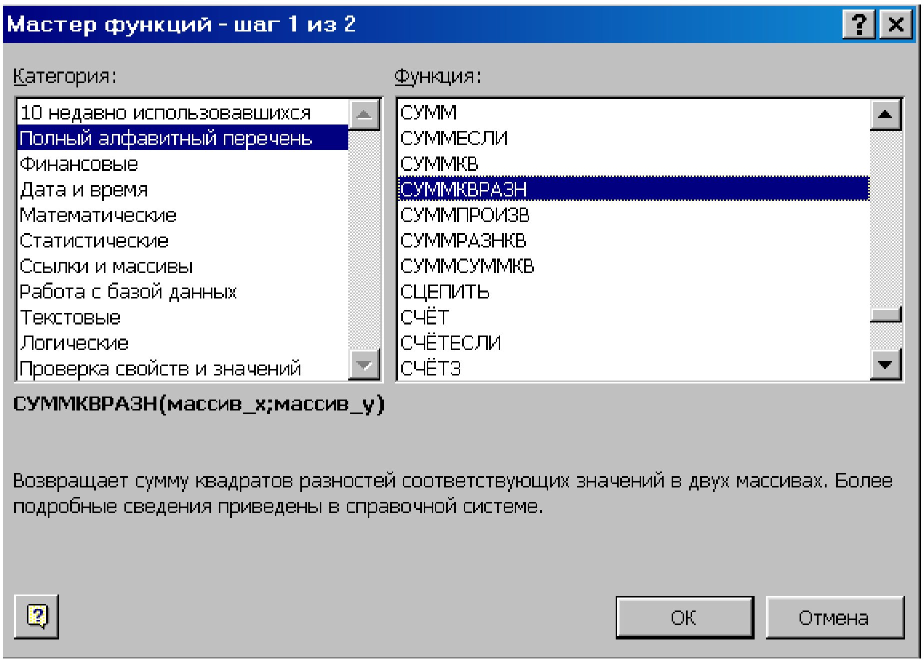


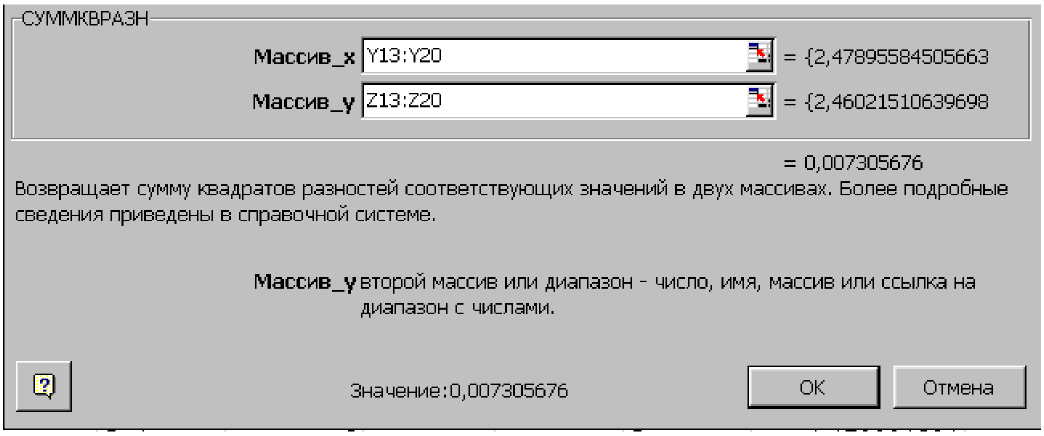
Получим адекватность полученной модели, составив вспомогательную таблицу вида:

№п/п	0	3	Ymi	Yi	(Ym-Y)^2
1	1	-1	2,478956	2,460215	0,000351215
2	1	1	2,259005	2,199524	0,003537952
3	1	-1	2,478956	2,50878	0,000889466
4	1	-1	2,9449	2,943252	2,71777E-06
5	1	1	2,259005	2,307402	0,002342319
6	1	1	2,724949	2,726517	2,45935E-06
7	1	-1	2,9449	2,935466	8,90119E-05
8	1	1	2,724949	2,734464	9,05342E-05

Получим остаточную дисперсию:

$$\Delta = \sum_{i=1}^{n} (Y_i^m - \overline{Y}_i)^2 \qquad 10.10$$





$$\Delta = 5.33729E - 0.5$$

Определим дисперсию адекватности:

$$S_{ад}^2 = \frac{k * \Delta}{n - p - 1} \qquad 10.10$$

т.е. в электронной таблице Excel формула запишется в виде:

Sад^2 =W33*$AA$22/4

тогда $S^2_{ад}$=0,005479257

Определим расчетный критерий Фишера и сравним его с табличным по формуле:

$$F_{рас} = \frac{S^2\{Y\}}{S^2_{ад}} \qquad 10.11$$

или в электронной таблице по формуле:

Fрас =($Y$21)^2/$Z$24

Fтабл 3,84

Модель адекватности рассчитывается так же как дисперсия однородности т.е. через команду Если.

Рассчитываем коэффициенты степенной модели по формуле:

$$C = 10^d$$

$$d = \sum_{j=0}^{p} b_j - 2\sum_{i=1}^{p} \frac{b_i \lg X_{\max}}{\lg X_{i\max} - \lg X_{i\min}} \qquad 10.11$$

или

d = =$B$22+$E$22-2*(($B$22*LOG10(6)/(LOG(6)-LOG(2)))+($E$22*LOG(0,5)/(LOG(0,5)-LOG(0,2))))

-

d=5,2998053

$$\alpha_i = \frac{2b_i}{\lg X_{i\max} - \lg X_{i\min}}$$

α_i= =2*($B$22/(LOG(6)-LOG(2)))

Окончательный степенной вид модели:

$$P_y = 5.01412E - 0.6 * S^{10.90688127} * t^{1.170891201}$$

Задания

1) Проверить однородность дисперсий параллельных опытов. Определить коэффициенты уравнения регрессии (степенная модель) по заданной матрице планирования эксперимента.

2)Проверить значимость коэффициентов уравнения регрессии.

3)Проверить адекватность полученной модели.

4)По результатам моделирования (пп. 1-3) сделать выводы о характере влияния факторов на исследуемый параметр и о степени пригодности и возможной области применения полученной модели.

Вариант №1

№ оп.	У. м мим	Шаг S, мм	t, мм	Sm, мкм	
1	14,1	2	0,2	55	84
2	56,5	2	0,2	65	91
3	14,1	6	п.:	205	280
4	14,1	2	п.5	14	27
5	56,5	6	0,2	218	292
6	56.5	2	0,5	19	32
7	14,1	6	0,5	87	143
8	56,5	6	0,5	95	155

Вариант №2

№ оп.	V, м/мин	Шаг S, мм	t. мм	Стойкость Т, мин	
1	15,7	2	0,2	74	95
2	61,5	2	0,2	in	17
3	15.7	6	0,2	68	89
4	15,7	2	0,5	38	60
5	61,5	6	0,2	4	8

6	61.5	2	0,5	0,8	2
7	15.7	6	(1.5	12	28
8	61.5	6	0,5	и	2,3

Вариант №3

№оп.	V, м/мин	Шаг S, мм	t, мм	Стойкость T, мин	
1	14,8	2	0,2	191	196
2	59,8	2	0,2	2	5
3	14,8	5	0,2	114	137
4	14,8	2	0,5	88	96
5	59.К	5	о.:	3	2
6	59,8	2	0,5	2	4
7	14,8	5	0,5	56	66
8	59,8	5	0,5	0,5	1.5

Вариант №4

№оп.	п. оо чин	Шаг S, мм	t, мм	Ra, мкм		
1	31,5	2	*42*	12	17	22
2	125	2	0,2	11	15	21
3	31,5	6	0,2	21	28	34
4	31,5	2	0,5	35	40	44
5	125	6	0,2	27	23	18
6	125	2	0,5	33	38	42
7	31,5	6	0,5	4:	49	56
8	125		0,5	36	41	52

Вариант №5

№оп.	V, м/мин	Шаг S, мм	t, мм	Pz.

1	9	2	0,2	260	280	330
2	37	2	0,2	160	160	155
3	9	6	0,2	320	350	300
4	9	:	0,5	810	860	770
5	37	6	o.:	220	200	190
6	37	2	0,5	540	500	560
7	9	6	0,5	870	Sim	920
8	37	6	0.5	540	580	510

Вариант №6

№оп.	V, м/мин	Шаг S, мм	t, мм	Стойкость T, мин	
1	15,1	2	0,2	134	170
2	61,5	2	0,2	10	18
3	15,1	6	0,2	32	40
4	15,1	2	0,5	44	54
5	61 5	6	0,2	4	8
6	61,5	2	0,5	3	5
7	15,1	6	0.5	24	34
8	61,5	6	0,5	2	1

Вариант №7

№оп.	V, м мим	S. мм	мм	Ra, мкм	
1	14,1	2	0,2	2,4	2,5
2	56,5	2	0,2	2,0	2,1

3	14.1	6	0,2	2,4	2,7
4	14.1	2	0,5	4,7	4,7
5	56,5	6	0,2	2,0	2,3
6	56,5	2	0,5	4,0	4,2
7	14.1	6	0,5	4,4	5,0
8	56,5	6	0,5	4.1	4,5

Вариант №8

№оп.	V. м мин	llS мм	мм	Стойкость Т, мин	
1	14 ,8	2	0,2	187	198
:	59,8	2	0,2	19	23
3	14,8	6	0,2	57	69
4	14,8	2	0,5	95	116
5	59,8	6	0,2	5	7
6	59,8	2	0,5	9	13
7	14.8	6	0,5	40	52
8	59,8	6	0.5	*2*	4

Рекомендуемая литература

1. Дворецкий, С.И. Компьютерное моделирование и оптимизация технологических процессов и оборудования: Учеб. пособие. / С.И. Дворецкий, А.Ф. Егоров, Д.С. Дворецкий – Тамбов: Изд-во Тамб. гос. техн. ун-та, 2003. 224 с.ISBN 5-8265-0213-4
2. Дерканосова, Н.М. Моделирование и оптимизация технологических процессов пищевых производств. Практикум [Текст] : учеб. пособие / Н. М. Дерканосова, А. А. Журавлев, И. А. Сорокина; Воронеж. гос. технол. акад. - Воронеж : ВГТА, 2011. - 196 с.
3. Бывальцев, А. И. Практикум по курсу «Моделирование и оптимизация технологических процессов отрасли» [Текст]: учеб. пособие / А. И. Бывальцев, Н. М. Дерканосова, А. А. Журавлев; Воронеж. гос. технол. акад. – Воронеж, 2004. – 140 с
4. Грачев, Ю. П. Математические методы планирования эксперимента [Текст] / Ю. П. Грачев, Ю. М. Плаксин. – М. : ДеЛи принт, 2005. – 296 с.
5. Изучение хлебопекарных свойств муки с использованием метода многофакторного дисперсионного анализа [Текст] / Н. М. Дерканосова, Т. Н. Малютина, В. В. Сотникова, А. А. Журавлев / Прогрессивные технологии и оборудование для пищевой промышленности: материалы II Международной научно-практ. конф. / Воронеж. гос. технол. акад. – Воронеж. 2004. – Ч. 1. – С. 85 – 87.
6. Дерканосова, Н. М. Выбор сортов пшеницы для производства макаронных изделий в условиях конфликта частных показателей качества зерна [Текст] / Н. М. Дерканосова, А. А. Журавлев, И. А. Сорокина // Материалы IV Всероссийской научно-технической конференции «Теория конфликта и ее приложения». – Воронеж, 2006. – С. 206 – 212.

Приложение 1 Критерий Кохрена

Пятипроцентные пределы для отношения наибольшей эмпирической дисперсии к сумме К эмпирических дисперсий (критерий Кохрена)

f / К	Уровень значимости 0,05											
	1	2	3	4	5	6	7	8	9	10	16	36
2	0,9985	0,9750	0,9392	0,9057	0,8772	0,8534	0,8332	0,8159	0,8010	0,7880	0,7341	0,6602
3	0,9669	0,8709	0,7977	0,7457	0,7071	0,6771	0,6530	0,6333	0,6167	0,6025	0,5466	0,4780
4	0,9065	0,7679	0,6841	0,6287	0,5895	0,5598	0,5365	0,5175	0,5017	0,4884	0,4366	0,3720
5	0,8412	0,6838	0,5981	0,5441	0,5065	0,4873	0,4564	0,4387	0,4241	0,4118	0,3645	0,3066
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
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
8	0,6798	0,5157	0,4377	0,3910	0,3595	0,3362	0,3185	0,3043	0,2926	0,2829	0,2462	0,2022
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	0,6020	0,4450	0,3733	0,3311	0,3029	0,2823	0,2666	0,2541	0,2439	0,2353	0,2032	0,1655
2	0,5410	0,3924	0,3264	0,2880	0,2624	0,2439	0,2299	0,2187	0,2098	0,2020	0,1737	0,1403
5	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	0,3894	0,2705	0,2205	0,1921	0,1735	0,1602	0,1501	0,1422	0,1357	0,1303	0,1108	0,0879
4	0,3434	0,2354	0,1907	0,1656	0,1493	0,1[illegible]74	0,1286	0,1216	0,1160	0,1113	0,0942	0,0743
0	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	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	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	0,0998	0,0632	0,0495	0,0419	0,0371	0,0337	0,0312	0,0292	0,0279	0,0266	0,0218	0,0165

Приложение 2 Критерий Фишера

ПРИЛОЖЕНИЕ

Значения F для пятипроцентного уровня значимости
(односторонний критерий)

f_1 \ f_2	I	2	3	4	5	6	I2	24
I	I64,4	I99,5	215,7	224,6	230,2	234,0	244,9	249,0
2	I8,5	I9,2	I9,2	I9,3	I9,3	I9,3	I9,4	I9,5
3	I0,I	9,6	9,3	9,I	9,0	8,9	8,7	8,6
4	7,7	6,9	6,6	6,4	6,3	6,2	5,9	5,8
5	6,6	5,8	5,4	5,2	5,I	5,0	4,7	4,5
6	6,0	5,I	4,8	4,5	4,4	4,3	4,0	3,8
7	5,6	4,7	4,4	4,I	4,0	3,9	3,6	3,4
8	5,3	4,5	4,I	3,8	3,7	3,6	3,3	3,I
9	5,I	4,3	3,9	3,6	3,5	3,4	3,I	2,9
I0	5,0	4,I	3,7	3,5	3,3	3,2	2,9	2,7
II	4,8	4,0	3,6	3,4	3,2	3,I	2,8	2,6
I2	4,8	3,9	3,5	3,3	3,I	3,0	2,7	2,5
I3	4,7	3,8	3,4	3,2	3,0	2,9	2,6	2,4
I4	4,6	3,7	3,3	3,I	3,0	2,9	2,5	2,3
I5	4,5	3,7	3,3	3,I	2,9	2,8	2,5	2,3
I6	4,5	3,6	3,2	3,0	2,9	2,7	2,4	2,2
I7	4,5	3,6	3,2	3,0	2,8	2,7	2,4	2,2
I8	4,4	3,6	3,2	2,9	2,8	2,7	2,3	2,I
I9	4,4	3,5	3,I	2,9	2,7	2,6	2,3	2,I
20	4,4	3,5	3,I	2,9	2,7	2,6	2,3	2,I
22	4,3	3,4	3,I	2,8	2,7	2,6	2,2	2,0
24	4,3	3,4	3,0	2,8	2,6	2,5	2,2	2,0
26	4,2	3,4	3,0	2,7	2,6	2,5	2,2	2,0
28	4,2	3,3	3,0	2,7	2,6	2,4	2,I	I,9
30	4,2	3,3	2,9	2,7	2,5	2,4	2,I	I,9
40	4,I	3,2	2,9	2,6	2,5	2,3	2,0	I,8
60	4,0	3,2	2,8	2,5	2,4	2,3	I,9	I,7

I26

Приложение 3 Последовательность случайных чисел

ПОСЛЕДОВАТЕЛЬНОСТЬ СЛУЧАЙНЫХ ЧИСЕЛ

[illegible]	40	04	34	63	98	99	89	31	16	12	90	50	28	96
[illegible]	02	91	27	52	98	72	03	45	65	30	89	71	54	91
[illegible]	49	14	98	53	41	92	36	07	76	85	37	84	37	47
[illegible]	03	33	48	84	37	37	29	38	37	89	76	25	09	69
[illegible]	59	96	20	30	87	31	33	69	45	58	48	00	83	48
11	[illegible]	83	24	82	24	07	78	61	89	42	58	88	22	16
[illegible]	62	41	11	59	85	18	42	61	29	88	76	04	21	80
[illegible]	05	71	70	31	31	99	99	06	96	53	99	25	13	63
19	74	15	50	17	44	80	13	86	38	40	45	82	13	44
20	84	56	19	49	59	14	85	42	99	71	16	34	33	79
[illegible]	20	68	72	98	94	62	63	59	44	03	89	06	15	87
[illegible]	08	83	83	98	40	90	88	25	26	85	74	55	80	85
23	38	25	43	32	98	94	65	35	35	16	91	07	12	43
99	71	14	12	64	51	68	50	60	78	22	69	51	98	37
33	65	95	48	75	00	06	65	25	90	16	29	34	14	43
[illegible]	06	64	75	27	29	17	06	11	30	68	70	97	87	21
42	10	25	37	30	08	27	75	43	97	54	20	69	93	50
84	11	12	66	87	48	21	06	86	08	35	39	52	28	09
92	68	50	88	17	37	92	02	23	43	63	24	69	80	91
93	08	43	30	41	86	45	74	33	78	84	33	38	76	73
46	26	39	96	33	60	20	73	30	79	17	19	03	47	28
27	98	99	24	08	94	19	15	81	29	82	14	35	88	03
01	67	56	12	69	07	89	55	63	31	50	72	20	33	36
77	80	04	24	54	67	60	10	79	26	21	60	03	48	14
96	93	67	69	37	72	22	43	46	32	56	15	75	25	12
12	67	46	72	02	59	06	17	49	12	73	28	23	52	48
39	07	46	96	40	20	86	79	11	81	74	11	15	23	17
60	21	59	12	07	04	99	88	22	39	75	16	69	13	84

Приложение 4 Критерий Стьюдента

Значения t для различных уровней значимости
(двухсторонний критерий)

Число степеней свободы	Уровни значимости				
	0,I0	0,05	0,02	0,0I	0,00I
I	6,3I	I2,7I	3I,82	63,66	636,62
2	2,92	4,30	6,97	9,93	3I,60
3	2,35	3,I8	4,54	5,84	I2,94
4	2,I3	2,78	3,75	4,60	8,6I
5	2,02	2,57	3,37	4,03	6,86
6	I,94	2,45	3,I4	3,7I	5,96
7	I,90	2,37	3,00	3,50	5,4I
8	I,86	2,3I	2,90	3,36	5,04
9	I,83	2,26	2,82	3,25	4,78
I0	I,8I	2,23	2,76	3,I7	4,59
II	I,80	2,20	2,72	3,II	4,44
I2	I,78	2,I8	2,68	3,06	4,32
I3	I,77	2,I6	2,65	3,0I	4,22
I4	I,76	2,I5	2,62	2,98	4,I4
I5	I,75	2,I3	2,60	2,95	4,07
I6	I,75	2,I2	2,58	2,92	4,02
I7	I,74	2,II	2,57	2,90	3,97
I8	I,73	2,I0	2,55	2,88	3,92
I9	I,73	2,09	2,54	2,86	3,88
20	I,73	2,09	2,53	2,85	3,85
2I	I,72	2,08	2,52	2,83	3,82
22	I,72	2,07	2,5I	2,82	3,79
23	I,7I	2,07	2,50	2,8I	3,77
24	I,7I	2,06	2,49	2,80	3,75
25	I,7I	2,06	2,48	2,79	3,73
26	I,7I	2,06	2,48	2,78	3,7I
27	I,70	2,05	2,47	2,77	3,69
28	I,70	2,05	2,47	2,76	3,67

Приложение 5 Многофакторные регулярные планы

Многофакторные регулярные планы*

Таблица 3.1

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	1	0	0	1	1	1	1	0	0	1	1	1	1	1	0	0
3	2	0	0	1	1	1	1	0	0	2	2	2	2	0	0	0
4	0	1	0	1	2	0	0	1	1	1	2	1	2	0	1	0
5	1	1	0	2	0	1	1	1	1	2	0	2	0	1	1	0
6	2	1	0	0	1	2	2	1	1	0	1	0	1	0	1	0
7	0	2	0	2	1	0	0	2	2	2	1	2	1	0	0	0
8	1	2	0	0	2	1	1	2	2	0	2	0	2	1	0	0
9	2	2	0	1	0	2	2	2	2	1	0	1	0	0	0	0
10	0	0	1	0	0	1	2	1	2	1	1	2	2	0	0	1
11	1	0	1	1	1	2	0	1	2	2	2	0	0	1	0	1
12	2	0	1	2	2	0	1	1	2	0	0	1	1	0	0	1
13	0	1	1	1	2	1	2	2	0	2	0	0	1	0	1	1
14	1	1	1	2	0	2	0	2	0	0	1	1	2	1	1	1
15	2	1	1	0	1	0	1	2	0	1	2	2	0	0	1	1
16	0	2	1	2	1	1	2	0	1	0	2	1	0	0	0	1
17	1	2	1	0	2	2	0	0	1	1	0	2	1	1	0	1
18	2	2	1	1	0	0	1	0	1	2	1	0	2	0	0	1
19	0	0	2	0	0	2	1	2	1	2	2	1	1	0	0	0
20	1	0	2	1	1	0	2	2	1	0	0	2	2	1	0	0
21	2	0	2	2	2	1	0	2	1	1	1	0	0	0	0	0
22	0	1	2	1	2	2	1	0	2	0	1	2	0	0	1	0
23	1	1	2	2	0	0	2	0	2	1	2	0	1	1	1	0
24	2	1	2	0	1	1	0	0	2	2	0	1	2	0	1	0
25	0	2	2	2	1	2	1	1	0	1	0	0	2	0	0	0
26	1	2	2	0	2	0	2	1	0	2	1	1	0	1	0	0
27	2	2	2	1	0	1	0	1	0	0	2	2	1	0	0	0

Продолжение таблицы 3.1

	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1
3	0	0	0	0	0	0	0	0	0	0	2	2	2	2	2	2
4	1	0	0	0	1	1	1	0	1	0	3	3	3	3	3	3
5	0	0	1	1	1	1	0	0	0	0	0	4	4	4	4	4
6	0	1	0	0	1	1	0	1	0	1	1	0	5	5	5	5
7	0	1	0	0	0	0	0	1	0	1	2	1	0	6	6	6
8	0	0	1	1	0	0	0	0	0	0	3	2	1	0	7	7
9	1	0	0	0	0	0	1	0	1	0	0	3	2	1	0	8
10	0	0	1	0	1	0	1	1	0	0	0	0	0	0	0	0
11	1	1	0	0	1	0	0	0	0	0	1	1	1	1	1	1
12	0	0	0	1	1	0	0	0	1	1	2	2	2	2	2	2
13	1	0	1	0	0	0	0	0	0	1	3	3	3	3	3	3
14	0	0	0	0	0	0	0	1	1	0	0	4	4	4	4	4
15	0	1	0	1	0	0	1	0	0	0	1	0	5	5	5	5
16	0	1	1	0	0	1	0	0	1	0	2	1	0	6	6	6
17	0	0	0	0	0	1	1	0	0	1	3	2	1	0	7	7
18	1	0	0	1	0	1	0	1	0	0	0	3	2	1	0	8
19	0	0	0	1	0	1	0	0	1	1	0	0	0	0	0	0

	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
20	1	1	0	0	0	1	0	0	0	0	1	1	1	1	1	1
21	0	0	1	0	0	1	1	1	0	0	2	2	2	2	2	2
22	1	0	0	1	0	0	0	1	0	0	3	3	3	3	3	3
23	0	0	0	0	0	0	1	0	0	1	0	4	4	4	4	4
24	0	1	1	0	0	0	0	0	1	0	1	0	5	5	5	5
25	0	1	0	1	1	0	1	0	0	0	2	1	0	6	6	6
26	0	0	0	0	1	0	0	1	1	0	3	2	1	0	7	7
27	1	0	1	0	1	0	0	0	0	1	0	3	2	1	0	8

Таблица 3.2

<table>
<tr><td>3</td><td>3</td><td>3</td><td>3</td><td rowspan="4">3</td><td rowspan="4">3</td><td rowspan="4">3</td><td rowspan="4">3</td><td rowspan="4">3</td><td rowspan="4">3</td><td rowspan="4">3</td><td rowspan="4">3</td><td rowspan="4">3</td></tr>
<tr><td>2</td><td>2</td><td>2</td><td>2</td></tr>
<tr><td colspan="4">4</td></tr>
<tr><td colspan="4">5</td></tr>
<tr><td colspan="4">6</td><td rowspan="4">2</td><td rowspan="4">2</td><td rowspan="4">2</td><td rowspan="4">2</td><td rowspan="4">2</td><td rowspan="4">2</td><td rowspan="4">2</td><td rowspan="4">2</td><td rowspan="4">2</td></tr>
<tr><td colspan="4">7</td></tr>
<tr><td colspan="4">8</td></tr>
<tr><td colspan="4">9</td></tr>
</table>

Таблица 3.3

<table>
<tr><td>1</td><td>2</td><td>4</td><td>5</td><td rowspan="4">3</td><td rowspan="4">6</td><td rowspan="4">7</td><td rowspan="4">8</td><td rowspan="4">9</td><td rowspan="4">10</td><td rowspan="4">11</td><td rowspan="4">12</td><td rowspan="4">13</td></tr>
<tr><td>14</td><td>15</td><td>16</td><td>17</td></tr>
<tr><td colspan="4" rowspan="2">27
28</td></tr>
<tr></tr>
<tr><td colspan="4" rowspan="4">29
30
31
32</td><td rowspan="4">16</td><td rowspan="4">19</td><td rowspan="4">20</td><td rowspan="4">21</td><td rowspan="4">22</td><td rowspan="4">23</td><td rowspan="4">24</td><td rowspan="4">25</td><td rowspan="4">26</td></tr>
<tr></tr>
<tr></tr>
<tr></tr>
</table>

Таблица 3.4

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
1	0	0	0	0	0	0	0	0	0	0	0
2	0	1	0	1	1	0	1	1	1	1	1
3	0	2	0	2	2	0	0	0	2	2	2
4	0	0	1	1	2	1	1	0	0	0	0
5	0	1	1	2	0	1	0	0	1	1	1
6	0	2	1	0	1	1	0	1	2	2	2
7	0	0	2	2	1	0	0	1	0	0	0
8	0	1	2	0	2	0	0	0	1	1	1
9	0	2	2	1	0	0	1	0	2	2	2
10	1	0	0	0	0	0	0	0	3	3	3
11	1	1	0	1	1	0	1	1	0	4	4
12	1	2	0	2	2	0	0	0	1	0	5
13	1	0	1	1	2	1	1	0	3	3	3
14	1	1	1	2	0	1	0	0	0	4	4
15	1	2	1	0	1	1	0	1	1	0	5
16	1	0	2	2	1	0	0	1	3	3	3
17	1	1	2	0	2	0	0	0	0	4	4
18	1	2	2	1	0	0	1	0	1	0	5

Таблица 3.5

<table>
<tr><td colspan="2">6</td><td rowspan="2">3</td><td rowspan="2">3</td><td rowspan="2">3</td></tr>
<tr><td colspan="2">5</td></tr>
<tr><td colspan="2">4</td><td rowspan="2">2</td><td rowspan="2">2</td><td rowspan="2">2</td></tr>
<tr><td>3</td><td>2</td></tr>
</table>

Таблица 3.6

<table>
<tr><td colspan="2">11</td><td rowspan="2">3</td><td rowspan="2">4</td><td rowspan="2">5</td></tr>
<tr><td colspan="2">10</td></tr>
<tr><td colspan="2">9</td><td rowspan="2">6</td><td rowspan="2">7</td><td rowspan="2">2</td></tr>
<tr><td>2</td><td>1</td></tr>
</table>

Как пользоваться таблицами

В таблицах 3.1 и 3.4 первый столбец — номер опыта; первая строка — номер фактора.

Сначала вы выбираете таблицу с наиболее подходящим числом опытов. Затем из нее необходимо выбрать столбцы с нужным количеством уровней варьирования. Здесь следует воспользоваться вспомогательными таблицами (3.2 и 3.3 к 3.1; 3.5 и 3.6 к 3.4). Первая из таблиц показывает количество уровней варьирования; вторая — номера соответствующих столбцов. Следует помнить, что в один план не могут входить столбцы, номера которых находятся по вертикали в одной колонке. Например (см. табл. 3.6), не могут одновременно использоваться 11 и 10 или 3 и 6. Может быть взят только один из них.