Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное
образовательное учреждение высшего образования
«Уфимский государственный авиационный технический университет»

Кафедра вычислительной математики кибернетики

ЛАБОРАТОРНЫЙ ПРАКТИКУМ

по дисциплинам
«Теория планирования эксперимента»,
«Планирование эксперимента»

Уфа 2022

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное
образовательное учреждение высшего образования
«Уфимский государственный авиационный технический университет»

Кафедра вычислительной математики кибернетики

ЛАБОРАТОРНЫЙ ПРАКТИКУМ

по дисциплинам
«Теория планирования эксперимента»,
«Планирование эксперимента»

Учебное электронное издание сетевого доступа

Уфа 2022

Автор-составитель О. С. Нургаянова

Лабораторный практикум по дисциплинам «Теория планирования эксперимента», «Планирование эксперимента» [Электронный ресурс] / Уфимск. гос. авиац. техн. ун-т ; [авт.-сост. О. С. Нургаянова]. – Уфа : УГАТУ, 2022. – URL: https://www.ugatu.su/media/uploads/MainSite/Ob%20universitete/Izdateli/El_izd/2022-2.pdf

Цель лабораторного практикума – ознакомиться с принципами многофакторного планирования эксперимента, получить навыки построения планов экспериментов и расчета функции отклика с интерпретацией результатов эксперимента.

Предназначен для студентов, обучающихся по направлениям подготовки бакалавров 02.03.03 Математическое обеспечение и администрирование информационных систем, 09.03.01 Информатика и вычислительная техника, 09.03.04 Программная инженерия.

Рецензент канд. физ.-мат. наук, доц. Р. П. Абдрахманова

При подготовке электронного издания использовались следующие программные средства:

- Adobe Acrobat – текстовый редактор;
- Microsoft Word – текстовый редактор.

Автор-составитель *О. С. Нургаянова*

Редактирование и верстка *Р. М. Мухамадиева*
Программирование и компьютерный дизайн *О. М. Толкачёва*

Подписано к использованию: 21.01.2022
Объем: 0,95 Мб.

ФГБОУ ВО «Уфимский государственный авиационный технический университет»
450008, Уфа, ул. К. Маркса, 12.
Тел.: +7-908-35-05-007
e-mail: rik@ugatu.su

Введение

Эксперимент является важнейшей частью научного исследования. Проведение эксперимента и обработка его результатов тесно связаны с такими математическими дисциплинами, как теория вероятностей и математическая статистика. «Планирование эксперимента – это задание некоторого комплекса условий, допускающего неограниченное число повторений и изучение определенного круга событий, которые могут наступать в результате осуществления этих условий. Обработка результатов – выбор некоторых данных из условий» (А. Н. Колмогоров) [2].

Лабораторный практикум по дисциплинам «Теория планирования эксперимента», «Планирование эксперимента» предназначен для студентов всех форм обучения и дополнительного образования по направлениям подготовки бакалавров 02.03.03 Математическое обеспечение и администрирование информационных систем, 09.03.01 Информатика и вычислительная техника, 09.03.04 Программная инженерия.

Целью выполнения лабораторных работ по данным дисциплинам является систематизация, расширение и закрепление общепрофессиональных и профессиональных компетенций, связанных математическим планированием эксперимента, в том числе вычислительного, а также развитие навыков практического применения полученных знаний и самостоятельного планирования и организации эксперимента – определения параметра(-ов) оптимизации и факторов, на них влияющих, количестве возможных опытов, разработки методики их рандомизации и проведения, выборе способов получения экспериментальных данных и статистической обработке с выводами об адекватности математической модели и интерпретацией соответствующих результатов, а также оформлению научного отчета по результатам исследований.

Лабораторная работа № 1

Построение планов ПФЭ типа a^k и ДФЭ типа a^{k-p}

Цель работы – научиться строить матрицы планирования полного и дробного факторного экспериментов, ознакомиться с их свойствами.

Теоретическая часть

Матрица ПФЭ реализует все возможные неповторяющиеся комбинации уровней k независимых факторов, каждый из которых может варьироваться на двух и более уровнях. Таким образом, эксперимент, в котором реализуются все возможные сочетания уровней факторов, называется *полным факторным экспериментом.* Число этих комбинаций $N=a^k$ определяет число точек факторного пространства и равно числу опытов.

Если число уровней каждого фактора равно двум, то имеем полный факторный эксперимент типа 2^k, где цифра 2 указывает на число уровней факторов в планируемом эксперименте.

Нетрудно написать все сочетания уровней в эксперименте с двумя факторами. Напомним, что в планировании эксперимента используются кодированные значения факторов: +1 и -1 (часто для простоты записи единицы опускают). Условия эксперимента записывают в виде табл. 1, где строки соответствуют различным опытам, а столбцы – значениям факторов и параметру оптимизации. Каждый столбец в матрице планирования называют *вектор-столбцом,* а каждую строку – *вектор-строкой.* Таким образом, имеем два вектор-столбца независимых переменных и один вектор-столбец параметра оптимизации.

Таблица 1

Матрица планирования для двух факторов

Номер опыта, N	Факторы, k		y
	x_1	x_2	
1	-1	-1	y_1
2	+ 1	-1	y_2
3	-1	+ 1	y_3
4	+ 1	+ 1	y_4

Если для двух факторов все возможные комбинации уровней легко найти прямым перебором, то с ростом числа факторов возникает необходимость в применении существующих приемов построения матриц.

Наиболее часто используется следующие три приема, основанные на переходе от матриц меньшей размерности к матрицам большей размерности.

Первый прием основан на правиле чередования знаков. В первом столбце знаки меняются поочередно, во втором столбце они чередуются через 2, в третьем – через 4, в четвертом – через 8 и т. д.

Второй прием. При добавлении нового фактора каждая комбинация уровней исходного плана встречается дважды: в сочетании с нижним и верхним уровнями нового фактора. То есть сначала надо записать все опыты плана для одного уровня нового фактора, допустим, верхнего (+1), а затем повторить его для нижнего уровня (-1). Эффекты x_1, x_2, и x_3 называются *линейными эффектами.* Вот как это выглядит при переходе от эксперимента 2^2 к 2^3 (табл. 2).

Таблица 2

Матрица планирования ПФЭ 2^3

Номер опыта	x_1	x_2	x_3	y
1	-	-	+	y_1
2	+	-	+	y_2
3	-	+	+	y_3
4	+	+	+	y_4
5	-	-	-	y_5
6	+	-	-	y_6
7	-	+	-	y_7
8	+	+	-	y_8

Этот прием распространяется на построение матриц любой размерности.

Третий прием основан на правиле перемножения столбцов матрицы. При построчном перемножении двух столбцов матрицы произведение единиц с одноименными знаками дает +1, а с разноименными –1. Воспользовавшись этим правилом, получим для случая, который мы рассматриваем, вектор-столбцы произведений x_1x_2, x_2x_3 и x_1x_3 (табл. 3).

Эффект взаимодействия двух факторов называется *эффектом взаимодействия первого порядка.*

Таблица 3

План полного трехфакторного эксперимента

Номер опыта, N	Матрица планирования								Вектор результатов $У$
	Факторы, X_j				Факторы взаимодействия				
	x_0	x_1	x_2	x_3	$x_1\ x_2$	$x_1\ x_3$	x_2x_3	$x_1x_2x_3$	
1	+	-	-	-	+	+	+	-	$у_1$
2	+	+	-	-	-	-	+	+	$у_2$
3	+	-	+	-	-	+	-	+	$у_3$
4	+	+	+	-	+	-	-	-	$у_4$
5	+	-	-	+	+	-	-	+	$у_5$
6	+	+	-	+	-	+	-	-	$у_6$
7	+	-	+	+	-	-	+	-	$у_7$
8	+	+	+	+	+	+	+	+	$у_8$

Далее повторим еще раз перемножение знаков уже трех факторов и получим эффект взаимодействия трех факторов, вектор-столбец $x_1x_2x_3$ которого называется *эффектом взаимодействия второго порядка. Нулевой уровень* x_0 вводится для удобства дальнейших расчетов. Вообще, *эффект взаимодействия максимального порядка* в ПФЭ имеет порядок, на единицу меньший числа факторов.

Число всех возможных эффектов (включая x_0, линейные и эффекты взаимодействия всех порядков) равно числу опытов полного факторного эксперимента.

Чтобы найти число *возможных взаимодействий* (C^*) некоторого порядка, можно воспользоваться обычной формулой числа сочетаний:

$$C_n^k = \frac{k!}{n!(k-n)!},$$

где k – число факторов; n – число элементов во взаимодействии. Так, для плана 2^3 число взаимодействий первого порядка равно трем, а для плана 2^4 – шести:

$$C_2^4 = \frac{4!}{2!(4-2)!} = 6.$$

Физический смысл эффекта взаимодействия заключается в следующем: например, пусть на процесс дублирования деталей одежды влияют два фактора: температура и время дублирования. В области низких температур увеличение времени повышает прочность клеевого соединения [2]. При переходе в область высоких температур эта закономерность нарушается. Здесь, напротив, необходимо уменьшать время дублирования, чтобы получить высокую прочность клеевого соединения. Это и есть проявление эффекта взаимодействия.

Зависимость количества взаимодействий различного порядка от числа факторов ПФЭ приведена в табл. 4. Полное число всех возможных эффектов, включая и x_0, равно числу опытов ПФЭ.

Таблица 4

Взаимосвязь характеристик ПФЭ

Число факторов, k	Число линейных эффектов (коэффициентов)	Число опытов, N=2^k	Число и порядок взаимодействий (коэффициентов)					
			1	2	3	4	5	6
2	2	4	1					
3	3	8	3	1				
4	4	16	6	4	1			
5	5	32	10	10	5	1		
6	6	64	15	20	15	6	1	
7	7	128	21	35	35	21	7	1

С увеличением числа факторов количество опытов в полном факторном эксперименте резко возрастает. Так, при трех факторах следует поставить 2^3=8 опытов, при пяти – 2^5=32 опыта, а уже при восьми – 256 опытов. Разность между числом опытов и числом коэффициентов во многих случаях оказывается очень велика, и возникает естественное желание сократить число необходимых опытов. Заманчиво сократить их число за счет той информации, которая несущественна при построении линейных моделей.

Итак, нам заранее известно, что объект описывается линейным уравнением (предполагается, что эффекты взаимодействия отсутствуют) или нас интересуют только линейные члены. Возникает вопрос: как построить ортогональный план, позволяющий определить коэффициенты линейного уравнения, который содержит меньше число опытов, чем полный факторный эксперимент, т.е. целесообразно сократить число опытов за счет информации, которую несут эффекты взаимодействия факторов и которая для построения постулируемой линейной модели несущественна. Попробуем построить такие планы для случая k=3 и k=4. Рассмотрим снова ПФЭ 2^2 (план этого эксперимента задан табл. 5).

Таблица 5

Номер опыта	x_1	x_2	x_3	$x_1 x_2 = x_3$	y
1	+1	-1	-1	+1	y_1
2	+1	+1	-1	-1	y_2
3	+1	-1	+1	-1	y_3
4	+1	+1	+1	+1	y_4

По нему можно построить модель $y=b_0+b_1x_1+b_2x_2+b_{12}x_1x_2$. Пусть имеется достаточная уверенность, что в выбранных интервалах варьирования x_1 и x_2 значение b_{12} будет незначим. Тогда столбец x_1x_2 можно использовать для нового фактора x_3. Проведя такой эксперимент, сможем сделать оценку четырех коэффициентов, т.е. получить зависимость $y=b_0+b_1x_1+b_2x_2+b_3x_3$.

Рассмотрим теперь второй случай: k=4. План полного факторного эксперимента состоит из 16 точек. Как и в предыдущем примере рассмотрим ПФЭ 2^3 (k=3), представленный в табл. 6.

Таблица 6

Номер опыта	x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3	$x_1x_2x_3$
1	+1	1	1	1	1	1	1
2	+1	-1	1	-1	1	-1	-1
3	-1	1	1	-1	-1	1	-1
4	-1	-1	1	1	-1	-1	1
5	+1	1	-1	1	-1	-1	-1
6	+1	-1	-1	-1	-1	1	1
7	-1	1	-1	-1	1	-1	1
8	-1	-1	-1	1	1	1	-1

В связи с тем, что взаимодействия принимаются равными нулю, можно воспользоваться любым из столбцов, характеризующих эффекты взаимодействия для четвертой переменной, например, столбец $x_1x_2x_3$ (или $-x_1x_2x_3$). Приняв для фактора x_4 столбец $x_1x_2x_3$ получим план, приведенный в табл. 7.

Таблица 7

Номер опыта	x_1	x_2	x_3	x_4
1	+1	1	1	1
2	+1	-1	1	-1
3	-1	1	1	-1
4	-1	-1	1	1
5	+1	1	-1	-1
6	+1	-1	-1	1
7	-1	1	-1	1
8	-1	-1	-1	-1

Рассмотренные планы содержат половину опытов ПФЭ и носят название *полуреплики*.

Таким образом, из всего множества точек полного факторного плана может быть отобрана лишь некоторая часть, представляющая так называемый *дробный факторный план* и содержащая подходящее число опытов.

Дробный факторный эксперимент составляет часть ПФЭ, которая называется дробной репликой. Используется ½реплики (полуреплика), ¼ реплики, 1/8 реплики и т. д. Условное обозначение реплик и количество опытов приведены в табл. 8.

Для построения дробного факторного плана типа 2^{k-p} из множества k факторов отбирают k-p основных, для которых строят полный факторный план с матрицей X_{k-p}. Этот план дополняют затем p столбцами (линейными эффектами), приравненных к эффектам взаимодействия. Каждый из этих столбцов получается как результат поэлементного перемножения не менее двух и не более k-p определенных столбцов, соответствующих основным факторам.

Перед постановкой эксперимента по дробным репликам необходимо решить, каким взаимодействием можно пренебречь, и к какому это приведет риску.

Таблица 8

Число факторов	Дробная реплика	Обозначение	Количество опытов	
			ДФЭ	ПФЭ
3	½ реплики от 2^3	2^{3-1}	4	8
4	½ реплики от 2^4	2^{4-1}	8	16
4	1/4 реплики от 2^4	2^{4-2}	4	16
5	½ реплики от 2^5	2^{5-1}	16	32
5	1/4 реплики от 2^5	2^{5-2}	8	32

Для определения способа образования каждого из p столбцов дробного плана вводится понятие генерирующего соотношения (*генератор плана)* – произведение основных факторов, определяющих значение элементов каждого из дополнительных р столбцов матрицы плана и *определяющие контрасты*. В случае плана типа 2^{k-p} может иметься p генераторов.

Целесообразно сократить число опытов за счет информации, которую несут эффекты взаимодействия факторов и которая для построения постулируемой линейной модели несущественна.

Рассмотренная матрица составляет полуреплику от 2^3, т.е. полуреплику 2^{3-1}.

В рассмотренном случае ДФЭ типа 2^{3-1} могут быть два генерирующих соотношения $x_3=x_1x_2$ и $x_3=-x_1x_2$.

Определяющий контраст получается умножением левой и правой частей генерирующих соотношений (ГС) на x_3 (для нашего случая), т.е. $x_3^2=1=x_1x_2x_3$ и $x_3^2=-1=-x_1x_2x_3$.

Таким образом, произведения столбцов матрицы равные +1 или -1 называются *определяющими контрастами*, которые помогают найти смешанные эффекты, не изучая матрицу планирования, и позволяет установить разрешающую способность дробной реплики (см. далее).

Дробные реплики, содержащие 2^{k-p} точек, называют *регулярными*. Существуют также нерегулярные дробные реплики, содержащие количество точек не равное 2^{k-p}, например $\frac{3}{4}2^k$.

При разработке технологических процессов дробные реплики очень широко используются на стадии крутого восхождения (при оптимизации), а также при математическом описании локальной области факторного пространства с узким интервалом изменения переменных (функцию на данном интервале можно заменить полиномом первого порядка, т.е. график функции заменить отрезком прямой). Существуют задачи, где из физических (экономических или других) соображений ясно, что эффекты взаимодействия существовать не могут.

Разрешающая способность ДФЭ задается системой смешивания данной реплики. Она будет (разрешающая способность) максимальной, если линейные эффекты смешаны с эффектами взаимодействия наибольшего возможного порядка.

Определяющий контраст позволяет определить систему смешивания дробной реплики. Для того чтобы определить, какой эффект смешан с данным, нужно помножить обе части определяющего контраста на столбец, соответствующий данному эффекту.

Для этого нужно последовательно помножить независимые переменные на *определяющий контраст*.

В рассмотренном случае ДФЭ типа 2^{3-1} могут быть два генерирующих соотношения $x_3=x_1x_2$ и $x_3=-x_1x_2$.

Определяющий контраст получается умножением левой и правой частей генерирующих соотношений (ГС) на x_3 (для нашего случая), т.е.
$x_3^2=1=x_1x_2x_3$ и $x_3^2=-1=-x_1x_2x_3$.
Для x_1 имеем $x_1=x_1^2x_2x_3=x_2x_3$,
для x_2 имеем $x_2=x_1x_2^2x_3=x_1x_3$,
для x_3 имеем $x_3=x_1x_2x=x_1x_2$.

Полученные соотношения указывают на равенство столбцов в матрице планирования, например, столбцы для x_1 и x_2x_3 одинаковы. Поэтому коэффициент b_1 будет оценивать сумму $\beta_1+\beta_{23}$. Это записывается следующим образом:
$b_1\to\beta_1+\beta_{23}$, $b_2\to\beta_2+\beta_{13}$, $b_3\to\beta_3+\beta_{12}$.

ДФЭ, в котором основные эффекты смешаны с двойными эффектами взаимодействия, носят название планов с *разрешающей способностью III* (по наибольшему числу факторов в определяющем контрасте). Чтобы получить высокую разрешающую способность реплик, нужно стремиться к такому их плану, чтобы линейные эффекты были смешаны с взаимодействиями более высокого порядка, т.к. они чаще бывают равными нулю.

Свойства полного и дробного факторных экспериментов

Полному и дробному факторным экспериментам присущи следующие свойства.

1. Симметричность относительно центра эксперимента: $\sum_{u=1}^{N} x_{iu} = 0$, т.е. сумма элементов любого столбца матрицы планирования равна нулю.
2. Нормировка: $\sum_{u=1}^{N} x_{iu}^2 = N$, т.е. сумма квадратов элементов любого столбца рана числу опытов.
3. Ортогональность: $\sum_{j=1}^{N} x_{iu} x_{ju} = 0$, т.е. сумма почленных произведений любых столбцов равна нулю.
4. Ротатабельность, т.е. способность математической модели, полученной в результате полного и дробного экспериментов, предсказывать значение параметра у с одинаковой точностью на равных расстояниях от центра эксперимента независимо от направления.

Свойства 1, 2 следуют непосредственно из способов построения матриц планирования, в частности из того, что значения факторов в матрице равны только +1 и –1.

Практическая часть

Реализовать генератор планов экспериментов ПФЭ и ДФЭ на любом языке высокого уровня. В интерфейсной части предусмотреть пользовательское задание количества факторов и уровней их варьирования, ограничив первые количеством 20, вторые – пять. Предусмотреть контроль ввода количества факторов, уровней варьирования и проверку степени дробности в ДФЭ.

Вопросы и задания для самопроверки

1. Что такое опыт, эксперимент и план эксперимента?
2. Перечислите требования, предъявляемые к факторам при планировании ПФЭ и ДФЭ.
3. Чему равно количество опытов в матрицах планирования 2^{3-1} и 2^{4-2}?
4. Перечислите свойства матрицы планирования ПФЭ.
5. О чем говорит разрешающая способность плана равная IV?
6. Сколько генерирующих соотношений можно построить для ДФЭ типа 2^{4-1}?
7. В чем заключается физический смысл взаимодействия факторов?

Требования к отчету

Отчет к лабораторной работе № 1 не требуется.

Лабораторная работа № 2

Расчет математической модели ПФЭ типа 2^3

Цель работы – научиться выполнять расчет математической модели ПФЭ типа 2^3 и интерпретировать полученные результаты.

Теоретическая часть

Задачи, для решения которых может использоваться планирование эксперимента, чрезвычайно разнообразны. К ним относится поиск оптимальных условий, построение интерполяционных формул, выбор существенных факторов, оценка и уточнение констант теоретических моделей, выбор наиболее приемлемых из некоторого множества гипотез о механизме явлений, исследование диаграмм «состав–свойство» и т. д.

Одной из наиболее распространенных научно-технических задач является поиск оптимальных условий. Она возникает в тот момент, когда установлена возможность проведения процесса и необходимо найти наилучшие с некоторой точки зрения условия его реализации. Примеров задач оптимизации множество – это и выбор оптимального состава многокомпонентных смесей и сплавов, повышение производительности действующих установок, повышение качества продукции, снижение затрат на ее получение и т. п.

Выделяются следующие этапы построения математической модели:

1) сбор и анализ априорной информации;
2) выбор факторов и выходных переменных, области экспериментирования;
3) выбор математической модели, с помощью которой будут представляться экспериментальные данные;
4) выбор критерия оптимальности и плана эксперимента;
5) определение метода анализа данных;
6) проведение эксперимента;
7) проверка статистических предпосылок для полученных экспериментальных данных;
8) обработка результатов;
9) интерпретация и рекомендации.

Факторы определяют состояние объекта. Основное требование к факторам – управляемость. Под *управляемостью* понимается установление нужного значения фактора (*уровня*) и поддержание его в течение всего опыта. В этом состоит особенность *активного эксперимента*. Факторы могут быть количественными и качественными. Примерами количественных факторов являются температура, давление, концентрация. Их уровням соответствует числовая шкала. Различные катализаторы, конструкции аппаратов, способы лечения, методики преподавания являются примерами качественных факторов. Уровням таких факторов не соответствует числовая шкала, и их порядок не играет роли.

Выходные переменные – это реакции (*отклики*) на воздействие факторов. Отклик зависит от специфики исследования и может быть экономическим (прибыль, рентабельность), технологическим (выход, надежность), психологическим, статистическим и т. д. Параметр оптимизации должен быть эффективным с точки зрения достижения цели, универсальным, количественным, выражаемым числом, имеющим физический смысл, а также простым и легко вычисляемым.

Затраты машинного времени можно значительно сократить, если на этапе оптимизации параметров использовать экспериментальную факторную математическую модель. Экспериментальные факторные модели, в отличие от теоретических, не используют физических законов, описывающих происходящие в объектах процессы, а представляют собой некоторые формальные зависимости выходных параметров от внутренних и внешних параметров объектов проектирования.

Экспериментальная факторная модель может быть построена на основе проведения экспериментов непосредственно на самом техническом объекте (физические эксперименты) либо вычислительных экспериментов на ЭВМ с теоретической моделью.

При построении экспериментальной факторной модели объект моделирования (проектируемая техническая система) представляется в виде «черного ящика» (рис. 1), на вход которого подаются некоторые переменные векторов X и U, а на выходе можно наблюдать и регистрировать переменные значения вектора Y.

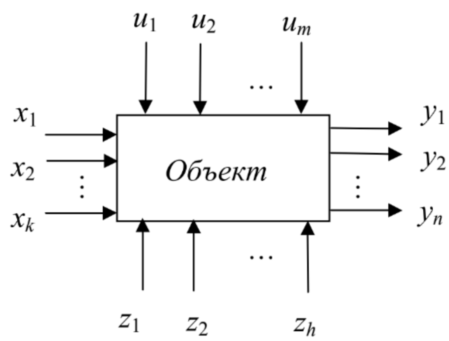

Рис. 1. Модель «черного ящика»

В процессе проведения эксперимента изменение переменных X и U приводит к изменениям выходных переменных Y. Для построения факторной модели необходимо регистрировать эти изменения и осуществить необходимую их статистическую обработку для определения параметров модели.

При проведении физического эксперимента переменными X можно управлять, изменяя их величину по заданному закону. Переменные U – неуправляемые, принимающие случайные значения. При этом значения переменных X и U можно контролировать и регистрировать с помощью соответствующих измерительных приборов. Переменные $X = (x_1, x_2, ..., x_n)$ называют контролируемыми управляемыми; переменные $U = (u_1, u_2, ..., u_m)$ – контролируемыми, но неуправляемыми, а переменные $Z = (z_1, z_2,..., z_l)$ – неконтролируемыми и неуправляемыми. Переменные X и U называют *факторами*. Факторы X являются управляемыми и изменяются как детерминированные переменные, а факторы U неуправляемые, изменяемые во времени случайным образом, т. е. U представляют собой случайные процессы. Пространство контролируемых переменных – факторов X и U – образует *факторное пространство* (на рис. 2 факторное пространство представлено двумя факторами X_1 и X_2).

Выходная переменная Y представляет собой вектор зависимых переменных моделируемого объекта. Ее называют *откликом*, а зависимость Y от факторов X и U – *функцией отклика*. Геометрическое представление функции отклика называется *поверхностью отклика* (рис. 3).

Переменная Z действует в процессе эксперимента бесконтрольно. Если предположить, что факторы X и U стабилизированы во времени и сохраняют постоянные значения, то под влиянием переменных Z функция отклика Y может меняться как систематическим, так и случайным образом. В первом случае говорят о систематической помехе, а во втором – о случайной помехе. При этом полагают, что случайная помеха обладает вероятностными свойствами, не изменяемыми во времени. Возникновение помех обусловлено ошибками методик проведения физических экспериментов, ошибками измерительных приборов, неконтролируемыми изменениями параметров и характеристик объекта и внешней среды.

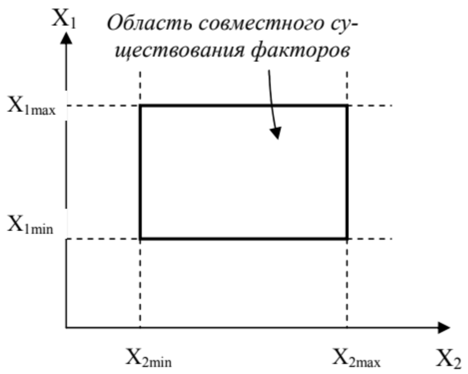

Рис. 2. Факторное пространство

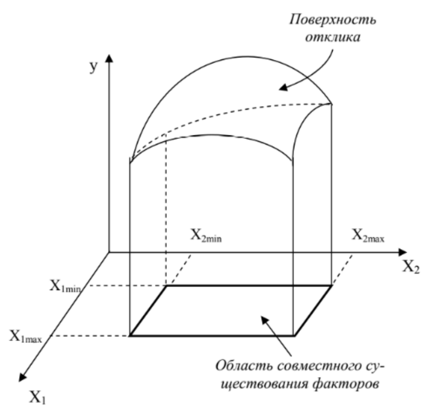

Рис. 3. Поверхность отклика

В вычислительных экспериментах объектом исследования является теоретическая математическая модель, на основе которой необходимо получить экспериментальную факторную модель. Для ее получения необходимо определить структуру и численные значения параметров модели [1].

Под структурой модели понимается вид математических соотношений между факторами X, U и откликом Y. Параметры представляют собой коэффициенты уравнений факторной модели. Структуру модели обычно выбирают на основе априорной информации об объекте с учетом назначения и последующего использования модели. Задача определения параметров модели полностью формализована. Она решается методами регрессионного анализа. Экспериментальные факторные модели называются также регрессионными моделями.

Для получения адекватной математической модели необходимо обеспечить выполнение определенных условий проведения эксперимента. Модель называют адекватной, если в оговоренной области варьирования факторов X полученные с помощью модели значения функций отклика Y отличаются от истинных не более чем на заданную величину. Цель планирования эксперимента – получение максимума информации о свойствах исследуемого объекта при минимуме опытов. Такой подход обусловлен высокой стоимостью экспериментов, как физических, так и вычислительных, и вместе с тем необходимостью построения адекватной модели. При планировании активных экспериментов используются следующие принципы:

1) отказ от полного перебора всех возможных состояний объекта;
2) постепенное усложнение структуры математической модели;
3) сопоставление результатов эксперимента с величиной случайных помех;
4) рандомизация опытов;
5) оптимальное планирование эксперимента.

Детальное представление о свойствах поверхности отклика может быть получено лишь при условии использования густой дискретной сетки значений факторов, покрывающей все факторное пространство. В узлах этой многомерной сетки находятся точки плана, в которых проводятся опыты. Выбор структуры факторной модели основан на постулировании определенной степени гладкости поверхности отклика. Поэтому с целью уменьшения количества опытов принимается небольшое число точек плана, для которых осуществляется реализация эксперимента.

При большом уровне случайной помехи получается большой разброс значений функции отклика Y в опытах, проведенных в одной и той же точке плана. В этом случае оказывается, что чем выше уровень помехи, тем с большей вероятностью простая модель окажется работоспособной. Чем меньше уровень помехи, тем точнее должна быть факторная модель.

Кроме случайной помехи при проведении эксперимента может иметь место систематическая помеха. Наличие этой помехи практически никак не обнаруживается, и результат ее воздействия на функцию не поддается контролю. Однако если путем соответствующей организации проведения опытов искусственно создать случайную ситуацию, то систематическую помеху можно перевести в разряд случайных. Такой принцип организации эксперимента называется *рандомизацией* систематически действующих помех [5].

Наличие помех приводит к ошибкам эксперимента. Ошибки подразделяются на систематические и случайные соответственно наименованиям вызывающих их факторов-помех.

Рандомизацию опытов осуществляют только в физических экспериментах. Следует отметить, что в этих экспериментах систематическую ошибку может порождать наряду с отмеченными ранее факторами также неточное задание значений управляемых факторов, обусловленное некачественной калибровкой приборов для их измерения (инструментальная ошибка), конструктивные или технологические факторы.

К факторам в активном эксперименте предъявляются определенные требования:

1) управляемость (установка заданных значений и поддержание постоянными в процессе опыта);
2) совместность (их взаимное влияние не должно нарушать процесс функционирования объекта);
3) независимость (уровень любого фактора должен устанавливаться независимо от уровней остальных);
4) однозначность (одни факторы не должны быть функцией других);
5) они должны непосредственно влиять на выходные параметры.

Выбор параметров оптимизации (критериев оптимизации) является одним из главных этапов работы на стадии предварительного изучения объекта исследования, так как правильная постановка задачи зависит от правильности выбора параметра оптимизации, являющегося функцией цели.

Под *параметром оптимизации* понимается характеристика цели, заданная количественно. Параметр оптимизации является реакцией (откликом) на воздействие факторов, которые определяют поведение выбранной системы [2–4].

Реальные объекты или процессы, как правило, очень сложны. Они часто требуют одновременного учета нескольких, иногда очень многих, параметров. Каждый объект может характеризоваться всей совокупностью параметров или любым подмножеством этой совокупности, или одним единственным параметром оптимизации. В последнем случае прочие характеристики процесса уже не выступают в качестве параметра оптимизации, а служат ограничениями. Другой путь – построение обобщенного параметра оптимизации как некоторой функции от множества исходных.

Параметр оптимизации (функции отклика) – это признак, по которому оптимизируется процесс. Множество значений, которые может принимать параметр оптимизации, называется *областью его определения*. Области определения могут быть непрерывными и дискретными, ограниченными и неограниченными. Например, выход реакции – это параметр оптимизации с непрерывной ограниченной областью определения. Он может изменяться в интервале от 0 до 100 %. Число бракованных изделий, число зерен на шлифе сплава, число кровяных телец в пробе крови – вот примеры параметров с дискретной областью определения, ограниченной снизу.

Количественная оценка параметра оптимизации на практике не всегда возможна. В таких случаях пользуются приемом, называемым ранжированием. При этом параметрам оптимизации присваиваются оценки (ранги) по заранее выбранной шкале: двухбалльной, пятибалльной и т. д. Ранговый параметр имеет дискретную ограниченную область определения. В простейшем случае область содержит два значения (да, нет; хорошо, плохо). Это может соответствовать, например, годной продукции и браку [5].

Процесс определения явного вида уравнения регрессии получил название *регрессионного анализа*. Для различных математических планов эксперимента уравнение регрессии содержит различные составляющие:

а) для планов первого порядка уравнение регрессии включает линейные эффекты и парные взаимодействия:

$$y=b_0+b_1X_1+b_2X_2+\ldots+b_kX_k+b_{12}X_1X_2+\ldots+b_{k-1,k}X_{k-1}X_k \tag{1}$$

б) для планов второго порядка уравнение регрессии включает линейные эффекты, парные взаимодействия и квадратичные эффекты:

$$y=b_0+b_1X_1+\ldots+b_kX_k+b_{12}X_1X_2+\ldots+b_{k-1,k}X_{k-1}X_k+b_{11}X_1^2+\ldots+b_{kk}X_k^2, \tag{2}$$

где b_0 – свободный член уравнения регрессии; b_k, $b_{12}\ldots b_{k-1,k}$, $b_{11}\ldots b_{kk}$ – коэффициенты регрессии; X_k – условное значения фактора x_k.

Предположим, что изучается влияние ряда факторов x_i (i=1,..., k) на некоторую величину y. Для этого проводятся эксперименты по определенному плану, который позволяет реализовать все возможные комбинации факторов. Причем каждый фактор рассматривается лишь на двух фиксированных уровнях (верхнем и нижнем). Число всех опытов в этом случае будет равно $N = 2^k$, где k – количество изучаемых факторов. Постановка опытов по такому плану называется полным факторным экспериментом типа 2^k (ПФЭ 2^k). План проведения экспериментов записывается в виде матрицы планирования, в которой в определенном порядке перечисляются различные комбинации факторов на двух уровнях. Например, в табл. 9 приведена матрица планирования ПФЭ 2^3 для трех факторов: x_1, x_2, x_3. Знак «+» говорит о том, что во время опыта значение фактора устанавливают на верхнем уровне, а знак «–» показывает, что значение фактора устанавливают на нижнем уровне.

Таблица 9

Матрица планирования ПФЭ 2^3

Номер опыта	x_1	x_2	x_3	y
1	-	-	+	y_1
2	+	-	+	y_2
3	-	+	+	y_3
4	+	+	+	y_4
5	-	-	-	y_5
6	+	-	-	y_6
7	-	+	-	y_7
8	+	+	-	y_8

При проведении экспериментов получают значения исследуемой величины y для каждого опыта (или серии опытов). Затем переходят к построению математической модели. Под моделью понимается вид функции $y = f(x_1, x_2, \ldots, x_k)$, которая связывает изучаемый параметр со значениями факторов, лежащих в интервале между верхним и нижним уровнями. Эту функцию называют *уравнением регрессии*. Для обработки результатов проведенных экспериментов и дальнейшего определения коэффициентов уравнения регрессии факторы приводят к одному масштабу. Это достигается путем кодирования переменных. Обозначим нижний уровень фактора x_i через «–», а верхний уровень – через «+». Тогда новые кодированные переменные X_i будут определяться через x_i по формуле:

$$X_i = \frac{x_i - x_i^0}{\Delta x}, \tag{3}$$

где x_i – натуральное значение i-го фактора; x_i^0 – натуральное значение i-го фактора на основном уровне; Δx – интервал варьирования i-го фактора. При таком кодировании все новые переменные будут принимать значения от –1 до +1. Линейное уравнение регрессии относительно новых переменных имеет вид:

$$y = b_0 + b_1 X_1 + b_2 X_2 + \ldots + b_k X_k \tag{4}$$

Если требуется изучить влияние парных взаимодействий различных факторов на исследуемый параметр, то уравнение регрессии записывают в виде (1). Прежде чем определять коэффициенты выбранной модели, матрицу планирования записывают относительно новых переменных, затем дополняют (если это требует вид выбранного уравнения регрессии) столбцами знаков «+» и «–», соответствующих уровням, на которых будут находиться взаимодействия факторов. Знаки этих столбцов получают с помощью исходной матрицы планирования (табл. 10).

Таблица 10

Расширенная матрица планирования для обработки результатов ПФЭ 2^3

Номер опыта, N	Порядок реализации опытов			Матрица планирования								Результаты воспроизведения серий			Вектор усредненных результатов параметра оптимизации
				Факторы, x_j				Факторы взаимодействия							
	I серия	*II серия*	*III серия*	x_0	x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3	$x_1x_2x_3$	*I*	*II*	*III*	$\bar{y}$
1	1	3	7	+	-	-	-	+	+	+	-	y_1^I	y_3^{II}	y_7^{III}	y_1
2	2	4	8	+	+	-	-	-	-	+	+	y_2^I	y_4^{II}	y_8^{III}	y_2
3	3	2	6	+	-	+	-	-	+	-	+	y_3^I	y_2^{II}	y_6^{III}	y_3
4	4	1	5	+	+	+	-	+	-	-	-	y_4^I	y_1^{II}	y_5^{III}	y_4
5	5	6	2	+	-	-	+	+	-	-	+	y_5^I	y_6^{II}	y_2^{III}	y_5
6	6	8	3	+	+	-	+	-	+	-	-	y_6^I	y_8^{II}	y_3^{III}	y_6
7	7	5	1	+	-	+	+	-	-	+	-	y_7^I	y_5^{II}	y_1^{III}	y_7
8	8	7	4	+	+	+	+	+	+	+	+	y_8^I	y_7^{II}	y_4^{III}	y_8

Обычно проводят несколько серий опытов для каждого эксперимента, что необходимо для проверки уравнения на адекватность. *Адекватность* – это способность модели предсказывать результаты эксперимента в некоторой области с требуемой точностью. Результаты опытов в каждом j-м эксперименте (j = 1, ..., n) записываются в правые столбцы матрицы планирования. В последнем столбце записываются средние выборочные значения полученных результатов для каждой серии опытов. Если каждый эксперимент повторяли m раз, то в матрице будет записано m столбцов y_1, y_2, ..., y_m. Например, в табл. 2 видно, что каждый эксперимент повторялся три раза, т. е. m = 3. Коэффициенты уравнения регрессии находятся с помощью метода наименьших квадратов. Так как матрица планирования ПФЭ 2^k удовлетворяет определенным требованиям, то формулы, определяющие коэффициенты уравнения регрессии, достаточно просты:

$$b_0 = \frac{1}{N}\sum_{j=1}^{N}\overline{y_J} \tag{5}$$

$$b_i = \frac{1}{N}\sum_{j=1}^{N} x_{ij}\cdot\overline{y_J} \tag{6}$$

$$b_{im} = \frac{1}{N}\sum_{j=1}^{N} x_{ij}\cdot x_{jm}\cdot\overline{y_J} \tag{7}$$

$$b_{ilm} = \frac{1}{N}\sum_{j=1}^{N} x_{ij}\cdot x_{jl}\cdot x_{jm}\cdot\overline{y_J} \tag{8}$$

Некоторые из коэффициентов регрессии могут оказаться пренебрежимо малыми – незначимыми. Чтобы установить, значим коэффициент или нет, необходимо вычислить оценку дисперсии, с которой он находится:

$$S_{\{y\}}^2 = \frac{1}{N}\sum_{j=1}^{N} S_j^2 \tag{9}$$

Для проверки воспроизводимости опытов находится отношение наибольшей из оценок дисперсий к сумме всех оценок дисперсий (расчетное значение критерия Кохрена):

$$G_p = \frac{\max S_j^2}{\sum_{j=1}^{N} S_j^2} \tag{10}$$

Табулированные значения критерия Кохрена $G_{\text{таб}}$ приведены в прил. 2. Для нахождения $G_{\text{таб}}$ необходимо знать уровень значимости p, общее количество оценок дисперсий N и число степеней свободы f, связанных с каждой из них, причем $f = k - 1$. При выполнении условия $G_{\text{расч}} \le G_{\text{таб}}$ опыты считаются воспроизводимыми, а оценки дисперсий – однородными. Если опыты невоспроизводимы, то можно попытаться достигнуть воспроизводимости выявлением и устранением источников нестабильности эксперимента, а также использованием более точных методов и средств измерений. Наконец, если никакими способами невозможно достигнуть воспроизводимости, то математические методы планирования к такому эксперименту применять нельзя.

Следует отметить, что с помощью ПФЭ все коэффициенты определяются с одинаковой погрешностью. Значимость каждого коэффициента уравнения регрессии устанавливается с помощью критерия Стьюдента (прил. 3) вычислением его расчетного значения:

$$t_{\text{расч}} = \frac{|b|}{\sqrt{S^2_{\{y\}}}}, \tag{11}$$

где b – коэффициент уравнения регрессии, для которого устанавливается значимость. Каждое рассчитанное значение $t_{\text{расч}}$ сравнивают с $t_{\text{таб}}$ – табличным значением критерия Стьюдента, которое выбирают для заданного уровня значимости p при числе степеней свободы f=N(k–1). Если выполняется условие $t_{\text{расч}} \geq t_{\text{таб}}$ то коэффициент считается значимым. В противном случае коэффициент регрессии незначим, и соответствующий член можно исключить из уравнения регрессии. Полученное уравнение регрессии, остается только проверить на адекватность с помощью критерия Фишера (прил. 4), который представляет собой отношение:

$$F_{\text{расч}} = \frac{\max(S^2_{\text{ад}}; S^2_y)}{\min(S^2_{\text{ад}}; S^2_y)}, \tag{12}$$

где $S^2_{\text{ад}}$ – оценка дисперсии адекватности, которая вычисляется как:

$$S^2_{\text{ад}} = \frac{1}{N-B}\sum_{j=1}^{N}(y_j^{\text{эксп}} - y_j^{\text{расч}})^2, \tag{13}$$

где $y^{эксп}$, $y^{расч}$ – экспериментальное и расчетное значения функции отклика, полученные в j-м опыте; B – количество коэффициентов в уравнении регрессии, оставшееся после удаления незначимых коэффициентов. При вычислении расчетного значения критерия Фишера по формуле (12) в числителе указывается наибольшая, а в знаменателе – наименьшая из всех оценок дисперсий. Уравнение регрессии адекватно описывает результаты эксперимента, если выполняется условие $F_{\text{расч}} < F_{\text{табл}}$, где $F_{\text{табл}}$ – табличное значение критерия Фишера для принятого уровня значимости p и числа степеней свободы f_1 числителя и f_2 знаменателя. Если гипотеза об адекватности отвергается, необходимо перейти к более сложной форме или провести эксперимент с меньшим интервалом варьирования факторов. Анализ результатов предполагает интерпретацию полученной модели. Интерпретацию модели можно производить только тогда, когда она записана в кодированных переменных. Только в этом случае на коэффициенты не влияет масштаб факторов и можно по величине коэффициентов судить о степени влияния того или иного фактора. Чем больше абсолютная величина коэффициента, тем больше фактор влияет на отклик (изучаемый параметр). Следовательно, можно расположить факторы по величине их влияния. Знак «+» у коэффициента свидетельствует о том, что с увеличением значения фактора растет величина отклика, а при знаке «–» – убывает.

Для получения математической модели в натуральных переменных x_i в уравнение регрессии вместо X_i необходимо подставить их выражения.

При переходе к натуральным переменным коэффициенты уравнения изменяются, и в этом случае пропадает возможность интерпретации влияния факторов по величинам и знакам коэффициентов. Однако если уравнение адекватно, то с его помощью можно определять значения исследуемой величины, не проводя эксперимента и придавая факторам значения, которые должны лежать между нижним и верхним уровнем.

Практическая часть

Согласно заданному номеру варианта (прил.1; вариант определяется как сумма последних двух цифр зачетной книжки) построить математическую модель некоторого исследуемого процесса или объекта, проведя необходимые вычисления коэффициентов и проверку адекватности модели. Проинтерпретировать полученные результаты. Сформулировать соответствующие выводы.

Вопросы и задания для самопроверки

1. Что представляет собой факторное пространство, какие требования предъявляются к факторам и их взаимному влиянию?
2. Перечислите этапы построения математической модели.
3. Какая математическая модель соответствует матрице планирования полнофакторного эксперимента 2^4?
4. Перечислите требования, предъявляемые к параметру оптимизации и функции отклика.
5. Как проверяется воспроизводимость эксперимента?
6. Является ли план, построенный по матрице планирования ПФЭ ротатабельным?
7. Как можно определить значимость коэффициентов регрессионного уравнения?
8. С помощью какого критерия проверяется адекватность математической модели?

Требования к отчету

Отчет к лабораторной работе № 2 выполняется в программе Microsoft Word, в формате А4, шрифт Times New Roman, размер – 14 pt, междустрочный интервал – 1,0. Расчеты могут выполняться с использованием программ Microsoft Excel, Statistica и т. п.

Требования к содержательной части отчета: рекомендуется придерживаться тех же обозначений параметров, что приводятся в теоретической части. Каждый этап расчета сопровождать пояснениями и комментариями. В конце отчета сформулировать общие выводы по лабораторной (расчетно-графической) работе.

Лабораторная работа № 3

Построение планов на латинских квадратах и квадратах Юдена

Цель работы – научиться строить матрицы планирования на основе латинских квадратов и квадратов Юдена, ознакомиться с их свойствами.

Теоретическая часть

Среди множества планов экспериментов, используемых в исследованиях, особого внимания заслуживают планы типа так называемых латинских квадратах. Латинские квадраты представляют собой планы, в которых участвуют не все комбинации уровней факторов. Планы на латинских квадратах просты в применении и используются в тех случаях, когда интересующие факторы имеют одинаковое количество уровней *n* и заранее известно, что между этими факторами отсутствуют взаимодействия, либо этими взаимодействиями можно пренебречь. Пример плана типа латинского квадрата для трехфакторного четырех уровневого эксперимента приведен в табл. 11.

Таблица 11

План эксперимента типа латинского квадрата

Уровни фактора A	Уровни фактора B			
	b_1	b_2	b_3	b_4
a_1	c_1 y_{111}	c_2 y_{122}	c_3 y_{133}	c_4 y_{144}
a_2	c_2 y_{212}	c_3 y_{223}	c_4 y_{234}	c_1 y_{241}
a_3	c_3 y_{313}	c_4 y_{324}	c_1 y_{331}	c_2 y_{342}
a_4	c_4 y_{414}	c_1 y_{421}	c_2 y_{432}	c_3 y_{443}

Здесь два фактора *A* и *B* представлены соответственно строками и столбцами и как бы образуют шахматную доску. Каждая строка и каждый столбец данной таблицы отвечают одному из уровней соответствующего фактора. Третий фактор *C* отражен в клетках таблицы буквами-уровнями c_k. Эти буквы распределены в клетках таблицы таким образом, что в каждой строке и в каждом столбце соответствующая буква встречается только один раз и, следовательно, каким бы ни было нарушающее влияние источников неоднородностей (случайных неучтенных факторов), оно в равной мере скажется при подсчете средних значений как по строке, так и по столбцу. Поэтому латинский квадрат позволяет осуществлять двойной контроль дисперсии экспериментальных данных, т. е. контроль эффектов столбцов и строк. Значения y (с индексами i, j, k) записанные в клетках таблицы – это экспериментальный результат (параметр отклика), полученный при i-м уровне фактора A, j-м уровне фактора B и k-м уровне фактора C, где индексы i, j, k могут принимать значения от 1 до 4.

С точки зрения факторных планов латинский квадрат можно рассматривать как пример неполного факторного эксперимента. Наблюдения проводятся в n^2 из n^3 возможных совокупностей условий. Опытов требуется в n раз меньше, чем при полном факторном эксперименте. Ценой такого сокращения количества опытов, без потери возможности раздельной оценки эффектов изменения уровней каждого фактора, является не учет их взаимодействия.

При планировании по схеме латинского квадрата изучается влияние четырех источников дисперсии: первый источник – строка (фактор A), второй источник – столбец (фактор B), третий источник – латинская буква «с», записанная в квадратах таблицы (фактор C) и четвертый источник – ошибка эксперимента [4].

Схема проведения дисперсионного анализа для латинского квадрата $n \times n$, приведена в табл. 12.

Здесь n – количество уровней каждого из трех факторов A, B и C. Суммы квадратов наблюдений S_a, S_b, S_c для каждого из этих факторов вычисляются совершенно аналогично. Остаточная сумма квадратов $S_{ост}$ находится путем вычитания из общей суммы квадратов наблюдений $S_{общ}$ суммы S_a,$+S_b$,$+S_c$. Разделив каждую из этих пяти сумм квадратов наблюдений на соответствующее число степеней свободы получим так называемые исправленные выборочные дисперсии MS_a, MS_b, MS_c, $MS_{ост}$ и $MSобщ$. Здесь $MS_{общ}$ – это общая дисперсия, $MS_{ост}$ – дисперсия ошибки, предположительно нормально распределенной, а MS_a, MS_b, MS_c – дисперсии, обусловленные влиянием соответственно факторов A, B и C на результаты измерений. Фактически все эти три фактора входят в план эксперимента симметрично: в любом плане типа латинского квадрата факторы, представляемые строками, столбцами и буквами, можно переставлять в любом порядке, при этом также получая латинские квадраты.

Таблица 12

Схема дисперсионного анализа планов на латинских квадратах

Факторы	**Сумма квадратов**	**Число степеней свободы**	**Средний квадрат (дисперсия MS)**	**Отношение средних квадратов** ($F_{набл}$)
A	$S_a = n^{-1}\sum_{i=1}^{n}\left(\sum_{j,k=1}^{n} y_{ijk}\right)^2 - n^{-2}\left(\sum_{i,j=1}^{n} y_{ijk}\right)^2$	$n-1$	$MS_a = \frac{S_a}{(n-1)}$	$\frac{MS_a}{MS_{ост}}$
B	$S_b = n^{-1}\sum_{j=1}^{n}\left(\sum_{i,k=1}^{n} y_{ijk}\right)^2 - n^{-2}\left(\sum_{i,j=1}^{n} y_{ijk}\right)^2$	$n-1$	$MS_b = \frac{S_b}{(n-1)}$	$\frac{MS_b}{MS_{ост}}$
C	$S_c = n^{-1}\sum_{k=1}^{n}\left(\sum_{i,j=1}^{n} y_{ijk}\right)^2 - n^{-2}\left(\sum_{i,j=1}^{n} y_{ijk}\right)^2$	$n-1$	$MS_c = \frac{S_{\tilde{n}}}{(n-1)}$	$\frac{MS_c}{MS_{ост}}$
Сумма	$S_{общ} = \sum_{i,j=1}^{n}\left(y_{ijk}\right)^2 - n^{-2}\left(\sum_{i,j=1}^{n} y_{ijk}\right)^2$	n^2-1	$MS_{общ} = \frac{S_{общ}}{n^2-1}$	–
Остаток	$S_{ост} = S_{общ} - \left(S_a + S_b + S_c\right)$	$(n-1)\cdot$ $(n-2)$	$MS_{ост} =$ $= \frac{S_{ост}}{(n-1)(n-2)}$	–

При дисперсионном анализе значимости влияния факторов А, В, С на значения параметра отклика использовался критерий Фишера F, который позволяет сравнивать величины выборочных дисперсий MS двух независимых выборок. Для вычисления $F_{набл}$ – наблюдаемого значения критерия F нужно найти отношение дисперсий двух выборок, причем таким образом, чтобы большая по величине дисперсия находилась бы в числителе, а меньшая в знаменателе. Так как, согласно условию критерия, величина числителя должна быть больше или равна величине знаменателя, то значение $F_{набл}$ всегда будет больше или равно единице.

Если это не так и $F_{набл}$ <1, то в этом случае для применения критерия F следует использовать обратную величину $(F_{набл})^{-1}$.

По таблице критических точек распределения Фишера, для заданного уровня значимости α и чисел степеней свободы k_1 и k_2, где k_1 – число степеней свободы большей дисперсии, стоящей в числителе отношения двух дисперсий, а k_2 – число степеней свободы меньшей дисперсии, стоящей в знаменателе, находят критическую точку $F_{кр}$. Если, в результате эксперимента будет получено, что $F_{набл} > F_{кр}$, то нулевая гипотеза H_0 о значимости влияния факторов принимается, в противном случае принимается альтернативная гипотеза H_1 о не значимости влияния соответствующих факторов. Если же результаты дисперсионного анализа указывают на значимость влияния линейных эффектов (факторов), т.е. на значимость различия в средних, то возникает вопрос, какие именно средние значения различны? Для проверки различия средних значений применяются различные критерии, в частности можно воспользоваться ранговым критерием Дункана.

Существуют неполные латинские квадраты, в которых число столбцов не совпадает с числом строк и обработок. Большинство таких планов были разработаны Юденом, поэтому соответствующие квадраты носят его имя. Квадрат Юдена представляет собой, в общем случае, латинский квадрат, в котором не хватает хотя бы одного столбца (или строки, или диагонали), в табл. 3 приведен пример квадрата Юдена, полученный из матрицы латинского квадрата путем удаления второй строки [4].

Таблица 13

План эксперимента квадрата Юдена

Уровни фактора A	Уровни фактора B			
	b_1	b_2	b_3	b_4
a_1	c_1 y_{111}	c_2 y_{122}	c_3 y_{133}	c_4 y_{144}
a_3	c_3 y_{313}	c_4 y_{324}	c_1 y_{331}	c_2 y_{342}
a_4	c_4 y_{414}	c_1 y_{421}	c_2 y_{432}	c_3 y_{443}

При произвольном выбрасывании более одного столбца из латинского квадрата может нарушиться его сбалансированность. В общем случае квадрат Юдена соответствует симметричному сбалансированному неполноблочному плану (число обработок совпадает с числом блоков), при чем строки представляют собой блоки, а каждая обработка встречается в каждом столбце или каждом «положении» блока только один раз.

Практическая часть

Реализовать генератор планов экспериментов на латинских квадратах и квадратах Юдена на любом языке высокого уровня. В интерфейсной части предусмотреть пользовательское задание количества факторов и соответственно уровней их варьирования. Предусмотреть контроль ввода количества факторов и уровней варьирования.

Вопросы и задания для самопроверки

1. В каких случаях прибегают к планированию эксперимента по схеме латинских квадратов?
2. Перечислите требования, предъявляемые к факторам при планировании по схеме латинского квадрата.
3. В каких случаях можно перейти к планированию на квадратах Юдена?
4. Каким образом проверяется значимость влияния факторов на функцию отклика при планировании по схеме латинского квадрата и квадрата Юдена?
5. В чем заключаются идея блочного планирования?

Требования к отчету

Отчет к лабораторной работе № 3 не требуется.

Тест для самопроверки

1. Объект исследования – это:

а) то, на что направлена человеческая деятельность;
б) человек, проводящий эксперимент;
в) эксперименты и вычисления, проводимые субъектом.

2. Моделирование может быть определено как:

а) замещение одного объекта другим с целью получения информации о важнейших свойствах объекта-оригинала с помощью объекта-модели;
б) проведение экспериментов над объектом;
в) представление объекта моделью для получения информации об этом объекте путем проведения экспериментов с его моделью.

3. Математические модели бывают:

а) классические;
б) стационарные и динамические;
в) линейные.

4. Адекватность модели зависит от:

а) выбранных критериев;
б) прогноза, полученного при помощи данной модели;
в) применения ЭВМ в исследовании.

5. В планировании эксперимента функция *Y=F(X)* называется:

а) функцией активации;
б) функцией отклика;
в) функцией надежности.

6. Математическая модель и описываемый ею реальный процесс:

а) тождественны между собой;
б) математическая модель проще, чем реальный процесс;
в) математическая модель всегда адекватна реальному процессу.

7. План эксперимента является рототабельным, если

а) общее число опытов равно числу неизвестных коэффициентов полинома;
б) если дисперсия отклика одинакова на одном расстоянии от центра плана при любом направлении в факторном пространстве;
в) он (план) насыщенный.

8. План эксперимента является насыщенным, если

а) общее число опытов равно числу неизвестных коэффициентов полинома;

б) если дисперсия отклика одинакова на одном расстоянии от центра плана при любом направлении в факторном пространстве;

в) он (план) имеет более двух уровней варьирования.

9. Сколько различных значений принимает каждый фактор при реализации ПФЭ?

а) 1;
б) 2;
в) 3.

10. Для чего необходима рандомизация экспериментов?

а) для исключения искажения модели, вызываемого систематическими составляющими возмущения;

б) для снижения порядка модели;

в) для упрощения вычислений.

11. Какая статистика вычисляется при проверке значимости коэффициента модели?

а) G-статистика;
б) t-статистика;
в) F-статистика.

12. Какая статистика вычисляется при проверке адекватности модели?

а) G-статистика;
б) t-статистика;
в) F-статистика.

13. Оценки коэффициентов модели являются:

а) случайными величинами;
б) неслучайными величинами;
в) зависит от обстоятельств.

14. Коэффициенты модели являются:

а) случайными величинами;
б) неслучайными величинами;
в) зависит от обстоятельств.

15. Матрица плана ПФЭ является ортогональной?

а) да;

б) нет;

в) это зависит от числа факторов.

16. Матрица плана ПФЭ является симметричной?

а) да;

б) нет;

в) это зависит от числа факторов.

17. Кто определяет структуру и параметры модели?

а) заказчик;

б) исследователь;

в) необходимы дополнительные данные.

18. Модели какого типа позволяют получить методы планирования эксперимента?

а) любые;

б) детерминированные;

в) вероятностные.

19. Количество управляющих факторов в модели равно двум. Сколько строк должна содержать матрица планирования?

а) 2;

б) 4;

в) 8.

20. Количество управляющих факторов в модели равно трем. Сколько строк должна содержать матрица планирования?

а) 2;

б) 4;

в) 8.

Список литературы

1. Сидняев Н. И. Теория планирования эксперимента и анализ статистических данных: учеб. пособие. М.: Юрайт, 2012. 399 с.
2. Адлер Ю. П., Маркова Е. В., Грановский Ю. В. Планирование эксперимента при поиске оптимальных условий. М.: Наука, 1976. 279 с.
3. Федоров В. В. Теория оптимального эксперимента. М.: Наука, 1971. 312 с.
4. Хартман К., Лецкий Э., Шефер В. Планирование эксперимента в исследовании технологических процессов. М.: Мир, 1977. 552 с.
5. Хикс Ч. Основные принципы планирования эксперимента. М.: Мир, 1967. 406 с.

Задания для самостоятельного решения

На основании полученных экспериментальных данных вывести математическую модель (линейную и нелинейную) в кодированных и натуральных значениях факторов.

Вариант 1–6. Для изучения зависимости некоторой величины от воздействующих факторов были поставлены эксперименты по плану ПФЭ 23. В качестве факторов, влияющих на отклик, были выбраны следующие:

Уровни факторов	Факторы процесса		
	X_1	X_2	X_3
Нижний	6	40	0,22
Основной	10	80	0,40
Верхний	14	120	0,31

Вариант 1

Y_1	Y_2	Y_3
0,12	1,10	0,11
0,06	0,06	0,06
0,20	0,18	0,22
0,18	0,16	0,18
0,14	0,12	0,14
0,11	0,12	0,10
0,24	0,23	0,24
0,20	0,22	0,20

Вариант 2

Y_1	Y_2	Y_3
0,12	1,12	0,11
0,06	0,07	0,06
0,20	0,18	0,21
0,18	0,16	0,18
0,13	0,12	0,14
0,11	0,12	0,11
0,24	0,23	0,22
0,20	0,21	0,20

Вариант 3

Y_1	Y_2	Y_3
0,12	0,11	0,12
0,06	0,06	0,08
0,20	0,21	0,20
0,18	0,18	0,17
0,12	0,14	0,16
0,12	0,11	0,10
0,23	0,22	0,21
0,21	0,20	0,18

Вариант 4

Y_1	Y_2	Y_3
0,12	1,10	1,10
0,06	0,06	0,06
0,20	0,18	0,18
0,18	0,16	0,16
0,14	0,12	0,12
0,11	0,12	0,12
0,24	0,23	0,23
0,20	0,22	0,22

Вариант 5

Y_1	Y_2	Y_3
0,12	1,12	0,12
0,06	0,07	0,08
0,20	0,18	0,20
0,18	0,16	0,17
0,13	0,12	0,13
0,11	0,12	0,11
0,24	0,23	0,24
0,20	0,21	0,20

Вариант 6

Y_1	Y_2	Y_3
0,12	0,11	0,12
0,18	0,06	0,08
0,14	0,21	0,20
0,11	0,18	0,18
0,12	0,14	0,14
0,12	0,11	0,11
0,23	0,22	0,21
0,21	0,20	0,18

Вариант 7–12. Для изучения зависимости некоторой величины от воздействующих факторов были поставлены эксперименты по плану ПФЭ 23. В качестве факторов, влияющих на отклик, были выбраны следующие:

Уровни факторов	Факторы процесса		
	X_1	X_2	X_3
Нижний	3	11	0,2
Основной	7	26	0,25
Верхний	11	41	0,3

Вариант 7

Y_1	Y_2	Y_3
0,12	1,10	0,11
0,06	0,06	0,06
0,20	0,18	0,22
0,18	0,16	0,18
0,14	0,12	0,14
0,11	0,12	0,10
0,24	0,23	0,24
0,20	0,22	0,20

Вариант 8

Y_1	Y_2	Y_3
0,12	1,12	0,11
0,06	0,07	0,06
0,20	0,18	0,21
0,18	0,16	0,18
0,13	0,12	0,14
0,11	0,12	0,11
0,24	0,23	0,22
0,20	0,21	0,20

Вариант 9

Y_1	Y_2	Y_3
0,12	0,11	0,12
0,06	0,06	0,08
0,20	0,21	0,20
0,18	0,18	0,17
0,12	0,14	0,16
0,12	0,11	0,10
0,23	0,22	0,21
0,21	0,20	0,18

Вариант 10

Y_1	Y_2	Y_3
0,12	1,10	1,10
0,06	0,06	0,06
0,20	0,18	0,18
0,18	0,16	0,16
0,14	0,12	0,12
0,11	0,12	0,12
0,24	0,23	0,23
0,20	0,22	0,22

Вариант 11

Y_1	Y_2	Y_3
0,12	1,12	0,12
0,06	0,07	0,08
0,20	0,18	0,20
0,18	0,16	0,17
0,13	0,12	0,13
0,11	0,12	0,11
0,24	0,23	0,24
0,20	0,21	0,20

Вариант 12

Y_1	Y_2	Y_3
0,12	0,11	0,12
0,10	0,06	0,08
0,21	0,21	0,20
0,16	0,18	0,18
0,12	0,14	0,14
0,12	0,11	0,11
0,23	0,22	0,21
0,21	0,20	0,18

Вариант 13–18. Для изучения зависимости некоторой величины от воздействующих факторов были поставлены эксперименты по плану ПФЭ 23. В качестве факторов, влияющих на отклик, были выбраны следующие:

Уровни факторов	Факторы процесса		
	X_1	X_2	X_3
Нижний	5	30	0,2
Основной	10	80	0,4
Верхний	15	130	0,6

Вариант 13

Y_1	Y_2	Y_3
0,12	1,10	0,11
0,06	0,06	0,06
0,20	0,18	0,22
0,18	0,16	0,18
0,14	0,12	0,14
0,10	0,12	0,10
0,24	0,23	0,24
0,20	0,22	0,20

Вариант 14

Y_1	Y_2	Y_3
0,12	1,12	0,11
0,08	0,07	0,06
0,20	0,18	0,21
0,18	0,17	0,18
0,13	0,12	0,14
0,11	0,12	0,11
0,22	0,23	0,22
0,20	0,21	0,20

Вариант 15

Y_1	Y_2	Y_3
0,12	0,11	0,12
0,06	0,06	0,08
0,20	0,21	0,20
0,18	0,19	0,18
0,12	0,13	0,16
0,12	0,10	0,11
0,23	0,21	0,21
0,21	0,24	0,18

Вариант 16

Y_1	Y_2	Y_3
0,12	1,10	1,13
0,06	0,06	0,06
0,20	0,18	0,17
0,18	0,16	0,19
0,14	0,12	0,15
0,11	0,13	0,12
0,24	0,25	0,23
0,20	0,22	0,19

Вариант 17

Y_1	Y_2	Y_3
0,12	1,12	0,14
0,06	0,07	0,06
0,20	0,18	0,19
0,18	0,16	0,16
0,13	0,12	0,15
0,11	0,12	0,11
0,24	0,23	0,22
0,20	0,21	0,19

Вариант 18

Y_1	Y_2	Y_3
0,12	0,11	0,10
0,09	0,06	0,08
0,19	0,21	0,20
0,17	0,20	0,18
0,12	0,13	0,14
0,12	0,13	0,11
0,23	0,22	0,21
0,21	0,20	0,18

Приложение 2

ТАБЛИЦА *G*-РАСПРЕДЕЛЕНИЯ

G-случайная величина, распределенная по закону Кохрена с числом степеней свободы f_1=n_1 для числителя и f_2=n_2 для знаменателя. Таблица содержит значения e,, полученные из условия $P(|G|<e)$=0,95 (верхняя строка при всех n_2) $P(|G|<e)$=0,99 (нижняя строка при тех же n_2).

n_2 \ n_1	1	2	3	4	5	6	7	8	10
2	0,998	0,97	0,94	0,91	0,86	0,85	0,83	0,82	0,79
	0,999	0,99	0,99	0,98	0,96	0,94	0,92	0,90	0,88
3	0,966	0,87	0,80	0,75	0,71	0,68	0,65	0,63	0,60
	0,991	0,94	0,88	0,83	0,79	0,76	0,71	0,69	0,67
4	0,91	0,77	0,68	0,63	0,59	0,56	0,54	0,52	0,49
	0,97	0,96	0,78	0,72	0,68	0,64	0,61	0,58	0,55
5	0,84	0,68	0,60	0,54	0,51	0,48	0,46	0,44	0,41
	0,93	0,79	0,70	0,63	0,59	0,55	0,52	0,50	0,47
6	0,78	0,62	0,53	0,48	0,45	0,42	0,40	0,38	0,36
	0,88	0,72	0,63	0,56	0,52	0,49	0,46	0,44	0,40
7	0,73	0,56	0,48	0,43	0,39	0,37	0,36	0,34	0,32
	0,84	0,66	0,57	0,51	0,47	0,44	0,46	0,39	0,36
8	0,68	0,52	0,44	0,39	0,36	0,34	0,32	0,30	0,28
	0,79	0,62	0,52	0,46	0,43	0,39	0,37	0,35	0,33
9	0,64	0,48	0,40	0,36	0,33	0,31	0,29	0,28	0,26
	0,75	0,57	0,48	0,43	0,39	0,36	0,34	0,32	0,30
10	0,60	0,45	0,37	0,33	0,30	0,28	0,27	0,25	0,24
	0,72	0,54	0,45	0,39	0,36	0,33	0,31	0,30	0,27
12	0,54	0,39	0,33	0,29	0,26	0,24	0,23	0,22	0,21
	0,65	0,48	0,39	0,34	0,31	0,29	0,27	0,25	0,23
15	0,47	0,33	0,28	0,24	0,22	0,20	0,19	0,18	0,17
	0,57	0,41	0,33	0,29	0,26	0,24	0,22	0,21	0,19
20	0,39	0,27	0,22	0,19	0,17	0,16	0,15	0,14	0,130
	0,48	0,33	0,27	0,23	0,20	0,19	0,17	0,16	0,150
24	0,34	0,24	0,19	0,17	0,15	0,14	0,13	0,116	0,111
	0,42	0,29	0,23	0,20	0,18	0,16	0,15	0,142	0,133
30	0,29	0,20	0,16	0,14	0,12	0,11	0,106	0,100	0,092
	0,36	0,24	0,19	0,16	0,15	0,13	0,123	0,116	0,105
40	0,24	0,16	0,13	0,11	0,10	0,089	0,083	0,078	0,071
	0,29	0,19	0,15	0,13	0,11	0,103	0,095	0,090	0,082
60	0,17	0,11	0,09	0,08	0,068	0,062	0,058	0,055	0,050
	0,22	0,14	0,11	0,09	0,080	0,072	0,067	0,063	0,057
120	0,10	0,06	0,05	0,042	0,037	0,034	0,030	0,029	0,027
	0,027	0,08	0,06	0,049	0,043	0,039	0,036	0,033	0,030

Примечание. Допускается линейная интерполяция по аргументу n_2. Погрешность интерполяции не превышает 0,01.

Приложение 3

ТАБЛИЦА *t*-РАСПРЕДЕЛЕНИЯ

t-случайная величина, распределенная по закону Стьюдента с числом степеней свободы $f=n$. Таблица содержит значения e, полученные из условия $P(|G|<e)=1-q$.

n \ $1-q$	0,99	0,95	0,90	0,80	0,50	0,20
1	63,657	12,706	6,314	3,078	0,727	0,325
2	9,935	4,303	2,920	1,886	0,617	0,289
3	5,841	3,182	2,353	1,638	0,584	0,277
4	4,604	2,776	2,132	1,533	0,569	0,271
5	4,032	2,571	2,015	1,476	0,559	0,267
6	3,707	2,447	1,943	1,440	0,553	0,265
7	3,499	2,365	1,895	1,415	0,549	0,263
8	3,355	2,306	1,860	1,397	0,546	0,262
9	3,250	2,262	1,833	1,383	0,543	0,261
10	3,169	2,228	1,812	1,372	0,542	0,260
11	3,106	2,201	1,796	1,363	0,540	0,260
12	3,055	2,119	1,782	1,356	0,539	0,259
13	3,012	2,160	1,771	1,350	0,538	0,259
14	2,977	2,145	1,761	1,345	0,537	0,258
15	2,947	2,131	1,753	1,341	0,536	0,258
16	2,921	2,120	1,746	1,337	0,535	0,258
18	2,878	2,101	1,734	1,330	0,534	0,257
20	2,845	2,086	1,725	1,325	0,533	0,257
23	2,807	2,069	1,714	1,319	0,532	0,256
25	2,787	2,060	1,708	1,316	0,531	0,256
30	2,750	2,042	1,697	1,310	0,530	0,256
40	2,704	2,021	1,684	1,303	0,529	0,255
60	2,660	2,000	1,671	1,296	0,527	0,254
100	2,617	1,980	1,658	1,289	0,526	0,254
	2,576	1,960	1,645	1,282	0,524	0,253

Примечание. Допускается интерполяция только по аргументу n. Погрешность линейной интерполяции не превышает 0,007.

ТАБЛИЦА *F*-РАСПРЕДЕЛЕНИЯ

F – случайная величина, распределенная по закону Фишера с числом степеней свободы $f_1=n_1$ для числителя и $f_2=n_2$ для знаменателя. Таблица содержит значения e, полученные из условия $P(|F|<e)=0{,}95$ (верхняя строка при всех n_2) $P(|F|<e)=0{,}99$ (нижняя строка при тех же n_2).

	f_1										
f_2	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**15**
1	161,45	199,50	2115,71	224,58	230,16	233,99	236,77	238,88	240,54	241,88	245,95
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38	19,40	19,43
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,70
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,86
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,62
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	3,94
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	368	3,64	3,51
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,22
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,01
10	4,96	4.10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,85
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,72
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,62
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	2,53
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,46
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,40
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,35
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	2,31
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,27
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,23
20	4,35	3,49	3,10	2,87	2,21	2,60	2,51	2,45	2,39	2,35	2,20

Примечание. Допускается линейная интерполяция по аргументу n_2 и квадратичная по n_1. Погрешность интерполяции не превышает 0,01.