ВОЕННО-КОСМИЧЕСКАЯ АКАДЕМИЯ
имени А.Ф. Можайского

МОДЕЛИРОВАНИЕ СИСТЕМ УПРАВЛЕНИЯ

Практикум

Учебное издание «Моделирование систем управления» утверждено в качестве практикума по дисциплине «Моделирование систем управления» и рекомендовано кафедрой автоматизированных систем подготовки и пуска ракет космического назначения Военно-космической академии имени А.Ф. Можайского для обучающихся по основной профессиональной образовательной программе высшего образования – программе специалитета по специальности «Специальные организационно-технические системы», протокол от 26 июня 2017 г. № 20

Санкт-Петербург
2017

Рецензент:
доктор технических наук, старший научный сотрудник **И.С. Гурьев**

Моделирование систем управления: практикум / Н.А. Осипов, И.В. Дорожко, А.С. Шавин, К.А. Крупский, С.В. Калиниченко. – СПб.: ВКА имени А.Ф. Можайского, 2017. – 80 с.

Практикум содержит описание девяти работ, каждая из которых включает краткие теоретические сведения, методические рекомендации и порядок выполнения.

Практикум разработан авторским коллективом в составе: кандидата технических наук, доцента Н.А. Осипова – руководитель авторского коллектива (введение, работы 2–6), кандидата технических наук И.В. Дорожко (работа 9), кандидата технических наук А.С. Шавина (работа 8), кандидата технических наук К.А. Крупского (работа 7), кандидата технических наук С.В. Калиниченко (работа 1).

Подписано к печ. 18.07.2017	Формат печатного листа 445×300/8
Гарнитура Times New Roman	Уч.-изд. л. 5,00
Уч.-печ. л. 11,00	Заказ 3431/43 Бесплатно

Типография ВКА имени А.Ф. Можайского

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ

Трудно себе представить современную науку без применения математического моделирования. Сущность этой методологии состоит в замене исходного объекта его «образом» – математической моделью, и дальнейшем изучении модели с помощью реализуемых на компьютерах вычислительно-логических алгоритмов. Этот метод познания сочетает в себе многие достоинства, как теории, так и эксперимента. Работа не с самим объектом (явлением, процессом), а с его моделью дает возможность безболезненно, относительно быстро и без существенных затрат исследовать его свойства и поведение в любых ситуациях. Вычислительные эксперименты с моделями объектов позволяют, опираясь на мощь современных вычислительных методов и технических средств, подробно и глубоко изучать объекты в достаточной полноте, недоступной чисто теоретическим подходам.

В настоящее время методология математического моделирования бурно развивается, охватывая все новые сферы – от разработки технических систем и управления ими до анализа сложнейших экономических и социальных процессов.

Данное учебное пособие содержит описание практических занятий и лабораторных работ по дисциплине «Моделирование систем управления».

Первая работа посвящена исследованию свойств датчика случайных чисел: равномерности распределения и независимости чисел в псевдослучайной последовательности. Работа выполняется с использованием специально созданной на кафедре программы, генерирующей различные псевдослучайные последовательности. Планируемое время выполнения работы 4 часа.

Следующие шесть работ заключаются в изучении средств широко известного программного продукта Microsoft Excel, применяемых для решения основных задач моделирования:

– расчет основных характеристик марковских процессов;

– построение линейных оптимизационных моделей;

– изучение и исследование законов распределения случайных величин;

– генерация случайных чисел с требуемыми законами распределения для построения и анализа выборок данных;

– анализ временных рядов;

– регрессионный анализ данных.

На выполнение каждой работы отводится 4 часа.

Восьмая работа посвящена изучению метода наименьших квадратов, применяемого для построения математических моделей по экспериментальным данным. В работе рассматривается способ построения моделей при помощи базисных функций. Планируемое время выполнения работы 6 часов.

Целью девятой работы является идентификация динамических систем управления в интерактивных инструментах Matlab. Планируемое время выполнения работы 6 часов.

В написании руководства принимали участие: кандидат технических наук, доцент Н.А. Осипов – руководитель авторского коллектива (введение, работы 2–6), кандидат технических наук И.В. Дорожко (работа 9), кандидат технических наук А.С. Шавин (работа 8), кандидат технических наук К.А. Крупский (работа 7), кандидат технических наук С.В. Калиниченко (работа 1).

1. ИССЛЕДОВАНИЕ ДАТЧИКА СЛУЧАЙНЫХ ЧИСЕЛ

Цель работы

Исследование равномерности распределения и независимости чисел в псевдослучайной последовательности.

Теоретические основы

Моделирование случайностей основывается на использовании датчика случайных чисел (ДСЧ), который вырабатывает жёсткую периодическую последовательность чисел, находящихся в пределах от 0 до 1. Период последовательности достаточно большой, и в зависимости от программной среды составляет миллионы и миллиарды чисел, что, как правило, достаточно для решения задач имитационного моделирования. Для того чтобы каждый раз при запуске программы не генерировалась одна и та же последовательность чисел, в программу включают команду рандомизации (Randomize). Эта команда позволяет входить в псевдослучайную последовательность достаточно случайно, так как место входа зависит от текущего времени, задаваемого до наносекунд.

Требуемые в программе законы распределения вероятностей строятся на базе ДСЧ в предположении, что он обеспечивает равномерное распределение и что числа в вырабатываемой последовательности независимы. Для уверенности в результатах моделирования полезно знать, насколько эти предположения справедливы.

Гипотезу о равномерности распределения можно проверить следующим образом.

Надо разбить диапазон вырабатываемых ДСЧ чисел на m одинаковых интервалов. При числе обращений к ДСЧ (объёме выборки n), в каждый интервал должно попасть n/m чисел, т.е. оценка вероятности $\hat{p}$ (опытное значение вероятности) попадания в любой интервал равняется:

$$\hat{p} = \frac{n}{m \cdot n} = \frac{1}{m}.$$

При ограниченном значении n величина $\hat{p}$ случайна, причём, из математической статистики известно, что среднеквадратическое отклонение этой оценки определяется выражением:

$$\sigma[\hat{p}] = \sqrt{\frac{p(1-p)}{n}},$$

где p – истинное значение оцениваемой вероятности, которое при равномерном распределении равно $1/m$.

Таким образом, если при некотором значении n отклонение $\hat{p}$ от $1/m$ не выходит за пределы 3σ, т.е. если

$$\left|\hat{p}-1/m\right|<3\sqrt{\frac{1/m(1-1/m)}{n}}$$

для всех интервалов, то нет оснований гипотезу о равномерности отвергнуть.

Эту проверку необходимо выполнить для различных значений n.

Независимость чисел в последовательности можно проверить так же, как проверяется независимость случайных величин (x, y) – путём определения ковариации (момент связи, корреляционный момент), равной матожиданию произведения центрированных случайных величин:

$$K_{xy}=\mathrm{cov}(x,y)=M[(x-M[x])(y-M[y])].$$

После преобразований можно получить:

$$K_{xy}=M[xy]-M[x]M[y].$$

Ковариация может находиться в пределах от $-\sigma[x]\cdot\sigma[y]$ до $+\sigma[x]\cdot\sigma[y]$, принимая нулевое значение для независимых величин.

Для псевдослучайной последовательности подобной характеристикой служит автокорреляционная функция (или просто корреляционная функция), которая отражает корреляцию между элементами последовательности, расположенными на расстоянии s (s = 0,1,2,...).

Корреляционная функция $K(s)$ идеальной псевдослучайной последовательности должна быть равна:

$$K(s)=\begin{cases}D[r], & s=0\\ 0, & s>0\end{cases}$$

где r – случайная величина, значениями которой являются числа, генерируемые ДСЧ;

$D[r]$ – дисперсия.

Для случайной величины распределённой равномерно в диапазоне Δ дисперсия равна $\Delta^2/12$.

Оценку корреляционной функции можно вычислить по формуле:

$$\hat{K}(s)=\frac{1}{n-s}\sum_{i=1}^{n-s}r_i r_{i+s}-\left(\frac{1}{n-s}\sum_{i=1}^{n-s}r_i\right)\left(\frac{1}{n-s}\sum_{i=s+1}^{n}r_i\right).$$

Программа, для исследования равномерности распределения чисел в псевдослучайной последовательности и их независимости строит гистограмму (при фиксированном числе интервалов $m = 10$) и вычисляет оценку корреляционной функции при заданном s (рис. 1.1).

Объём выборки можно увеличивать постепенно, добавляя к уже выбранным числам ещё произвольное количество чисел. Оценку корреляционной функции можно получить при произвольном объёме выборки n в зависимости от числа s путём нажатия кнопки $K(s)$.

Гистограмма является графическим отображением оценок вероятностей попадания в интервалы. Весь диапазон чисел равен 10, что достигается умножением результата обращения к ДСЧ на 10. Интервалы обозначены «0», «1», «2», ..., «9». При прибавлении к объёму выборки (с помощью кнопки **Add**) по одному числу по гистограмме можно проследить, в какой из интервалов попало очередное число. Сами числа последовательности отображаются в специальном окне, в котором помещается 32000 чисел.

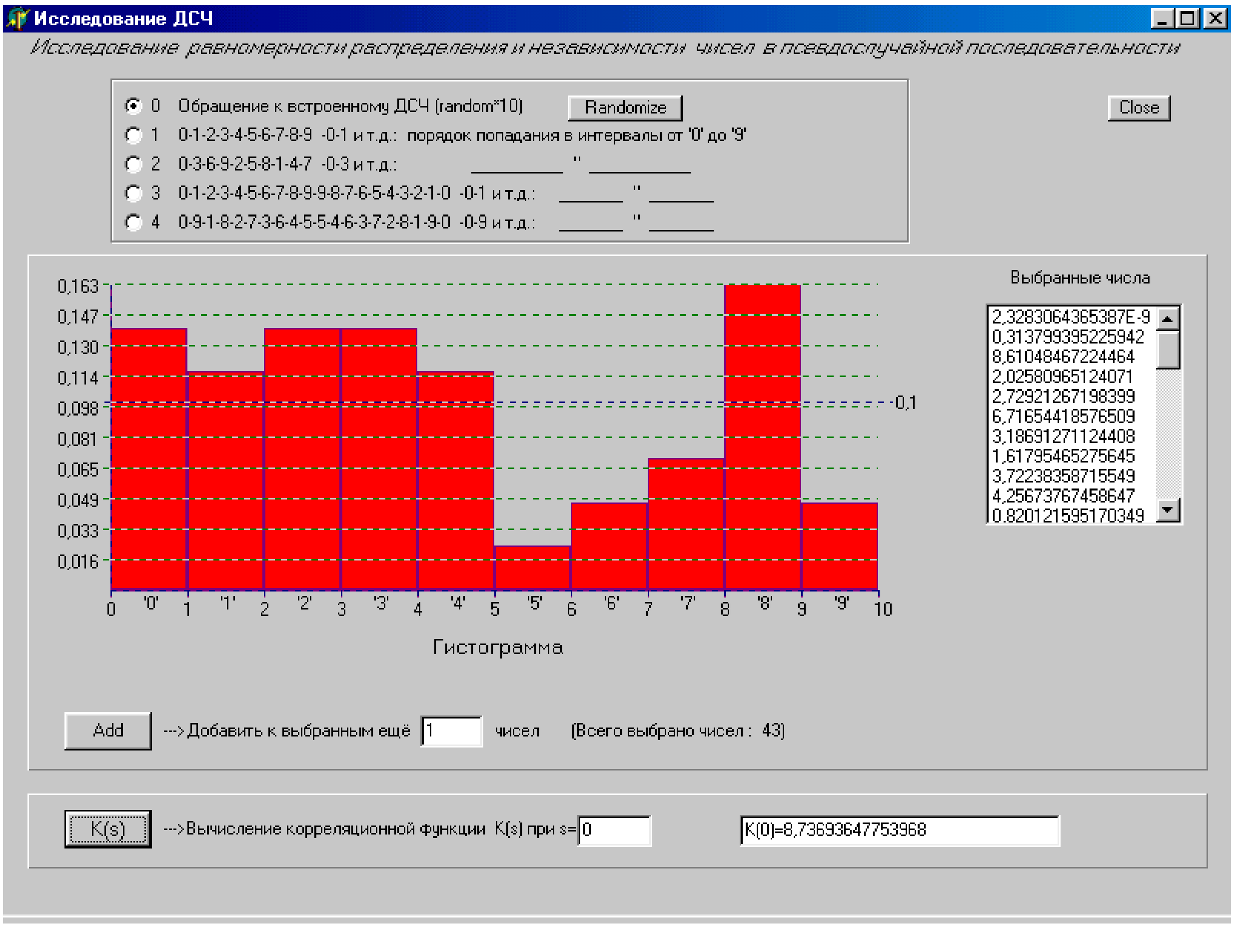


Рис. 1.1. Окно программы

Для сравнения значений корреляционной функции псевдослучайной последовательности со значениями корреляционной функции других возможных последовательностей, отличающихся от основной наличием явных закономерностей, в программе предусмотрена генерация последовательностей, содержащих различные закономерности. Выбираются они с помощью переключателей (радиокнопок). Рядом с кнопками указан характер закономерностей, отражающий порядок попадания чисел в интервалы. Закономерности распространяются на целые числа: 0, 1, …, 9. К ним прибавляется случайная величина, находящаяся в пределах от 0 до 1.

Для рандомизации псевдослучайной последовательности применяется кнопка **Randomize**.

Содержание работы

Основная часть работы состоит в проверке гипотез о равномерности распределения и независимости чисел в псевдослучайной последовательности.

Порядок выполнения

1. Запустить программу несколько раз, записывая в каждом случае первые 3 числа, генерируемых ДСЧ. Убедиться, что они одни и те же.

2. Повторить предыдущий пункт, но в отличие от него, до первого нажатия кнопки **Add** с помощью кнопки **Randomize** рандомизировать последовательность. Начало последовательности должно измениться. Записать начальные значения.

3. Проверить гипотезу о равномерности распределения чисел в псевдослучайной последовательности:

3.1. Построить графики доверительной границы равной

$$\sqrt[3]{\frac{1/m\cdot(1-1/m)}{n}}=\frac{0.9}{\sqrt{n}}$$

для принятия гипотезы о равномерности, в пределах изменения n от 1 до 100 и от 100 до 10000 (2 графика), откладывая по оси абсцисс $\sqrt{n}$.

3.2. Нанести на графики точки максимального отклонения оценок вероятностей попадания в интервалы от предполагаемого истинного значения вероятности, равного $1/m = 0{,}1$ для n = 1,4,9,16,25,36,49,64,81,100 на первый график и n = 400, 900, 1600, 2500, 3600, 4900, 6400, 8100, 10000 на второй график. Сделать выводы о справедливости гипотезы.

3.3. Нанести на эти же графики точки максимального отклонения оценок вероятностей при некоторых значениях n для последовательностей с явными закономерностями.

4. Проверить независимость чисел в псевдослучайной последовательности, для чего построить графики корреляционных функций $K(s)$ при изменении s от 0 до 30 для всех реализуемых в программе последовательностей. Объём выборки при определении $K(s)$ должен быть не менее 10000. Сравнить функции и сделать выводы.

Отчётность по работе

После выполнения работы обучаемый представляет отчет. Отчёт должен содержать:

1. Название и цель работы.
2. Начальные числа, вырабатываемые ДСЧ, без рандомизации и с рандомизацией по п.п. 1 и 2.
3. Графики доверительных границ с нанесёнными точками максимальных отклонений оценок вероятностей от теоретического значения, равного 0,1 (п. 3).
4. Графики корреляционных функций (п. 4).
5. Выводы по всем пунктам задания.

Контрольные вопросы

1. Как формируются числа псевдослучайной последовательности?
2. Как осуществляется рандомизация?
3. Как формируются числа последовательностей с явными закономерностями, используемых в программе?
4. Что собой представляет гистограмма?
5. Как определяется среднеквадратическое отклонение оценки вероятности?
6. Как проверяется гипотеза о равномерности распределения чисел в псевдослучайной последовательности?
7. Что такое ковариация между случайными величинами и что она характеризует?
8. Что такое ковариационная функция числовой последовательности?
9. Чему должна равняться ковариационная функция для псевдослучайной последовательности?
10. Как определяется оценка ковариационной функции?

2. РАСЧЕТ ХАРАКТЕРИСТИК МАРКОВСКИХ ПРОЦЕССОВ

Цель работы

Изучить средства программы Microsoft Excel для расчета основных характеристик марковских процессов.

Теоретические основы

В настоящее время имеется большое количество компьютерных программ для использования различных итерационных методов. Наиболее простые и удобные программы оформлены в виде функций MS Excel **Подбор параметра** и **Поиск решения**.

Перед тем как рассмотреть основные средства MS Excel полезно изучить возможности программы для решения менее сложных задач, например, задачи исследования функций.

Многие инженерные задачи сводятся к исследованию функций одной или нескольких переменных вида $\boldsymbol{Y = f(X)}$ или $\boldsymbol{Y = f(x_1, x_2, \ldots, x_N)}$.

Исследовать функцию, значит установить область ее существования (значения $\boldsymbol{X}$ при которых возможно вычислить $\boldsymbol{Y}$), определить области значений $\boldsymbol{X}$, при которых $\boldsymbol{Y}$ принимает положительные, отрицательные и аномально большие значения («уходит в бесконечность»), найти максимумы, минимумы, иногда точки перегиба графика функции, а также корни уравнения $\boldsymbol{Y = f(x)}$ – значения $\boldsymbol{x}$, при которых $\boldsymbol{Y}$ обращается в 0 (график функции пересекает ось абсцисс).

Наиболее простые методы исследования функциональных зависимостей с помощью компьютера – итерационные**,** основанные на многократном выполнении сравнительно простых операций.

Один из итерационных методов – табулирование функции, включающий следующие этапы:

– расчет значений $\boldsymbol{Y}$ при заданных $\boldsymbol{X}$ в большом диапазоне значений $\boldsymbol{X}$ с большим шагом;

– табулирование с небольшим шагом в наиболее близких диапазонах – вблизи корней, максимумов и минимумов;

– сужение диапазонов $\boldsymbol{X}$ и уменьшение шага для получения все более точных значений экстремумов и корней. Получаемые решения зависят от того, в каких диапазонах $\boldsymbol{X}$ и $\boldsymbol{Y}$ ведется их поиск, т.е. от их начальных значений.

Пример 1. Решение уравнений

Пример показывает использование функции **Подбор параметра** для решения уравнений.

Требуется: решить уравнение $Y = 0{,}1 \cdot x^2 - x - 11$.

Решение.

1. С помощью график функции определить количество действительных корней уравнения. Для этого следует:

1.1. Задать область определения (x) от –20 до +20: занести в соседние ячейки (например, A5 и A6) –20 и –19, выделить обе ячейки, поставить курсор на черный квадратик в правом нижнем углу, нажать левую клавишу мыши и потащить вниз до появления числа 20.

1.2. Создать область значений: в ячейку рядом с –20 вставить формулу = 0,1*A5^2-A5-11, скопировать ее вниз.

1.3. Выделить область значений и вызвать в меню **Вставка** команду **Диаграмма…** (Excel 2003) или на вкладке **Вставка** в группе **Диаграммы** выбрать кнопку **График** (Excel 2007). Обратите внимание, что на оси ***X*** указываются не значения аргумента, а порядковые номера.

2. Найти первый корень уравнения:

2.1. Сделать активной ячейку в диапазоне ***Y*** вблизи одного из корней (первое пересечение графика функции с осью абсцисс).

2.2. Вызвать **Подбор параметра:**

2.2.1. Excel 2003: в меню **Сервис**, выбрать команду **Подбор параметра**;

2.2.2. Excel 2007: на вкладке **Данные**, в группе **Работа с данными** выбрать кнопку **Анализ «что-если»** и указать команду **Подбор параметра**.

2.3. В окне *Значение* установить 0, в нижнем окне *Изменяя значение ячейки* указать адрес ячейки ***X***, соответствующей активной ячейки ***Y*,** после чего кликнуть по кнопке **ОК**.

3. Найти второй корень, выбрав значения *Y* и *X* вблизи него.

Пример 2. Решение системы уравнений

Расчет установившихся значений вероятностей состояний системы (финальных вероятностей), описанной марковским процессом, сводится к решению системы алгебраических уравнений. Для изучения принципа решения подобных систем рассмотрим решение системы из трех уравнений с тремя неизвестными вида $a_iX + b_iY + c_iZ = d_i$ (i = 1,2,3) с помощью команды **Поиск решения**.

Поиск решения – это надстройка, входящая в поставку Excel и предназначенная для решения задач линейной и нелинейной оптимизации. Для этого в ней используются методы и алгоритмы математического программирования, которые позволяют находить оптимальные решения задач оптимизации, представленные в виде табличных моделей. Пример подобной задачи рассматривается в следующей работе. В данном примере с помощью этой надстройки решается система уравнений.

Для решения системы уравнений следует:

1. Выполнить подготовительные действия (табл. 2.1):

1.1. Записать строкой начальные значения неизвестных *X, Y, Z*, например, нулевые значения в ячейки A10, B10, C10 соответственно.

1.2. Составить таблицу, содержащую значения коэффициентов при этих неизвестных a_i, b_i, c_i (i = 1,2,3), и значения соответствующих свободных членов d_i, например, в ячейках A12 : D14.

1.3. Перемножить начальные значения *X, Y, Z* на соответствующие коэффициенты (ячейки A16:C18) и просуммировать произведения по строкам (ячейки D16:D18).

2. Запустить **Поиск решения**:

2.1. Excel 2003: в меню **Сервис**, выбрать команду **Поиск решения**.

2.2. Excel 2007: на вкладке **Данные**, в группе **Анализ** выбрать команду **Поиск решения.**

3. Заполнить окно диалога **Поиск решения** данными решаемой задачи (рис. 2.1):

3.1. В качестве целевой ячейки установить первую сумму: задать *«Установить целевую ячейку* (в примере это ячейка D16) *равной значению»* первого свободного члена d_1.

3.2. На две другие суммы наложить ограничения: равенство двум другим свободным членам d_2 и d_3 (используя кнопку *«Добавить»*).

3.3. В окне *«Изменяя ячейки»* указать ячейки с начальными значениями неизвестных *X, Y, Z* (в этих ячейках окажется результат решения системы).

3.4. Нажать кнопку *«Параметры»* и ознакомиться с параметрами и методами, используемыми при оптимизационных расчетах (при необходимости можно параметры изменить), чтобы закрыть окно *«Параметры»*, следует нажать кнопку ОК.

4. Запустить выполнение программы нажатием кнопки *«Выполнить»*. В результате появится сообщение о нахождении или не нахождении решения.

Таблица 2.1

	A	B	C	D	*Комментарии*
10	X	Y	Z		Начальные значения (числа), после решения задачи на их месте появятся результирующие значения
11					
12	a_1	b_1	c_1	d_1	Коэффициенты (числа) в уравнениях и значения свободных членов
13	a_2	b_2	c_2	d_2	
14	a_3	b_3	c_3	d_3	
15					
16	$a_1 \cdot X$	$b_1 \cdot Y$	$c_1 \cdot Z$	=Сумм(A16:C16)	Начальный вид системы уравнений (числа)
17	$a_2 \cdot X$	$b_2 \cdot Y$	$c_2 \cdot Z$	=Сумм(A17:C17)	
18	$a_3 \cdot X$	$b_3 \cdot Y$	$c_3 \cdot Z$	=Сумм(A18:C18)	

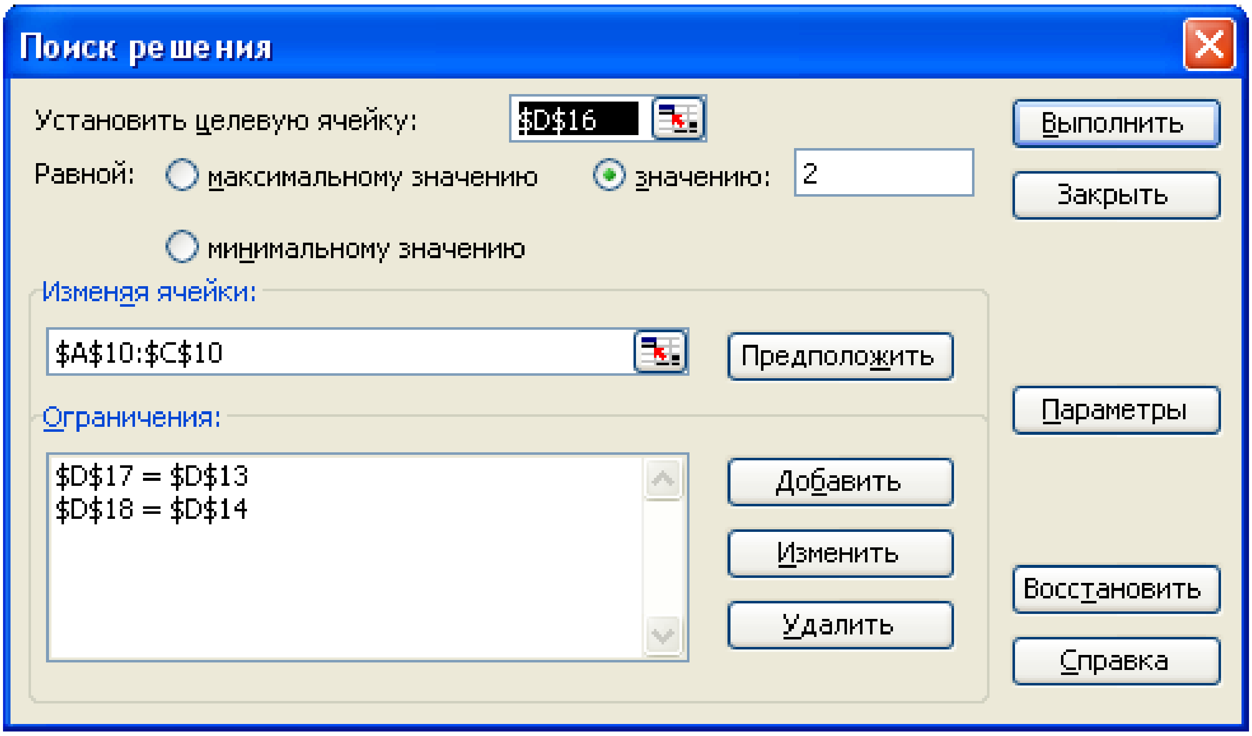


Рис. 2.1. Окно «Поиск решения»

Содержание работы

Основная часть работы состоит в расчете установившихся значений вероятностей состояний системы (финальных вероятностей), описанной марковским процессом.

Порядок выполнения

1. Изучить и решить примеры, описанные в теоретической части работы.

2. Выполнить задания для самостоятельного решения:

2.1. Найти корни уравнения третьего порядка. Рекомендации к решению: протабулируйте функцию (согласно номеру варианта) на достаточно большом интервале, постройте график, определите, сколько корней и где они примерно находятся, найдите корни через **Подбор параметра**.

Варианты заданий:

1. $Y=x^3-4x^2-5x+6=0$	5. $Y=x^3-5x^2-4x+8=0$
2. $Y=1{,}5x^3-x^2-4x+4=0$	6. $Y=1{,}5x^3-3x^2-6x+5=0$
3. $Y=1{,}2x^3-2x^2-x+4=0$	7. $Y=0{,}5x^3-2x^2-4x+6=0$
4. $Y=x^3-3x^2-4x+4=0$	8. $Y=1{,}1x^3-5x^2-3x+7=0$

2.2. Решить систему уравнений, используя надстройку **Поиск решения**.

Варианты заданий:

1) $x_1 + 2x_2 – x_3 = -15$
$- x_1 + 7 x_2 – 9 x_3 = 4$
$x_1 + 2 x_2 + 4 x_3 = 18$

2) $2 x_1 + x_2 – 3 x_3 = 5$
$-2 x_1 + 4 x_2 – 2 x_3 = 2$
$x_1 + x_2 + 4 x_3 = 8$

3) $x_1 + x_2 – x_3 = 2$
$x_1 + 3 x_2 – 5 x_3 = 5$
$x_1 + 4 x_2 - 3 x_3 = 10$

2.3. Определить финальные вероятности состояний процесса, заданного (по указанию преподавателя) в виде графа.

Отчётность по работе

После выполнения работы обучаемый представляет отчет. Отчёт должен содержать:

1. Название и цель работы.
2. Результаты решений примеров, описанных в теоретической части.
3. Результаты решения контрольных заданий.
4. Выводы по результатам работы.

Контрольные вопросы

1. Объясните назначение и принцип работы средства MS Excel **Подбор параметра.**
2. Какой процесс можно считать марковским?
3. Как можно определить установившиеся значения вероятностей состояний системы?
4. Сформулируйте правило, по которому составляется система уравнений для определения вероятностей состояний системы.

3. РЕШЕНИЕ ЗАДАЧ ЛИНЕЙНОЙ ОПТИМИЗАЦИИ

Цель работы

Изучить средства программы Microsoft Excel для решения задач линейной оптимизации.

Теоретические основы

В общем виде задачу оптимизации формулируют следующим образом.

Пусть $X = (x_1, x_2 \ldots x_n)$ – вектор действительных переменных. Необходимо минимизировать или максимизировать целевую функцию $z = f(X)$ при выполнении нескольких ограничений $g_j(X) \le b_j$, $(j = 1 \ldots m)$, которые задаются в виде неравенств или равенств. Могут быть также добавлены условия неотрицательности переменных ($x_i \ge 0$), которые включаются в указанные ограничения.

Если все функции $f(X)$ и $g_j(X)$ линейны относительно переменных x_i, то имеем задачу линейной оптимизации, если хотя бы одна из функций нелинейная, то получаем задачу нелинейной оптимизации.

Таким образом, задача оптимизации включает три элемента:

– переменные $x_1, x_2 \ldots x_n$ (в средстве **Поиск решения** ячейки, содержащие значения этих переменных, называются изменяемыми ячейками);

– целевая функция (ячейка, содержащая значение этой функции называется целевой ячейкой);

– ограничения (для применения средства **Поиск решения** ограничения могут быть записаны на рабочем листе и затем указаны в диалоговом окне либо заданы непосредственно в этом окне без записи на рабочем листе). При задании ограничений отдельно указываются функции ограничений $g_j(X)$ и вектор правых частей ограничений b_j.

После формулирования математической задачи оптимизации на рабочем листе Excel создается ее табличная модель, в которой в отдельных ячейках содержатся переменные решения, в отдельные ячейки записываются формулы, по которым будут вычисляться целевая функция и функции ограничений (левые части ограничений), также в отдельных ячейках указываются значения правых частей ограничений.

После создания табличной модели задачи оптимизации для нахождения оптимального решения применяют средство **Поиск решения**.

Пример. Минимизирование линейной функции

Дана функция: $z = 2x_1 + 3x_2 + 5x_3$.

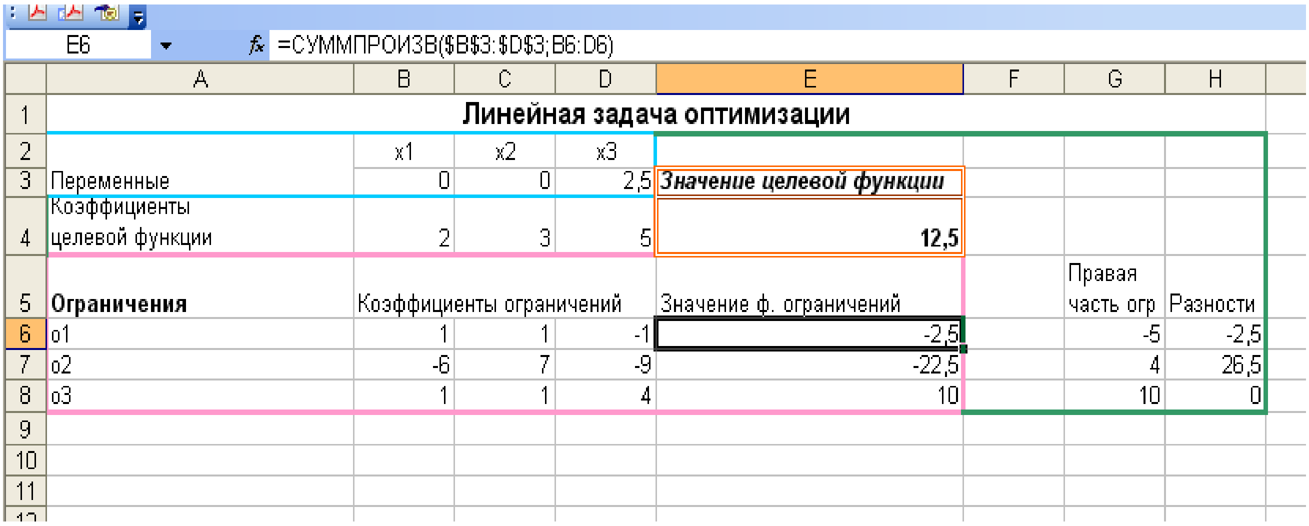
Требуется минимизировать эту функцию при следующих ограничениях:

$$x_1 + x_2 - x_3 \geq -5,$$
$$-6x_1 + 7x_2 - 9x_3 \leq 4,$$
$$x_1 + x_2 + 4x_3 = 10,$$

и на переменные поставлены условия неотрицательности.

Для решения данной задачи средство **Поиск решения** используется следующим образом:

1. Создать на рабочем листе Excel табличную модель решаемой задачи, например, как на рис. 3.1:



E6 =СУММПРОИЗВ($B$3:$D$3;B6:D6)

	A	B	C	D	E	F	G	H
1					Линейная задача оптимизации			
2		x1	x2	x3				
3	Переменные	0	0	2,5	*Значение целевой функции*			
4	Коэффициенты целевой функции	2	3	5	**12,5**			
5	**Ограничения**	Коэффициенты ограничений			Значение ф. ограничений		Правая часть огр	Разности
6	о1	1	1	-1	-2,5		-5	-2,5
7	о2	-6	7	-9	-22,5		4	26,5
8	о3	1	1	4	10		10	0
9								
10								
11								

Рис. 3.1. Табличная модель задачи линейной оптимизации

2. Открыть **Поиск решения**.

3. В открывшемся окне диалога **Поиск решения** указать данные, требуемые для процесса оптимизации (рис. 3.2):

– в поле *«Установить целевую ячейку»* ввести адрес ячейки, содержащей значение целевой функции. Для примера в это поле следует ввести E4 или щелкнуть указателем мыши по этой ячейке и адрес введется автоматически;

– в области *«Равной»* выбрать переключатель *минимальному значению*;

– в поле *«Изменяя ячейки»* указать ячейки, в которых содержатся переменные модели (в данном случае это диапазон B3:D3.

4. Задать ограничения – щелкнуть по кнопке *«Добавить»*, откроется окно диалога **Добавление ограничения**. Ввести ограничения поочередно, нажимая кнопку *«Добавить»* для каждого ограничения. После ввода всех ограничений нажать *ОК* для возврата в диалоговое окно **Поиск решения**.

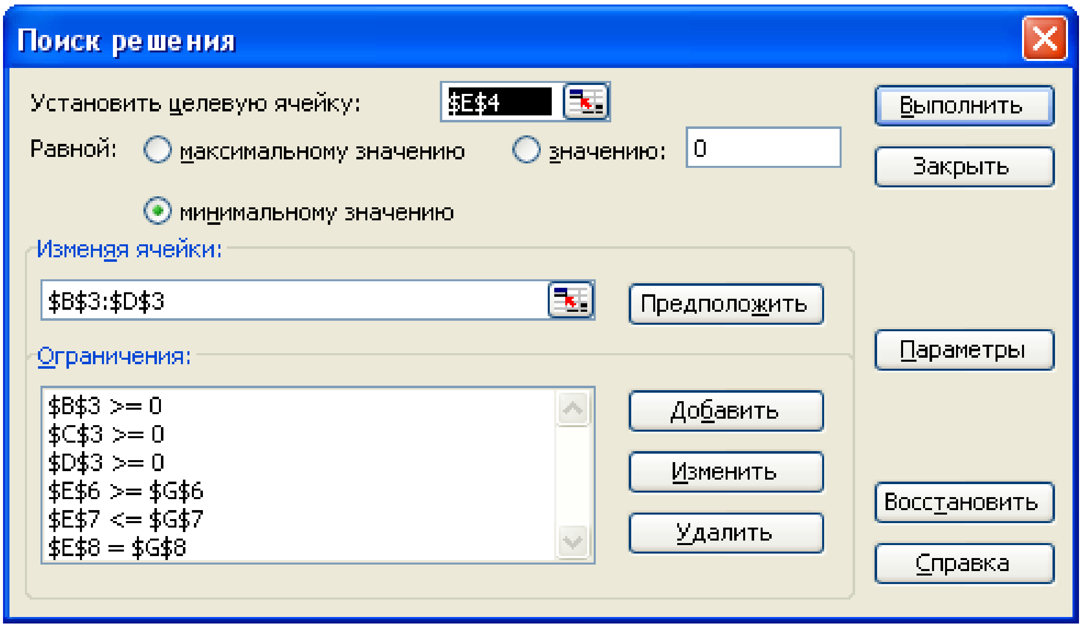


Рис. 3.2. Окно Поиск решения для задачи оптимизации

5. Нажать кнопку *«Параметры»*, откроется окно диалога **Параметры поиска решения**, предназначенное для задания дополнительных условий для поиска решений. Для рассматриваемого примера установить флажок *«Линейная модель»*, остальные параметры оставить без изменений. Нажать *ОК* для возврата в диалоговое окно **Поиск решения.**

6. После задания всех необходимых данных нажать кнопку *«Выполнить»*. Средство **Поиск решения** выполнит оптимизацию. В процессе вычислений в строке состояния отображаются число итераций и значения целевой функции при переборе множества допустимых решений задачи.

7. Если в табличной модели нет ошибок, то будет выведено на экран диалоговое окно **Результаты поиска решения** (рис. 3.3), которое сообщает о завершении поиска. При нахождении оптимального решения в диалоговом окне должны присутствовать два ключевых предложения: *Решение найдено. Все ограничения и условия оптимальности выполнены.* Если хотя бы одного из предложений нет, программе не удалось оптимизировать модель. В этом случае следует сначала проверить правильность ввода данных, затем табличную модель и, наконец, пересмотреть исходную формулировку задачи.

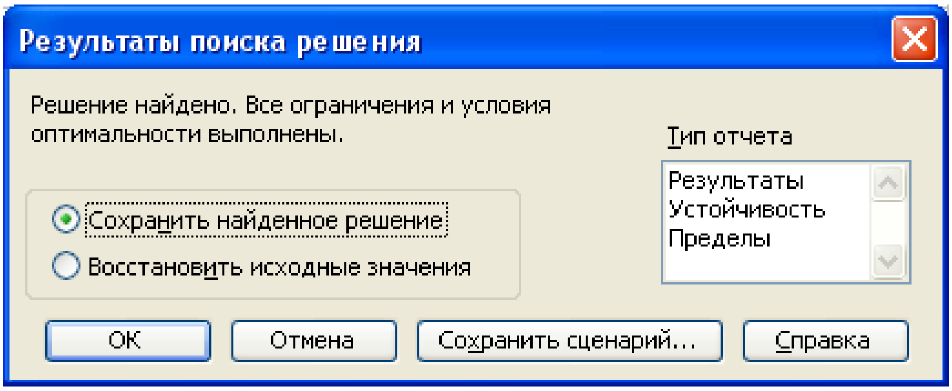


Рис. 3.3. Окно результатов поиска решения

8. В случае оптимального решения выбрать переключатель *Сохранить найденное решение* и нажать *ОК*. Кроме того есть возможность выбрать *Восстановить исходные значения*, что приведет к отказу от решения и восстановлению исходных значений в изменяемых ячейках, а также получить три типа отчетов (выбрав тип в списке *Тип отчета*) с результатами поиска решения.

Типы отчетов:

1. В отчете *Результаты* выводятся исходные и полученные в результате поиска решения значения изменяемых ячеек и целевой функции, а также сведения об ограничениях задачи.
2. Отчет *Устойчивость* дает основную информацию для анализа чувствительности моделей. Этот анализ показывает, насколько чувствительно найденное оптимальное решение к небольшим изменениям параметров модели.
3. Отчет *Пределы* показывает наименьшее и наибольшее значения, которые может принимать каждая переменная решения при удовлетворении ограничений и при постоянстве значений всех остальных переменных.

Содержание работы

Основная часть работы состоит в решении наиболее распространенных практических задач оптимизационных моделей.

Порядок выполнения

1. Изучить и решить пример, описанные в теоретической части работы.
2. Выполнить задания для самостоятельного решения:

2.1. Решить с помощью надстройки **Поиск решения** задачи «Распределения ресурсов» и «Рационального раскроя», описанные в курсе лекций[1].

2.2.Получить и проанализировать отчеты всех трех типов, описанные в теоретической части.

Отчётность по работе

После выполнения работы обучаемый представляет отчет. Отчёт должен содержать:

1. Название и цель работы.
2. Результаты решения примера, описанного в теоретической части.
3. Результаты решения контрольных заданий.
4. Выводы по результатам работы.

Контрольные вопросы

1. Объясните назначение и принцип работы средства MS Excel **Поиск решения.**
2. Какие элементы включает в себя задачу оптимизации?
3. Что такое целевая функция?
4. В каком случае модель является оптимизационной?
5. Из каких соображений формулируются ограничения?

[1] Н.А. Осипов, А.Б. Кузнецов. Моделирование систем. Часть 1. Математические модели: курс лекций. – СПб.:ВКА, 2007. – С.16–17, 20–21.

4. ИЗУЧЕНИЕ ЗАКОНОВ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

Цель работы

Изучить законы распределения случайных величин с помощью программы Microsoft Excel.

Теоретические основы

Биномиальное распределение

Случайная величина ξ имеет биномиальное распределение с параметрами n и p, где $0 \le p \le 1$, если ξ принимает значения $0, 1, 2, \ldots, n$ с вероятностями $P(\xi = k) = C_n^k p^k (1-p)^{n-k}$. Случайная величина с таким распределением определяет число успехов в n испытаниях с вероятностью успешного исхода p. Такое распределение, например, используется при определении числа отказов при испытаниях определенного объема изделий, количества заказов, которые будут получены в результате следующих звонков в магазин, а также количества людей среди опрошенных, которые выразят желание купить данный товар и т.д.

Пример. Необходимо определить вероятность того, что трое из пяти человек, зашедших в магазин, купят товар марки «А», причем известно, что 85% покупателей предпочитают именно товар этой марки.

Решение. Для определения вероятности $P(\xi = k)$ следует использовать формулу =БИНОМРАСП(*k*;*n*;*p*;0), для определения вероятности того, что результат окажется меньше или равным k ($P(\xi \le k)$) формулу =БИНОМРАСП(*k*;*n*;*p*;1).

Для вызова требуемой функции следует выполнить следующие действия:

– в MS Excel 2003: меню **Вставка → Функция;**

– в MS Excel 2007: вкладка **Формулы → Вставить функцию;**

– в окне диалога **Мастер функций Шаг 1 из 2** в списке **Категория: Статистические** выбрать функцию БИНОМРАСП и нажать кнопку **ОК**.

В диалоговом окне **Аргументы функции** (рис. 4.1) задать следующие параметры:

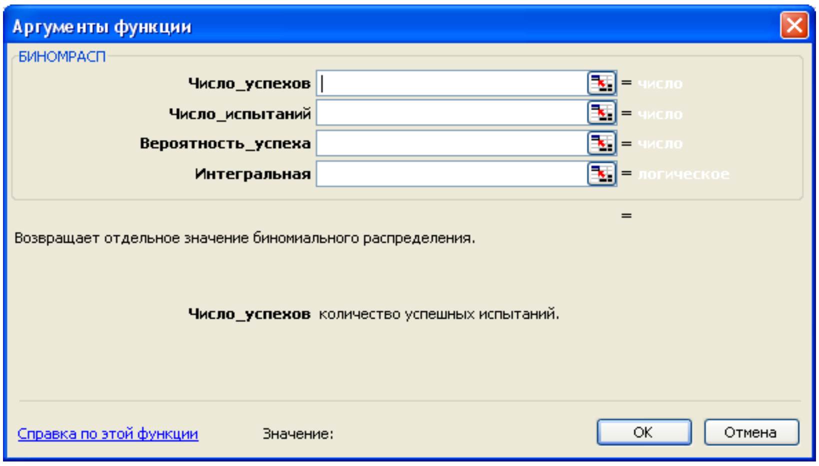


Рис. 4.1. Параметры биномиального распределения

1. *Число_успехов* – количество успешных событий.
2. *Число_испытаний* – количество независимых испытаний.
3. *Вероятность_успеха* – вероятность наступления успешного события.
4. *Интегральная* – логическое значение. Если 1, то функция рассчитывает интегральную (кумулятивную) функцию распределения, т.е. вероятность того, что число успешных событий не больше значения аргумента *Число_успехов*. Если 0, то рассчитывается дифференциальная функция распределения, т.е. вероятность того, что число успешных событий равно значению аргумента *Число_успехов*.

Для рассматриваемого примера значение функции =БИНОМРАСП(3; 5; 0,85; 1) равно 0,16479, т.е. вероятность того, что из пяти покупателей число купивших товар марки «А» будет не больше трех равно 0,16479. Значение функции =БИНОМРАСП(3;5;0,85;0) равно 0,138178, это значит, что трое из пяти человек, зашедших в магазин, купят товар марки «А».

Гипергеометрическое распределение

Гипергеометрическое распределение применяется при решении задач контроля качества продукции. Вероятность того, что из n изделий, выбранных случайным образом из партии объемом N, ровно k изделий с дефектом, имеет гипергеометрическое распределение. Случайная величина ξ имеет гипергеометрическое распределение с параметрами n, N, K, где $K \le N$, $n \le N$, если принимает целые значения от максимума $(0,N\text{-}K\text{-}n)$ до минимума (n,N) с вероятностями:

$$P(\xi = k) = \frac{C_K^k C_{N-K}^{n-k}}{C_N^n},$$

где k – число изделий с дефектом в выборке;

K – число изделий с дефектом в генеральной совокупности;

n – объем выборки;

N – объем генеральной совокупности.

Пример. Из партии, содержащей 40 изделий, случайным образом выбираются 5 и повергаются проверке на качество. Если в результате контроля только одно изделие оказывается бракованным, то партия принимается. В противном случае вся партия бракуется. Требуется определить вероятность того, что партия будет принята, если из 40 изделий 6 имеют дефекты. Вероятность получить k из выборки объемом n определяется с помощью функции MS Excel = ГИПЕРГЕОМЕТ() со следующими аргументами:

1. *Число_успехов_в_выборке* (k) – количество успешных событий в выборке.
2. *Размер_выборки* (n).
3. *Число_успехов_в_совокупности* (K).
4. *Размер_совокупности* (N).

Для примера искомая вероятность будет равна сумме двух функций:

=ГИПЕРГЕОМЕТ(0,5,6,40)+ГИПЕРГЕОМЕТ(1,5,6,40).

В итоге вероятность того, что партия будет принята, равна 0,846.

Распределение Пуассона

Данное распределение описывает число событий, происходящих в одинаковых промежутках времени при условии, что события происходят независимо друг от друга с постоянной интенсивностью λ, например, число отказавших изделий за определенный интервал, число требований выплаты страховых сумм в год, число вызовов в ремонтную службу за сутки. Случайная величина ξ имеет распределение Пуассона с параметром λ, если она принимает значения 0,1,2 … с вероятностями $P(\xi = k) = \frac{\lambda^k}{k!} e^{-\lambda}$.

Пример. Фирма выпускает оборудование высокого качества, благодаря чему каждый день ожидается возврат в среднем только 1,5 единиц товара на гарантийный ремонт. Требуется рассчитать, с какой вероятностью завтра на гарантийный ремонт не поступит ни одного изделия или, например, поступит не более одного. Для того чтобы получить вероятность $P(\xi = k)$ следует использовать функцию =ПУАССОН(k,λ,0), а для вычисления вероятности $P(\xi \le k)$ – =ПУАССОН(k,λ,1) со следующими параметрами:

1. *Количество событий* – Х.

2. *Среднее* – интенсивность появления событий.

3. *Интегральная* – логическое значение. Если 1, то функция рассчитывает интегральную (кумулятивную) функцию распределения, т.е. вероятность того, что число успешных событий не больше значения аргумента *X*. Если 0, то рассчитывается дифференциальная функция распределения, т.е. вероятность того, что число успешных событий равно значению *X*.

Таким образом, вероятность, что завтра на гарантийный ремонт не поступит ни одного изделия =ПУАССОН(0, 1,5,0) равна 0,22, а вероятность, что не более одного – =ПУАССОН(1, 1,5,1) = 0,56.

Нормальное распределение

Нормальное распределение часто применяется в тех случаях, когда возможные значения лежат вокруг некого среднего значения, например, при стрельбе по мишени, измерении какого-либо параметра и т.д.

Пример. Компания по производству верхней одежды в результате измерения роста группы людей получила среднее значение, равное 172 см и стандартное отклонение – 5 см. Требуется определить, какой процент общего числа производимых пальто должны составлять пальто 3-го роста (158–164 см). Для решения задачи следует использовать функцию =НОРМРАСП() со следующими параметрами:

1. *X* – значение, для которого рассчитывается вероятность.
2. *Среднее* – среднее значение.
3. *Стандартное_откл* – стандартное отклонение.
4. *Интегральная* – логическое значение. Если 1, то функция рассчитывает интегральную (кумулятивную) функцию распределения, т.е. вероятность того, результат не больше значения аргумента *X*. Если 0, то рассчитывается дифференциальная функция распределения, т.е. результат равен значению *X*.

Итоговое значение будет определяться как разность: =НОРМРАСП(164;172;5;1) – НОРМРАСП(158;172;5;1) = 0,0548 – 0,0025 = 0,052, т.е. пальто 3-го роста должны составлять 5,2% от общего числа.

χ^2*-распределение (Пирсона)*

Предположим, что $\xi_1, \xi_2 \ldots \xi_n$ – независимые случайные величины, каждая из которых имеет нормальное распределение. Закон распределения суммы $\xi_1^2 + \xi_2^2 + \ldots + \xi_n^2$ называется χ^2-распределением с n – степенями свободы.

Вероятность $p(\chi^2)$ можно определить с помощью функции =ХИ2РАСП() при следующих аргументах:

1. *X* – значение, для которого рассчитывается χ^2- распределение.
2. *Степени_свободы* – число степеней свободы.

Данное распределение используется при проверке статистических гипотез с помощью критерия согласия Пирсона. Для определения квантиля χ^2-распределения следует использовать функцию – =ХИ2ОБР(*уровень значимости; число степеней свободы*), возвращающее обратное значение вероятности.

F-распределение (распределение Фишера)

Распределение Фишера используется, например, при проверке адекватности уравнений регрессии. Для расчета значений вероятности используется функция =FРАСП() со следующими параметрами:

1. Х – значение, для которого рассчитывается функция.
2. *Степени_свободы1* – число степеней свободы первого набора данных.
3. *Степени_свободы2* – число степеней свободы второго набора данных.

Для определения квантиля *F*-распределения следует использовать функцию =FРАСПОБР(*уровень значимости*; *число степеней свободы1*; *число степеней свободы1*), возвращающее обратное значение вероятности.

Распределение Вейбулла

Распределение Вейбулла часто применяют в теории надежности при описании этапа старения аппаратуры. Для вычисления вероятности применяется функция – =ВЕЙБУЛЛ() с параметрами:

1. *X* – значение, для которого вычисляется функция.
2. *Альфа* – параметр распределения.
3. *Бета* – параметр распределения.
4. *Интегральная* – логическое значение. Если 1, то функция рассчитывает интегральную (кумулятивную) функцию распределения, т.е. вероятность того, что результат не больше значения аргумента *X*. Если 0, то рассчитывается дифференциальная функция распределения, т.е. результат равен значению *X*.

Экспоненциальное распределение

Экспоненциальное распределение применяют в теории надежности при описании режима нормальной эксплуатации техники. Для вычисления вероятности применяется функция – =ЭКСПРАСП() с параметрами:

1. *X* – значение, для которого вычисляется функция.
2. *Лямбда* – параметр распределения.
3. *Интегральная* – логическое значение. Если 1, то функция рассчитывает интегральную (кумулятивную) функцию распределения, т.е. вероятность того, что результат не больше значения аргумента *X*. Если 0, то рассчитывается дифференциальная функция распределения, т.е. результат равен значению *X*.

Содержание работы

Основная часть работы состоит в решении наиболее распространенных практических задач, использующих законы распределения случайных величин.

Порядок выполнения

1. Изучить и решить примеры, описанные в теоретической части работы.
2. Выполнить задания для самостоятельного решения:

2.1 Количество клиентов, которыми располагает фирма, равно 8, а вероятность, с которой может позвонить каждый из них, – 0,25. Чему равна вероятность, что завтра позвонят не менее 3-х клиентов?

2.2 Чему равна вероятность того, что завтра на ремонт поступит 4 детали, если в среднем возвращается 1,8 единиц.

2.3 Результаты опроса общественного мнения говорят о том, что наибольшей популярностью покупателей пользуются мужские пальто серого цвета (30% опрошенных), затем пальто черного цвета (26% опрошенных), синего (22%) и коричневого (22%). Имеется ли существенная разница между цветами пальто с точки зрения покупателей, если считать, что результаты получены при $n = 100$. Для проверки гипотезы о принадлежности данных опроса к равномерному распределению используйте критерий χ^2-квадрат при уровне значимости $\alpha = 0,05$.

2.4 Решите задачу из предыдущего пункта при $n = 1000$.

Отчётность по работе

После выполнения работы обучаемый представляет отчет. Отчёт должен содержать:

1. Название и цель работы.
2. Результаты решения примеров, описанных в теоретической части.
3. Результаты решения контрольных заданий.
4. Выводы по результатам работы.

Контрольные вопросы

1. Сформулируйте определение понятия «закон распределения случайной величины».
2. Приведите примеры применения экспоненциального закона распределения для описания случайных событий.
3. Покажите порядок ввода формул в Excel.
4. Перечислите параметры нормального распределения.
5. Какие случайные события можно описать с помощью нормального закона распределения?
6. Что такое дисперсия?
7. Раскройте понятие квантиля.
8. Каким образом определяется степень свободы?
9. Перечислите действия при применении критерия согласия.
10. В чем причина отличия результатов решения задач из пунктов 2.3 и 2.4?

5. ГЕНЕРАЦИЯ СЛУЧАЙНЫХ ЧИСЕЛ И АНАЛИЗ ВЫБОРКИ ДАННЫХ

Цель работы

Изучить средства программы Microsoft Excel для генерации случайных чисел с требуемыми законами распределения, для построения и анализа выборок данных.

Теоретические основы

Генерация случайных чисел

Для того чтобы сформировать массив случайных чисел, распределенных по требуемому закону распределения, можно воспользоваться функцией **Генерация случайных чисел**.

Данная функция входит в состав надстройки **Анализ данных.** Для вызова этой надстройки необходимо:

– Excel 2003: в меню **Сервис**, выбрать команду **Анализ данных**;

– Excel 2007: на вкладке **Данные**, в группе **Анализ** выбрать команду **Анализ данных.**

В окне диалога **Анализ данных** необходимо указать функцию **Генерация случайных чисел** и нажать **OK.** Появится соответствующее диалоговое окно (рис. 5.1).

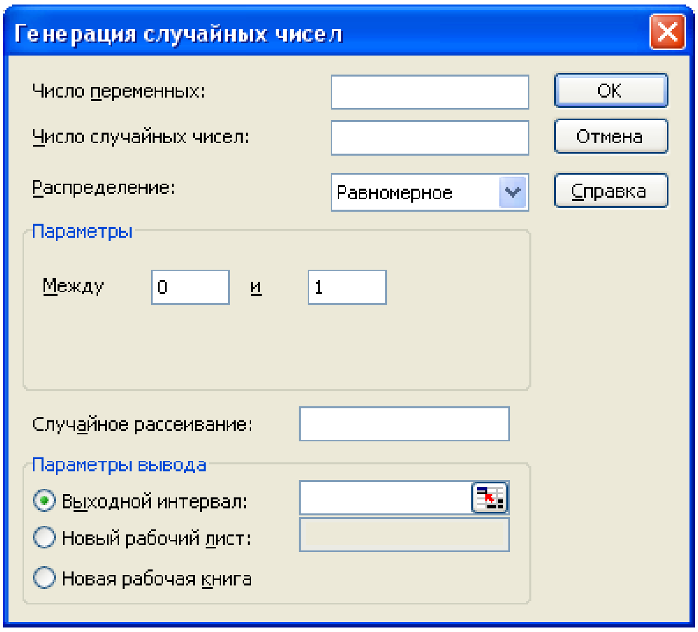


Рис. 5.1. Генерация случайных чисел

В диалоговом окне (рис. 5.1) необходимо задать следующие параметры:

1. *Число переменных* – вводится число столбцов значений, которые необходимо разместить в выходном диапазоне.

2. *Число случайных чисел* – вводится число значений, которое необходимо вывести в каждом столбце выходного диапазона.

3. *Распределение* – в выпадающем списке выбирается тип распределения, необходимый для генерации случайных чисел.

4. *Параметры* – вводятся параметры выбранного распределения.

5. *Случайное рассеивание* – вводится произвольное значение, для которого необходимо генерировать случайные числа. Впоследствии можно снова использовать это значение для получения тех же самых случайных чисел.

6. *Параметры вывода* – указывается место, где будет сгенерирован массив.

Построение выборки данных

На практике при проведении исследований часто ограничиваются только выборочным набором объектов из генеральной совокупности, т.е. **выборкой**. Работа с выборкой существенно экономит время и средства.

Случайная выборка строится таким образом, чтобы каждый объект генеральной совокупности имел одинаковую вероятность быть выбранным, и при этом объекты отбираются независимо друг от друга. Существует два основных типа случайной выборки: выборка без возврата (любой объект не может попасть в выборку более одного раза) и выборка с возвратом (объект может попасть в выборку более одного раза, т.е. после выбора объект снова возвращается в генеральную совокупность).

Для получения систематической выборки в генеральной совокупности определяют случайную начальную точку и отбирают элементы, начиная с этой точки через постоянный интервал (с постоянным шагом отбора). Для того чтобы из совокупности N получить систематическую выборку размером n, необходимо выбирать элементы с шагом N/n. Данный метод используется в том случае, когда есть предварительно пронумерованные списки, платежные поручения, приходные и расходные чеки и т.д., в подобных случаях он быстрее и проще, чем метод простой случайной выборки.

В Excel реализованы случайная выборка с возвратом и систематическая выборка с помощью функции **Выборка**. Вызвать эту функцию можно в окне диалога **Анализ данных.** В этом случае появится диалоговое окно **Выборка** (рис. 5.2).

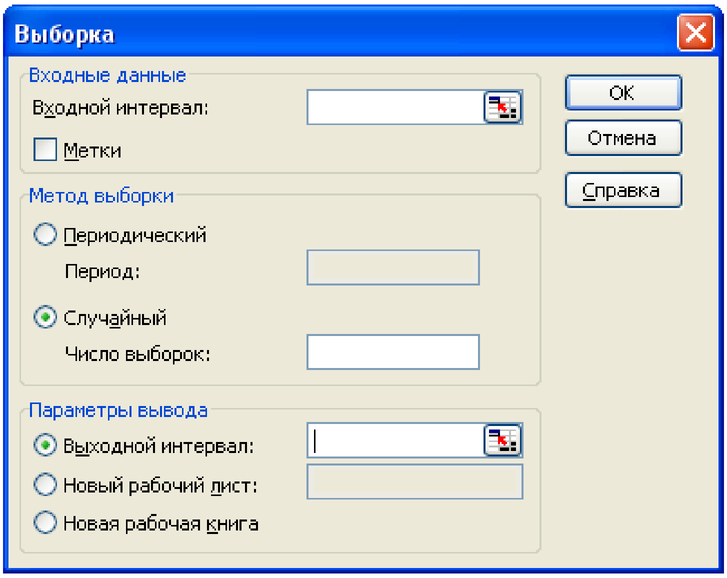


Рис. 5.2. Моделирование выборки

В диалоговом окне необходимо задать следующие параметры:

1. *Входной интервал* – вводится диапазон ячеек, содержащих анализируемые данные.

2. *Переключатель* в группе **Метод выборки** для случайной выборки с возвратом должен быть установлен в положение **Случайный**. В поле **Число выборок** вводится количество размещаемых в выходном столбце случайных значений (размер выборки). Для формирования систематической выборки следует выбрать переключатель **Периодический** и в поле **Период** указать шаг отбора, элементы будут отбираться через постоянный интервал.

3. *Параметры вывода* – указывается место, где будет сгенерирована выборка.

Анализ данных

Для того чтобы обнаружить общие свойства совокупности данных, выявить закономерности, тенденции развития процесса и в результате прийти к правильным выводам, необходимы обобщающие количественные показатели. Эти показатели можно условно разделить на четыре группы.

1. Показатели, которые описывают закон распределения данных: таблицы частот, полигоны, гистограммы.

2. Показатели уровня – описывают положение данных на числовой оси: минимальный и максимальный элементы выборки, верхний и нижний квартили, перцентиль, различные средние и другие характеристики.

3. Показатели рассеивания – описывают степень разброса данных относительно своего центра: дисперсия, среднеквадратичное отклонение, размах выборки.

4. Показатели асимметрии – характеризуют симметрию распределения данных около своего центра: коэффициент асимметрии, эксцесс.

Гистограмма

Для графического отображения данных, представляющих вариационный ряд можно построить **гистограмму** – столбчатую диаграмму частот. По оси абсцисс откладываются значения интервалов, а по оси ординат – частоты в виде столбиков, высота которых соответствует частоте попадания случайной величины в интервал. В Excel для построения гистограммы применяется функция **Гистограмма**. Вызвать эту функции можно в окне диалога **Анализ данных.** Появится диалоговое окно **Гистограмма** (рис. 5.3).

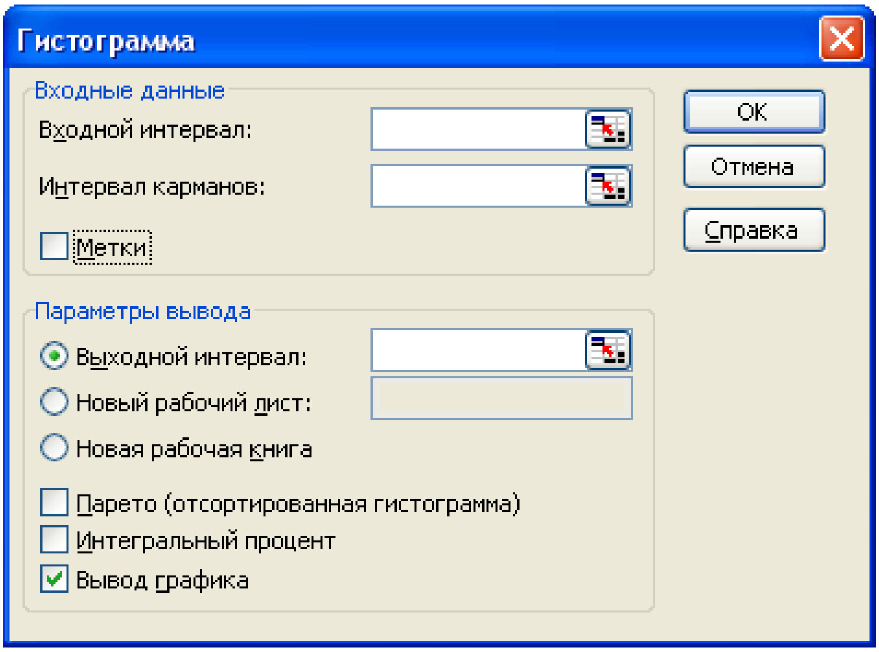


Рис. 5.3. Построение гистограммы

В диалоговом окне **Гистограмма** задаются следующие параметры:

1. *Входной интервал* – вводится диапазон ячеек, содержащих анализируемые данные.

2. В поле *Интервал карманов* вводится диапазон ячеек, содержащих значения границ интервалов (параметр является необязательным, в этом случае набор интервалов создается автоматически).

3. *Параметры вывода* – указывается место, где будет указана таблица частот.

4. Для *вывода гистограммы* следует установить флажок опции Вывод графика. Флажки опций Парето (отсортированная гистограмма) и Интегральный процент (накопленные) частоты следует оставить сброшенными.

Показатели уровня

Средняя арифметическая: для расчета используется функция =СРЗНАЧ(*число1;число2*;…) из категории Статистические.

Средняя геометрическая: $\bar{x} = \sqrt[n]{\prod_{i=1}^{n} x_i}$, для расчета используется функция =СРГЕОМ(*число1;число2*;…) из категории Статистические.

Медиана: величина, которая соответствует варианту, находящемуся в середине ранжированного ряда.

Пример. Для ряда 1,2,3,3,6,7,8,8,10 медианой будет величина, которая расположена в центре ряда, т.е. пятая величина, действительно значение функции =МЕДИАНА(1;2;3;3;6;7;8;8;10) равно 6.

Мода: значение признака, которое встречается наиболее часто среди элементов совокупности, для расчета используется функция =МОДА(*число1;число2*;…) из категории Статистические.

Минимум: для расчета используется функция =МИН(*число1;число2*;…) из категории Статистические.

Максимум: для расчета используется функция =МАКС(*число1;число2*;…) из категории Статистические.

Ранг: номер (порядковое место) значения случайной величины в наборе данных.

Перцентиль: обобщает информацию о рангах, характеризуя значение, достигаемое заданным процентом общего количества данных, т.е. являются характеристиками набора данных, которые выражают ранги элементов в виде процентов.

Для расчета рангов и перцентилей используется функция **Ранг и перцентиль**.

Для вызова этой функции необходимо в окне диалога **Анализ данных** выбрать соответствующую функцию. Появится диалоговое окно **Ранг и перцентиль** (рис. 5.4).

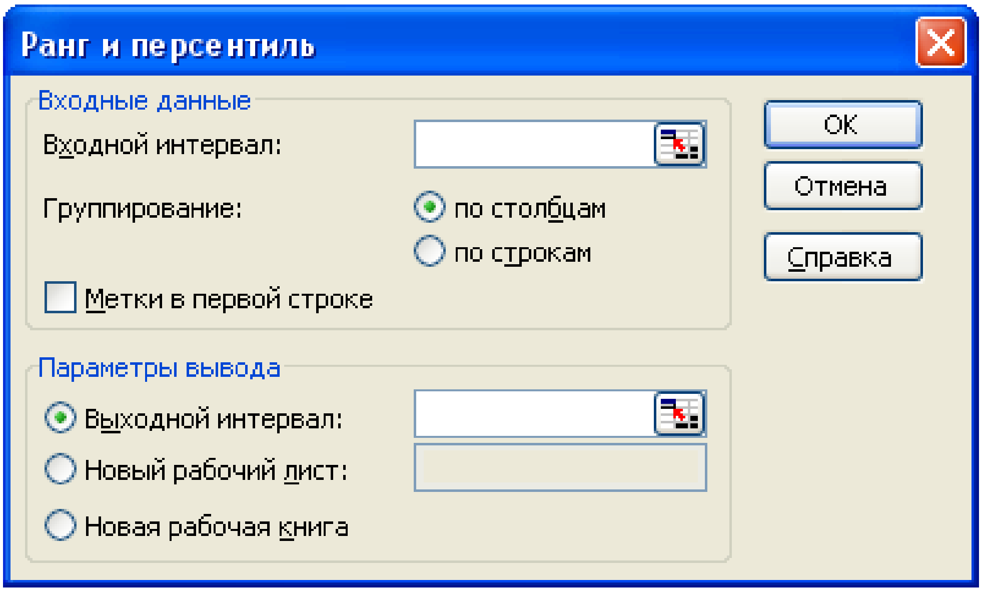


Рис. 5.4. Ранг и перцентиль

В диалоговом окне (рис. 5.4) задаются следующие параметры:

1. *Входной интервал* – вводится диапазон ячеек, содержащих анализируемые данные.
2. *Группирование* – по строкам и столбцам в зависимости от расположения данных во входном диапазоне.
3. *Метки в первой строке* – флажок ставится, если первая строка содержит заголовок, в противном случае будут созданы стандартные заголовки автоматически.
4. *Параметры вывода* – указывается место, где будет указана таблица рангов и перцентилей.

Показатели рассеяния

Размах выборки – разности между максимальным и минимальным значением.

Среднее линейное отклонение – среднее арифметическое из абсолютных значений отклонений от средней: $(\sum|x_i - \bar{x}|)/n$. Для расчета среднего линейного отклонения используется функция =СРОТКЛ(*число1;число2*;…) из категории Статистические.

Среднее квадратическое отклонение (ско). Для расчета используется функция =СТАНДОТКЛОН(*число1;число2*;…) из категории Статистические.

Дисперсия. Для расчета используется функция =ДИСП(*число1;число2*;…) из категории Статистические.

Показатели асимметрии

Эксцесс – характеризует островершинность (плосковершинность) распределения. Если эксцесс больше нуля, то распределение островершинное, если меньше нуля – плосковершинное. Для расчета используется функция =ЭКСЦЕСС(*число1;число2*;…) из категории Статистические.

Асимметрия – характеризует меру несимметричности (скошенности) распределения. Если коэффициент асимметрии больше нуля, то асимметрия правосторонняя, если меньше нуля – левосторонняя. Для расчета используется функция =СКОС(*число1;число2*;…) из категории Статистические.

Описательная статистика

Для получения статистического отчета одновременно по всем основным показателям используют специальный инструмент анализа: **Описательная статистика**, который находится в надстройке **Анализ данных.** При вызове функции **Описательная статистика** появится соответствующее диалоговое окно (рис. 5.5).

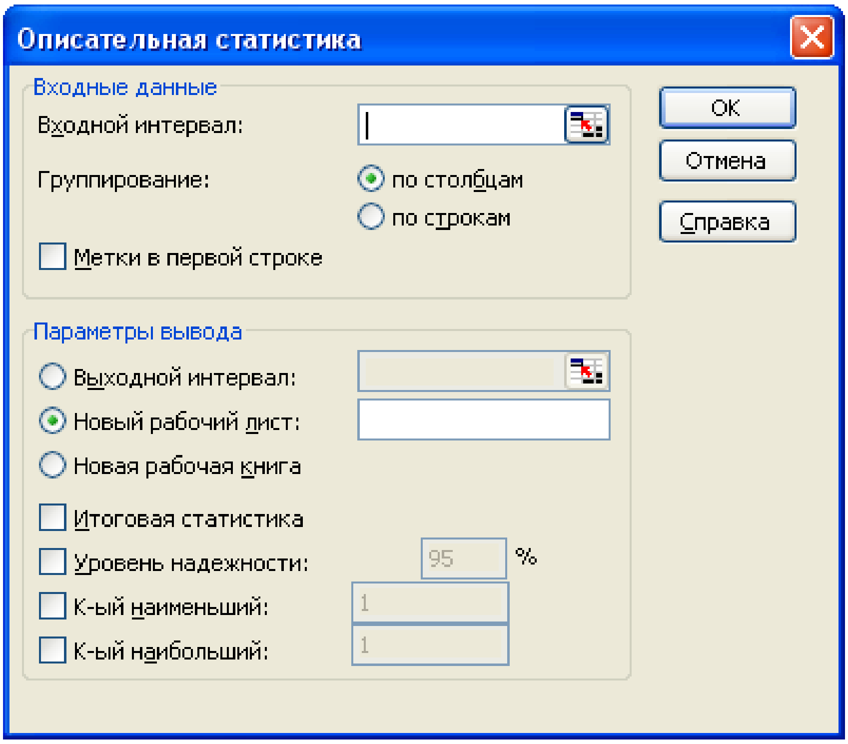


Рис. 5.5. Описательная статистика

В диалоговом окне (рис. 5.5) задаются следующие параметры:

1. *Входной интервал* – вводится диапазон ячеек, содержащих анализируемые данные.

2. *Группирование* – по строкам и столбцам в зависимости от расположения данных во входном диапазоне.

3. *Метки в первой строке* – флажок ставится, если первая строка содержит заголовок, в противном случае будут созданы стандартные заголовки автоматически.

4. *Параметры вывода* – указывается место, где будет указана таблица результатов анализа.

5. *Итоговая таблица* – флажок ставится, если необходимо получить результаты по каждому показателю.

6. *Уровень надежности* – флажок устанавливается, если требуется вывести значение ошибки выборки при установленном уровне надежности.

7. *К-ый наименьший и К-ый наибольший* – флажки устанавливаются, если требуется получить определенный наименьший элемент (начиная с минимального значения среди элементов выборки) или наибольший (начиная с максимального значения выборки).

Содержание работы

Основная часть работы состоит в решении практических задач, использующих возможности программы Excel для генерации случайных чисел с требуемыми законами распределения и анализа выборки данных.

Порядок выполнения

1. Сформировать в столбцах массивы случайных чисел, распределенных по требуемому закону распределения, число переменных = 1, число случайных чисел = 100:

1.1. Распределение Бернулли. Исходные данные: р = 0,3.

1.2. Биномиальное распределение. Исходные данные: р = 0,85; число испытаний – 25.

1.3. Нормальное распределение. Исходные данные: среднее – 100; стандартное отклонение – 20.

2. На основе построенных в п.1.1–1.3 массивов получить случайную и систематическую выборки размером 20.

3. Построить гистограммы для нормального распределения (п.1.3) и выборок, полученных в п.2, сравнить вид гистограмм.
4. Для случайных выборок, полученных в п.2 рассчитать показатели анализа данных, описанные в теоретической части с помощью функций Excel.
5. Для случайных выборок, полученных в п.2 рассчитать показатели анализа данных, описанные в теоретической части с помощью инструмента «Описательная статистика». Сделать выводы по результатам анализа данных.

Контрольные вопросы

1. Сформулируйте определение понятия «закон распределения случайной величины».
2. Приведите примеры применения нормального закона распределения для описания случайных событий.
3. Что показывают показатели асимметрии и эксцесса?
4. Перечислите этапы процесса построения гистограммы.
5. В чем отличие систематической выборки от случайной?

6. АНАЛИЗ ВРЕМЕННЫХ РЯДОВ

Цель работы

Изучить средства программы Microsoft Excel для анализа временных рядов.

Теоретические основы

Временной ряд представляет собой последовательность данных, описывающих объект в последовательные моменты времени.

Метод скользящего среднего

Для сглаживания на основе простого скользящего среднего используется функция **Скользящее среднее**. Её следует вызвать в окне диалога **Анализ данных,** указав имя соответствующей функции. Появится диалоговое окно (рис. 6.1).

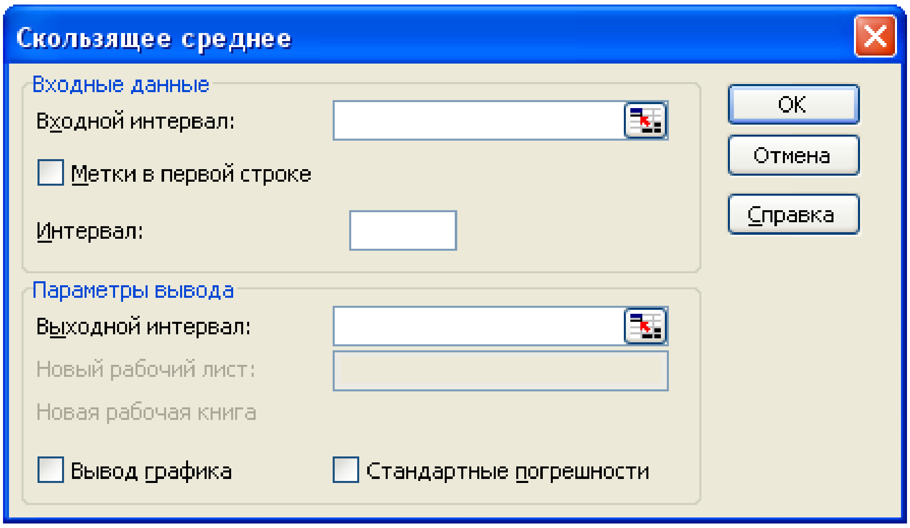


Рис. 6.1. Скользящее среднее

В диалоговом окне (рис. 6.1) задаются следующие параметры:

1. *Входной интервал* – вводится диапазон ячеек, содержащих анализируемые данные.

2. *Метки в первой строке* – флажок ставится, если первая строка содержит заголовок, в противном случае будут созданы стандартные заголовки автоматически.

3. *Интервал* – в поле вводится число уровней, входящих в интервал сглаживания. По умолчанию значение равно 3.

4. *Параметры вывода* – указывается место, где будет указана таблица результатов анализа.

5. Для *вывода графика* и стандартных погрешностей необходимо установить соответствующие опции.

Метод экспоненциального сглаживания

На практике часто возникает ситуация, когда последние данные временного ряда являются более значимыми, чем предыдущие и, следовательно, при вычислении скользящего среднего наблюдениям необходимо приписывать различные веса. При экспоненциальном сглаживании более старым наблюдениям приписываются экспоненциально убывающие веса, при этом, в отличие от скользящего среднего, учитываются все предшествующие наблюдения ряда. Формула экспоненциального сглаживания имеет вид:

$S_t = \alpha \cdot y_t + (1-\alpha) \cdot S_{t-1}$ и применяется рекурсивно, т.е. каждое новое сглаженное значение S_t вычисляется как взвешенное среднее текущего наблюдения y_t и сглаженного ряда S_{t-1} (предыдущие наблюдения). Очевидно, что если $\alpha = 1$, то предыдущие наблюдения полностью игнорируются, а если $\alpha = 0$, то игнорируются текущие наблюдения.

Для экспоненциального сглаживания используется функция **Экспоненциальное сглаживание**. Для ее вызова необходимо выбрать соответствующее имя в окне диалога **Анализ данных.** Появится диалоговое окно (рис. 6.2).

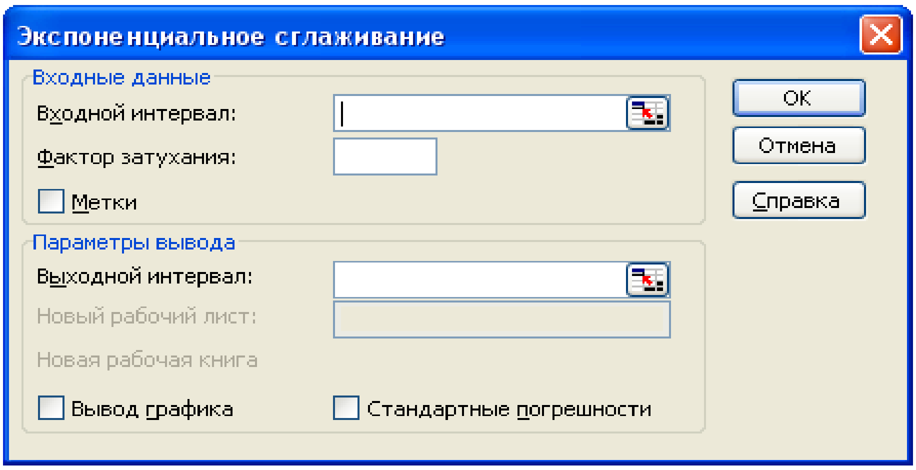


Рис. 6.2. Экспоненциальное сглаживание

В диалоговом окне (рис. 6.2) задаются следующие параметры:

1. *Входной интервал* – вводится диапазон ячеек, содержащих анализируемые данные.

2. *Фактор затухания* – в поле вводится значение коэффициента экспоненциального сглаживания α. По умолчанию значение принимается равным 0,3.

3. *Метки в первой строке* – флажок ставится, если первая строка содержит заголовок, в противном случае будут созданы стандартные заголовки автоматически.

4. *Параметры вывода* – указывается место, где будет указана таблица результатов анализа.

5. Для *вывода графика* и стандартных погрешностей необходимо установить соответствующие опции.

Метод аналитического выравнивания

Основным содержанием метода аналитического выравнивания временных рядов является расчет общей тенденции развития (тренда) как функции времени на основе математической модели, которая наилучшим образом отображает основную тенденцию развития временного ряда. В качестве моделей обычно используют линейную, показательную, степенную и логарифмическую функции в зависимости от особенностей временного ряда. Для оценки точности трендовой модели используется коэффициент детерминации, равный отношению дисперсии теоретических данных, полученных по трендовой модели к дисперсии эмпирических данных. Считается, что трендовая модель адекватна изучаемому процессу и отражает тенденцию его развития при значениях коэффициента детерминации близких к единице.

Порядок построения трендовой модели с помощью Excel 2003 следующий:

1. Эмпирический временной ряд следует представить в виде диаграммы одного из следующих типов: гистограмма, линейная диаграмма, график, точечная диаграмма, диаграмма с областями, а затем щелкнуть правой кнопкой мыши на одном из маркеров данных и в открывшемся контекстном меню выбрать команду **Добавить линию тренда**.

Откроется окно диалога (рис. 6.3).

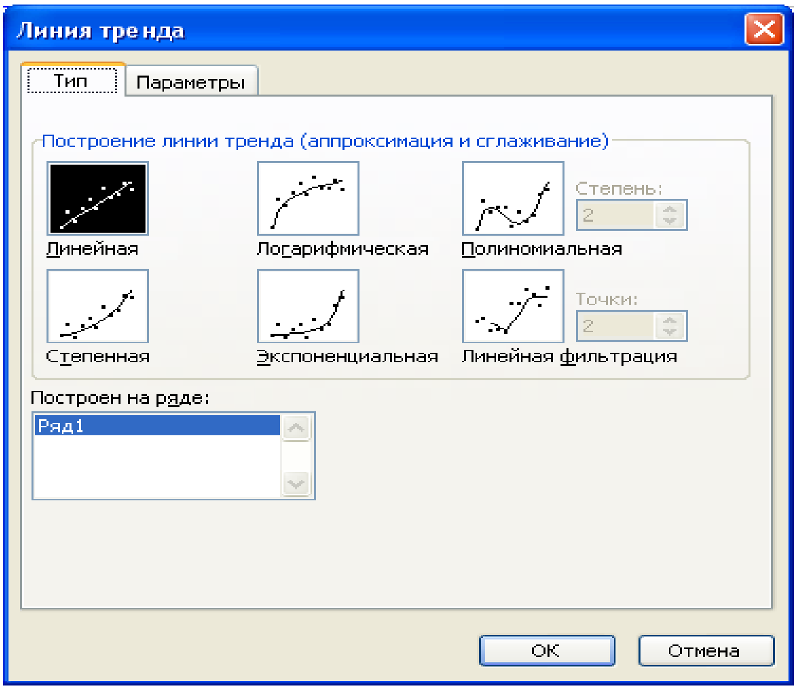


Рис. 6.3. Окно выбора модели тренда

2. На вкладке **Тип** выбрать требуемый вид трендовой модели.

3. На вкладке **Параметры** (рис. 6.4) задать дополнительные параметры тренда:

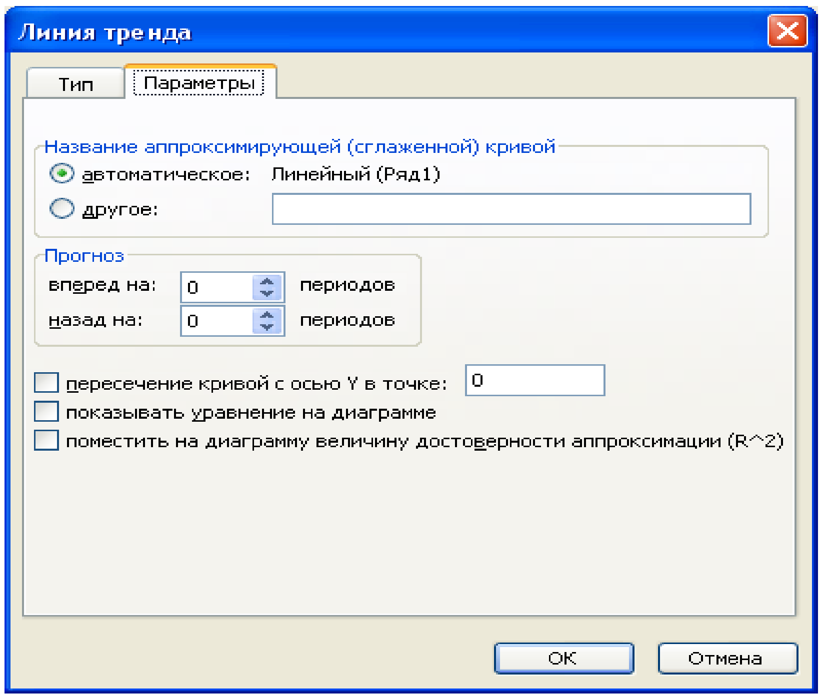


Рис. 6.4. Дополнительные параметры модели

– *название сглаженной кривой* – автоматическое по типу модели или собственное название функции;

– *прогноз* – можно указать, на сколько периодов вперед (назад) требуется спроектировать линию тренда в будущее (прошлое);

– если необходимо указать точку пересечения кривой с осью ординат, показать уравнение аппроксимирующей функции и вывести значение коэффициента детерминации следует установить соответствующие флажки.

Реализация метода аналитического выравнивания в Excel 2007 проводится аналогичным образом. Для открытия окна диалога с настройками параметров тренда необходимо использовать команду контекстного меню **Формат линии тренда**. Само диалоговое окно содержит те же параметры, что и в версии Excel 2003.

Использование статистических функций

Спрогнозировать данные без использования диаграмм, соответствующие линейным и экспоненциальным линиям, можно с использованием статистических функций ТЕНДЕНЦИЯ() и РОСТ().

Функция ТЕНДЕНЦИЯ возвращает значения в соответствии с линейным трендом. Аппроксимирует прямой линией (по методу наименьших квадратов) массивы *известные_значения_y* и *известные_значения_x*. Возвращает значения y, в соответствии с этой прямой для заданного массива *новые_значения_x*.

Синтаксис функции: ТЕНДЕНЦИЯ(y; x; n_x; конст):

y – *известные_значения_y* – множество значений y, для которых уже известна линейная зависимость;

x – *известные_значения_x* – множество значений x, для которых уже известна линейная зависимость;

n_x – *новые_значения_x* – новые значения x, для которых функция возвращает соответствующие значения y.

конст – логическое значение, если оно равно 0, то свободный член равен нулю, в противном случае свободный член вычисляется обычным образом.

Функция РОСТ рассчитывает прогнозируемый экспоненциальный рост на основании имеющихся данных. Синтаксис функции: РОСТ(y; x; n_x; конст):

y – *известные_значения_y* – множество значений y, для которых уже известна экспоненциальная зависимость;

x – *известные_значения_x* – множество значений x, для которых уже известна экспоненциальная зависимость;

n_x – *новые_значения_x* – новые значения x, для которых функция возвращает соответствующие значения y.

конст – логическое значение, если оно равно 0 или отсутствует, то константа равна единице, в противном случае вычисляется обычным образом.

Перед вводом формул ТЕНДЕНЦИЯ() и РОСТ() необходимо выделить результирующие ячейки, ввести ссылки на известные и новые значения, при необходимости ввести единицу в строку *Константа* и нажать комбинацию клавиш <Ctrl+Shift+Enter> для ввода данных в массив.

Содержание работы

Основная часть работы состоит в решении практических задач, использующих возможности программы Excel для анализа временных рядов.

Порядок выполнения

1. Для ряда, изображенного на графике (рис. 6.5) применить методы скользящего среднего и экспоненциального сглаживания средствами Excel, сравнить результаты и сделать выводы.

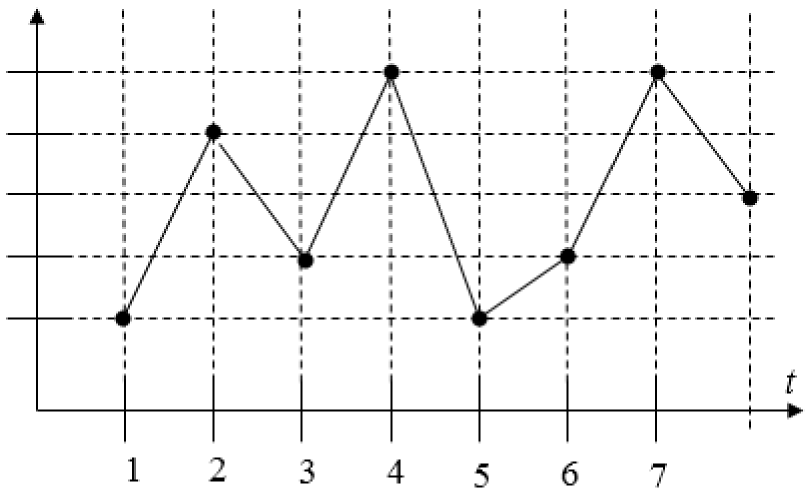


Рис. 6.5. Временной ряд

2. Сформировать в столбцах массивы случайных чисел, распределенных по требуемому закону распределения, число переменных = 1, число случайных чисел = 100:

2.1. Биномиальное распределение. Исходные данные: $p = 0,8$; число испытаний – 25.

2.2. Нормальное распределение. Исходные данные: среднее – 100; стандартное отклонение – 15.

3. На основе построенных массивов п.2 получить случайную выборку размером 14.

4. Для полученных исходных данных п.3 построить несколько вариантов типов моделей и выбрать наиболее адекватную модель тренда.

5. Для имеющихся данных продаж за первое полугодие спрогнозируйте объем продаж во втором полугодии, используя линейную и экспоненциальную зависимости с помощью функций ТЕНДЕНЦИЯ() и РОСТ() соответственно. Сравните результаты.

Месяц (x)	1	2	3	4	5	6	7	8	9	10	11	12
тыс. р.(y)	380	250	285	420	440	480						

Контрольные вопросы

1. Дать определение временного ряда. Привести примеры.
2. Перечислить задачи, решаемые при изучении временных рядов.
3. На какие части можно разделить модель временного ряда?
4. Что такое тренд? Какими моделями он может быть описан?
5. Какими способами можно выделить тренд?
6. Раскрыть принцип скользящего среднего.

7. РЕГРЕССИОННЫЙ АНАЛИЗ

Цель работы

Изучить средства программы Microsoft Excel для регрессионного анализа данных.

Теоретические основы

На практике встречается достаточно большое количество взаимосвязей, например, между заработной платой работника и его образованием, объемом выпускаемой продукции и затратами на производство, между производительностью работника и его отношением к труду и т.д. В связи с этим возникает необходимость в анализе двухмерных данных. Для этого используют два базовых инструмента: корреляционный анализ, позволяющий оценить степень взаимосвязи между факторами и регрессионный анализ, который показывает, как можно предсказать поведение одной из двух переменных или управлять ею с помощью другой.

Коэффициент корреляции

Наиболее важным для практического использования является случай, когда связь между признаками линейная. Мера силы линейной связи признаков называется коэффициентом корреляции. Величина коэффициента, близкая к 1, указывает, что зависимость между данными случайными величинами почти линейная. Значения, близкие к нулю, означают, что связь между величинами либо слабая, либо не носит линейного характера.

Для расчета коэффициента корреляции можно использовать функцию Excel =КОРРЕЛ(*массив1;массив2*) из категории статистические.

Для вывода результатов в виде таблицы применяют функцию **Корреляция** из пакета **Анализ данных.** После ее вызова из окна диалога **Анализ данных** откроется соответствующее диалоговое окно (рис. 7.1).

В диалоговом окне (рис. 7.1) задаются следующие параметры:

1. *Входной интервал* – вводится диапазон ячеек, содержащих анализируемые данные.

2. *Группирование* – переключатель устанавливается в требуемое положение в зависимости от расположения исходных данных.

3. *Метки в первой строке* – флажок ставится, если первая строка содержит заголовок, в противном случае будут созданы стандартные заголовки автоматически.

4. *Параметры вывода* – указывается место, где будет указана таблица результатов анализа.

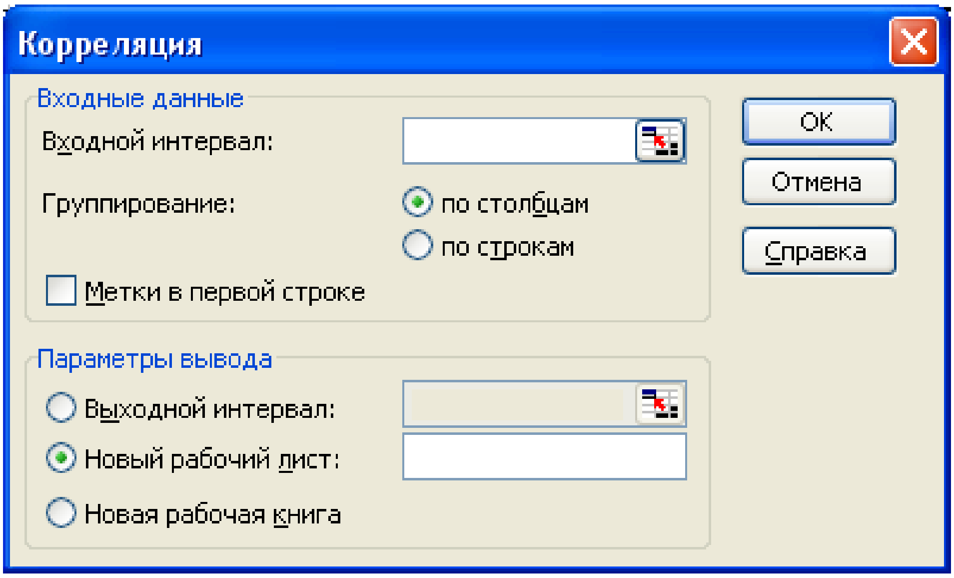


Рис. 7.1. Корреляция

Линейная регрессия

Форма связи результативного признака Y с факторами $X_1, X_2, \ldots X_m$ называется уравнением регрессии. В зависимости от типа выбранного уравнения различают линейную и нелинейную регрессию, а в зависимости от количества факторов – парную (простую, $m = 1$) и множественную (многофакторную, $m > 1$).

На этапе регрессионного анализа решаются следующие задачи:

1. Выбор общего вида уравнения регрессии и определение параметров регрессии.

2. Определение степени взаимосвязи результативного признака и факторов, проверка общего качества уравнения регрессии.

3. Проверка статистической значимости каждого коэффициента уравнения регрессии и определение их доверительных интервалов.

Уравнение простой линейной регрессии имеет вид: $y = b_0 + b_1 x$, множественная линейная регрессия описывается следующим уравнением: $y = b_0 + b_1 x_1 + b_2 x_2 + \ldots + b_m x_m$.

Параметры уравнений парной и множественной регрессий могут быть определены с помощью метода наименьших квадратов, который реализован в Excel. Для этого используется функция **Регрессия**. Для ее вызова необходимо выбрать требуемое имя в окне диалога **Анализ данных.** В результате появится диалоговое окно (рис. 7.2).

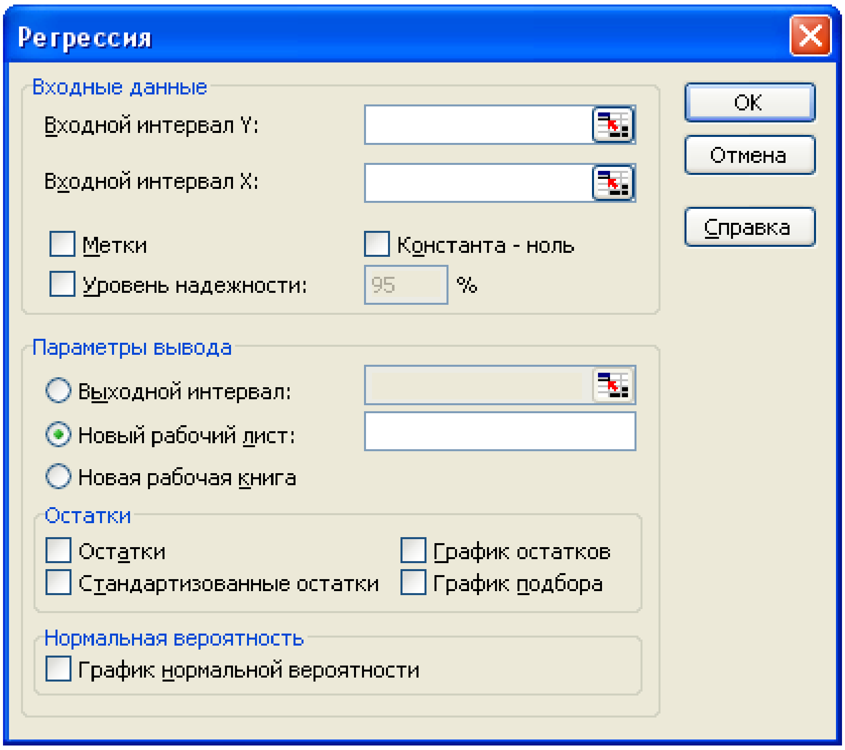


Рис. 7.2. Регрессия

В диалоговом окне (рис. 7.2) задаются следующие параметры:

1. *Входной интервал Y* – вводится диапазон ячеек (один столбец), содержащих исходные данные по результирующему признаку.

2. *Входной интервал X* – вводится диапазон ячеек (число столбцов равно количеству признаков), содержащих исходные данные факторного признака.

3. *Метки* – флажок ставится, если первая строка содержит заголовок, в противном случае будут созданы стандартные заголовки автоматически.

4. *Уровень надежности* – флажок устанавливается, если требуется ввести значение уровня отличное от 95%. При выключенном флажке уровень надежности принимается равным 95%.

5. *Константа-ноль* – флажок устанавливается в том случае, когда требуется, чтобы линия регрессии прошла через начало координат, т.е. $b_0 = 0$.

6. *Параметры вывода* – указывается место, где будут указаны таблицы результатов анализа.

7. *Остатки* – при необходимости вывода столбцов остатков и графиков остатков и подбора необходимо включить соответствующие флажки.

8. *Нормальная вероятность* – флажок устанавливается, если не требуется вывести график зависимости наблюдаемых значений от автоматически формируемых интервалов персентилей.

Пример. Для исходных данных (табл. 7.1) построить регрессионную линейную однофакторную модель зависимости затрат на ТО от срока службы с помощью функции **Регрессия**.

Таблица 7.1

Y	X
затраты на ТО	срок службы
6	2
13	5
23	9
5	3
22	8

Результаты решения с помощью функции **Регрессия** выводятся в виде нескольких отдельных таблиц.

Результаты расчета регрессионной статистики выводятся в табл. 7.2:

Таблица 7.2

Регрессионная статистика	
Множественный R	0,984535285
R-квадрат	0,969309728
Нормированный R-квадрат	0,959079637
Стандартная ошибка	1,724792855
Наблюдения	5

В табл. 7.2 указаны следующие элементы:

– *множественный R* – коэффициент корреляции;

– *R-квадрат* – коэффициент детерминации;

– *нормированный R-квадрат* – нормированное значение коэффициента корреляции;

– *стандартная ошибка* – стандартное отклонение для остатков;

– *наблюдения* – количество исходных наблюдений.

В табл. 7.3 представлены результаты дисперсионного анализа, которые используются для проверки значимости коэффициента детерминации.

Таблица 7.3

Дисперсионный анализ					
	df	*SS*	*MS*	*F*	*Значимость F*
Регрессия	1	281,8752688	281,8752688	94,75084337	0,002303227
Остаток	3	8,924731183	2,974910394		
Итого	4	290,8			

В табл. 7.3 указаны следующие элементы:

df – число степеней свободы. Для строки *Регрессия* это количество факторных признаков, для строки *Остаток* – число наблюдений минус количество переменных в уравнении регрессии, для строки *Итого* – сумма степеней свободы для строк *Регрессия* и *Остаток*.

SS – сумма квадратов отклонений. Для строки *Регрессия* это значение определяется как сумма квадратов отклонений теоретических данных от среднего, для строки *Остаток* это сумма квадратов отклонений эмпирических данных от теоретических, для строки *Итого* это сумма квадратов отклонений эмпирических данных от среднего.

MS – дисперсии. Для строки *Регрессия* это факторная дисперсия, для строки *Остаток* это остаточная дисперсия.

F – расчетное значение *F*-критерия Фишера, определяемое как отношение факторной дисперсии к остаточной.

Значимость F – значение уровня значимости, соответствующее вычисленному значению *F*.

Полученные значения коэффициентов регрессии и их статистические оценки сводятся в табл. 7.4:

Таблица 7.4

	Коэффициенты	*Стандартная ошибка*	*t-статистика*	*P-Значение*
Y-пересечение	-1,064516129	1,710826692	-0,622223241	0,577881718
срок службы	2,752688172	0,282790929	9,734004488	0,002303227

В табл. 7.4 указаны следующие элементы:

– *коэффициенты* – значения коэффициентов модели.

– *стандартная ошибка* – стандартные ошибки коэффициентов.

t-статистика – расчетные значения *t*-критерия, вычисляемого как отношение значений коэффициентов к соответствующим стандартным ошибкам.

P-Значение – значения уровней значимости, соответствующие вычисленным значениям t_p.

В экранной таблице Excel также указываются нижние и верхние границы доверительных интервалов для коэффициентов регрессии – *Нижние 95%, Верхние 95%* (ввиду ограниченности места в табл. 7.4 они опущены).

На основе данных из полученных таблиц можно сделать следующие выводы:

1. Уравнение регрессии имеет вид: $y = -1{,}06 + 2{,}75 \cdot x$.
2. Значение коэффициента детерминации, равного 0,97 показывает, что срок службы существенно влияют на затраты на ТО, что подтверждает правильность включения его в построенную модель.
3. Рассчитанный уровень значимости *Значимость F* = 0,002 меньший 0,05 подтверждает значимость величины коэффициента детерминации.
4. *P-Значение* для срока службы, равное 0,002 и меньшее 0,05 подтверждает значимость коэффициента b_1.
5. *P-Значение* для коэффициента b_0 превышает 0,05, это означает, что данный коэффициент для модели не является значимым и его можно опустить, т.е. график модели будет проходить через точку начала координат.

Для получения новой модели без коэффициента b_0 необходимо еще раз запустить функцию **Регрессия**, в окне (см. рис.7.2) поставить флажок **Константа-ноль**. В результате построятся новые таблицы. Приведем таблицу для значений коэффициентов регрессии (табл. 7.5).

Таблица 7.5

	Коэффициенты	*Стандартная ошибка*	*t-статистика*	*P-Значение*
Y-пересечение	0	–	–	–
срок службы	2,595628415	0,11732738	22,12295556	2,47108E-05

Анализ новых полученных таблиц показывает значимость коэффициента модели и коэффициента детерминации, что подтверждает адекватность полученного уравнения.

В итоге модель получится следующего вида: $y = 2{,}6 \cdot x$.

Если в результате анализа незначимыми окажутся коэффициенты $b_1, b_2, \ldots$, то следует пересчитать результаты регрессии, не указывая в поле *Входной интервал X* (см. рис.7.2) диапазон ячеек с данными соответствующего фактора.

Содержание работы

Основная часть работы состоит в решении практических задач, использующих возможности программы Excel для регрессионного анализа данных.

Порядок выполнения

1. Для исходных данных (табл. 7.1) вычислить корреляцию попарно между признаками, применяя функцию =КОРРЕЛ(*массив1;массив2*) и с помощью функции **Корреляция**. Сделать выводы.
2. Чему равен коэффициент корреляции двух случайных величин, представленных в табл. 7.6.

Таблица 7.6

X	12,1	14,7	20,5	16,6	19,0
Y	53,2	44,2	51,4	45,5	34,0

3. На основе данных (табл. 7.7) построить линейную модель и провести ее анализ.

Таблица 7.7

Количество работников (X)	Объем производства (Y)
7	483
6	489
7	486
8	563
9	570
9	559
9	594
9	575
6	464
9	647

4. Построить модель зависимости величины заработной платы от стажа работы и пола сотрудника (табл. 7.8). Проверить адекватность модели.

Таблица 7.8

Заработная плата, Y	Стаж работы, X_1	Пол (0-муж., 1-жен) X_2
38900	15	0
28700	2	1
31600	4	1
33800	13	0
31890	16	0
45000	35	0
24313	10	1
22700	8	1
36300	20	0
32350	7	0
31800	5	0

5. Определить по данным (табл. 7.9) параметры уравнения линейной регрессии и провести его анализ.

Таблица 7.9

Предприятие	Прибыль, Y	Оборотные средства, X_1	Основные фонды, X_2
1	188	129	510
2	78	64	190
3	93	69	240
4	152	87	470
5	55	47	110
6	161	102	420

Контрольные вопросы

1. Что называется регрессионной моделью?
2. Привести общий вид регрессионной модели.
3. Каким образом можно проверить значимость коэффициента регрессии?
4. Какой метод обычно используется при определении коэффициентов регрессионной модели?

8. ПОСТРОЕНИЕ МАТЕМАТИЧЕСКИХ МОДЕЛЕЙ НА ОСНОВЕ БАЗИСНЫХ ФУНКЦИЙ

Цель работы

Изучение метода наименьших квадратов для построения математических моделей по экспериментальным данным при помощи базисных функций.

Теоретические основы

При обработке экспериментальных (или статистических) данных часто требуется проводить кривые заданного вида, проходящие поблизости от заданных точек.

Пусть имеется набор экспериментальных точек (x_i, y_i), где $i = 1, 2, \dots, m$.

Требуется через данные точки провести кривую $F(x)$, которая является линейной комбинацией заранее выбранных базисных функций $f_j(x)$, $j = 1, 2, \dots, n$:

$$F(x) = \sum_{j=1}^{n} b_j f_j(x).$$

Как правило, $n << m$. Коэффициенты b_j необходимо определить, выбрав определенный критерий для сравнения функций. Используем критерий, называемый методом наименьших квадратов.

Построим матрицу значений базисных функций в заданных точках:

$$A = \begin{pmatrix} f_1(x_1) & f_2(x_1) & \cdots & f_n(x_n) \\ f_1(x_2) & f_2(x_2) & \cdots & f_n(x_n) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(x_m) & f_2(x_m) & \cdots & f_n(x_m) \end{pmatrix}.$$

Матрица A, как правило, не будет квадратной матрицей.

Пусть B – вектор из n искомых коэффициентов. Тогда можно построить вектор из n значений, через которые проходит данная кривая:

$$y^* = AB = \begin{pmatrix} f_1(x_1) & f_2(x_1) & \cdots & f_n(x_n) \\ f_1(x_2) & f_2(x_2) & \cdots & f_n(x_n) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(x_m) & f_2(x_m) & \cdots & f_n(x_m) \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} F(x_1) \\ F(x_2) \\ \vdots \\ F(x_m) \end{pmatrix}.$$

Согласно методу наименьших квадратов коэффициенты вектора *B* определяются минимумом «расстояния» между векторами *y* и *y**, т.е. чтобы было выполнено условие минимума квадратов невязок:

$$\left(\sum_{i=1}^{m} \left(y_i - y_i^*\right)^2 \right) \to \min. \tag{8.1}$$

Рассмотрим общий алгоритм, пригодный для любого выбора базисных функций.

Чтобы найти минимум функции (8.1), необходимо продифференцировать ее по всем переменным b_j и приравнять соответствующие производные нулю. Тогда полученную систему уравнений можно записать в следующем виде:

$$\sum_{i=1}^{n} 2\left(\sum_{j=1}^{n} a_{ij} b_j - y_i \right) a_{ik} = 0,$$

где $k = 1, 2, \ldots, n$.

Эту систему из *n* уравнений можно записать в виде матричного уравнения

$$(AB - Y)^T A = 0,$$

которое эквивалентно уравнениям

$$A^T (AB - Y) = 0 \text{ или } A^T AB = A^T Y. \tag{8.2}$$

Уравнение (8.2) в математической статистике называется нормальным уравнением. Очевидно, что матрица A^TA является симметрической и, согласно теории матриц, если ее столбцы являются линейно независимыми, существует обратная матрица $(A^TA)^{-1}$. Тогда решение системы (8.2) относительно неизвестного вектора *B* является единственным и выражается формулой

$$B = ((A^T A)^{-1} A^T) Y = A^+ Y. \tag{8.3}$$

Матрица $A^+ = (A^TA)^{-1}A^T$ называется псевдообратной матрицей по аналогии с обратной матрицей для систем алгебраических линейных уравнений.

Пример

Требуется построить по точкам (1, 1), (2, 1), (3, 3), (4, 8) следующую зависимость

$$F(x) = b_1 f_1(x) + b_2 f_2(x) + b_3 f_3(x).$$

Анализ графика функции дает возможность предположить, что в качестве базисных функций можно выбрать, например, следующие функции: 1, $x\log_2(x)$, e^x соответственно, т.е. $F(x) = b_1 + b_2 x\log_2(x) + b_3 e^x$.

Матрица А после вычислений значений базисных функций будет иметь следующий вид:

$$A = \begin{vmatrix} f_1(x_1) & f_2(x_1) & f_3(x_1) \\ f_1(x_2) & f_2(x_2) & f_3(x_2) \\ f_1(x_3) & f_2(x_3) & f_3(x_3) \\ f_1(x_4) & f_2(x_4) & f_3(x_4) \end{vmatrix} = \begin{vmatrix} 1 & 0 & 2.7 \\ 1 & 2 & 7.3 \\ 1 & 4.75 & 19.7 \\ 1 & 8 & 53 \end{vmatrix}.$$

Искомые коэффициенты модели можно определить по формуле (8.3).

Программное обеспечение

Для выполнения матричных операций можно воспользоваться специальной программой Mattools или любыми математическими пакетами, например, Mathcad, Maple и др.

Содержание работы

Работа заключается в построении математической модели по экспериментальным данным с помощью базисных функций.

Порядок выполнения

1. По экспериментальным данным (в соответствии с вариантом задания) в виде набора точек (x_i, y_i) нарисовать график функции.
2. На основе анализа графика подобрать вид функции $F(x)$, являющейся линейной комбинацией базисных функций $f_j(x)$ (рекомендуется использовать две базисные функции, не считая единичной).
3. Используя метод наименьших квадратов определить коэффициенты функции $F(x)$.
4. Построить график теоретической функции и проанализировать полученные результаты. В случае если максимальная разница между теоретической кривой и экспериментальными данными составит более 10%, то следует изменить модель добавлением новой базисной функции (или заменой имеющейся базовой функции на новую базовую функцию).

5. Определить коэффициенты новой модели, построить ее график и сравнить точность модельных зависимостей. При необходимости еще раз изменить модель пока расхождения не составят менее 10%.

Отчетность по работе

Отчет должен содержать:

1. Название и цель лабораторной работы.
2. Исходные данные.
3. Основные формулы и результаты вычислений неизвестных коэффициентов.
4. Матрицы базисных функций.
5. Графики экспериментальной и модельных функций.
6. Выводы по работе.

Контрольные вопросы

1. Что называется математической моделью?
2. В чем состоит основная идея метода наименьших квадратов?
3. Что представляют собой базисные функции?
4. Что такое псевдообратная матрица?
5. Перечислить особенности пассивного эксперимента?
6. Что называется регрессионной моделью?
7. Привести общий вид регрессионной модели.
8. Перечислить основные виды базисных функций.
9. Какой метод обычно используется при определении коэффициентов регрессионной модели?
10. Как можно проверить адекватность модели?

Варианты заданий

1	x	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0
	y	0.4	0.3	1.0	1.7	2.1	3.4	4.1	5.8
2	x	4.5	5.0	5.5	6.0	6.5	7.0	7.5	8.0
	y	7.7	9.4	11.4	13.6	15.6	18.6	21.2	24.1
3	x	0.4	0.8	1.2	1.6	2.0	2.4	2.8	3.2
	y	0.43	0.94	1.91	3.01	4.0	4.56	6.45	8.59
4	x	4.0	4.4	4.8	5.2	5.6	6.0	6.4	6.8
	y	13.88	16.93	20.47	24.15	28.29	32.61	37.41	42.39
5	x	1	2	3	4	5	6		

	y	2.11	2.45	2.61	2.73	2.75	2.81		
6	x	0.3	0.6	0.9	1.2	1.5	1.8		
	y	4.39	4.75	4.98	5.11	5.12	5.18		
7	x	1	2	3	4	5	6		
	y	0.1	0.21	0.43	0.51	0.62	0.81		
8	x	1.0	1.5	2.0	2.5	3.0	3.5	4.0	
	y	4.11	4.16	4.23	4.29	4.36	4.42	4.53	
9	x	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0
	y	2.47	2.86	3.01	2.91	2.55	2.11	2.61	1.25
10	x	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
	y	0.9	0.8	1.5	2.2	2.6	3.9	4.6	6.3
11	x	4.0	4.5	5.0	5.5	6.0	6.5	7.0	7.5
	y	7.2	9.0	11.0	13.1	15.2	18.0	20.5	23.6
12	x	0.5	0.78	1.3	1.71	2.1	2.5	2.9	3.33
	y	0.55	0.99	2.0	3.1	4.15	4.6	6.5	8.7
13	x	4.2	4.6	5.0	5.4	5.8	6.2	6.62	6.98
	y	14.0	17.3	20.6	24.4	28.5	32.8	37.6	42.6
14	x	1	2	3	4	5	6		
	y	3.1	3.5	3.6	3.7	3.5	3.85		
15	x	0.5	0.8	1.1	1.4	1.7	2.1		
	y	5.4	5.7	5.9	6.1	6.1	6.2		
16	x	1	2	3	4	5	6		
	y	0.2	0.31	0.55	0.62	0.72	0.93		
17	x	1.0	1.5	2.0	2.5	3.0	3.5	4.0	
	y	4.1	4.2	4.2	4.32	4.41	4.4	4.5	

9. ИДЕНТИФИКАЦИИ ДИНАМИЧЕСКИХ СИСТЕМ УПРАВЛЕНИЯ В ИНТЕРАКТИВНЫХ ИНСТРУМЕНТАХ MATLAB

Цель работы

1. Организация экспериментальных данных и их препроцессорная обработка.
2. Выбор структуры модели.
3. Определение параметров модели.
4. Оценка полученных результатов и выводы.

Теоретические основы работы

Задача ***идентификации*** является неотъемлемой частью синтеза и анализа систем управления. От качества идентификации может зависеть успех реализации целевой системы, ее работа в штатных, критических и оптимальных режимах эксплуатации. Многообразие форм представления систем, алгоритмов идентификации и методов проверки модели на адекватность ставит перед инженерами и исследователями нетривиальные и творческие задачи, решение которых требует обширных знаний и умений.

Рассмотрим:

– основные концепции, лежащие в идентификации моделей с использованием измеренных данных;

– как оценить передаточные функции, уравнения пространства состояний и другие модели динамических систем с помощью графического интерфейса;

– как сравнивать модели между собой, проверять их качество и устойчивость к возмущениям;

– как использовать полученные модели для анализа и синтеза систем управления.

Система представляет собой реальную физическую систему, например, ракету-носитель, где входами являются сигналы, например, сигналы от рулей, тяги, а выходами могут быть значения скорости, ускорения, угла атаки (рис. 9.1).

Реальную систему заменяем моделью – уравнением, которое ставит в соответствие выходы (Y) входам (U), которые можно измерить и из которых можно получить уравнения «вход-выход» или параметры этих уравнений (рис. 9.2).

Рис. 9.1. Пример физической системы

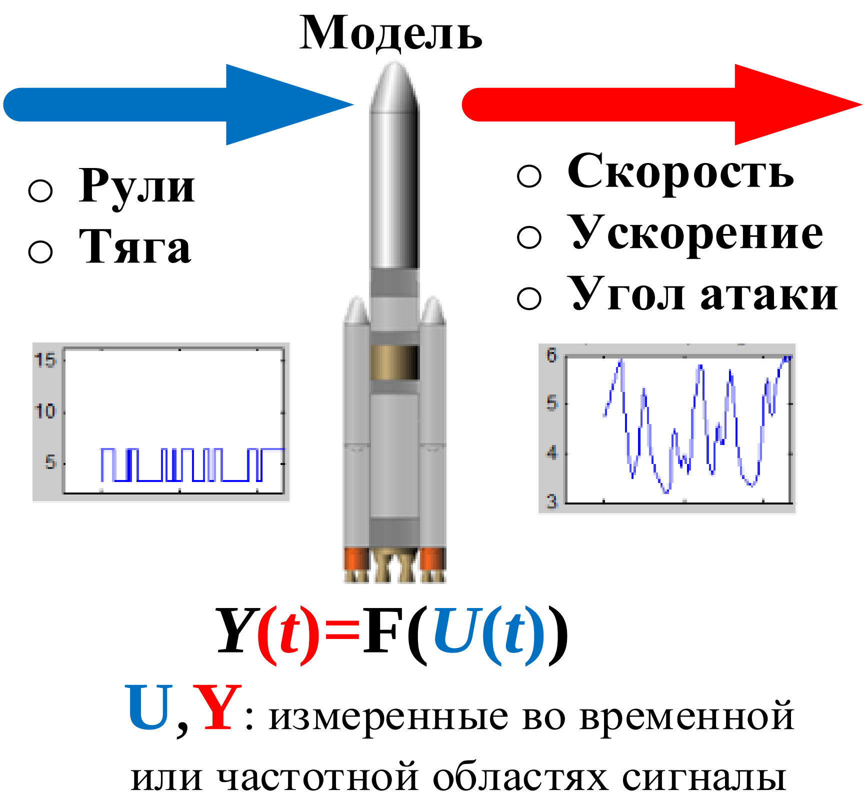


Рис. 9.2. Пример модели физической системы

Подход при создании модели заключается в воссоздании отклика модели при одинаковых условиях и подачи входного воздействия на систему и сравнения смоделированного сигнала и измеренного. В процессе определяется ошибка моделирования и происходит настройка (калибровка) параметров модели, которые минимизируют данную ошибку (рис. 9.3).

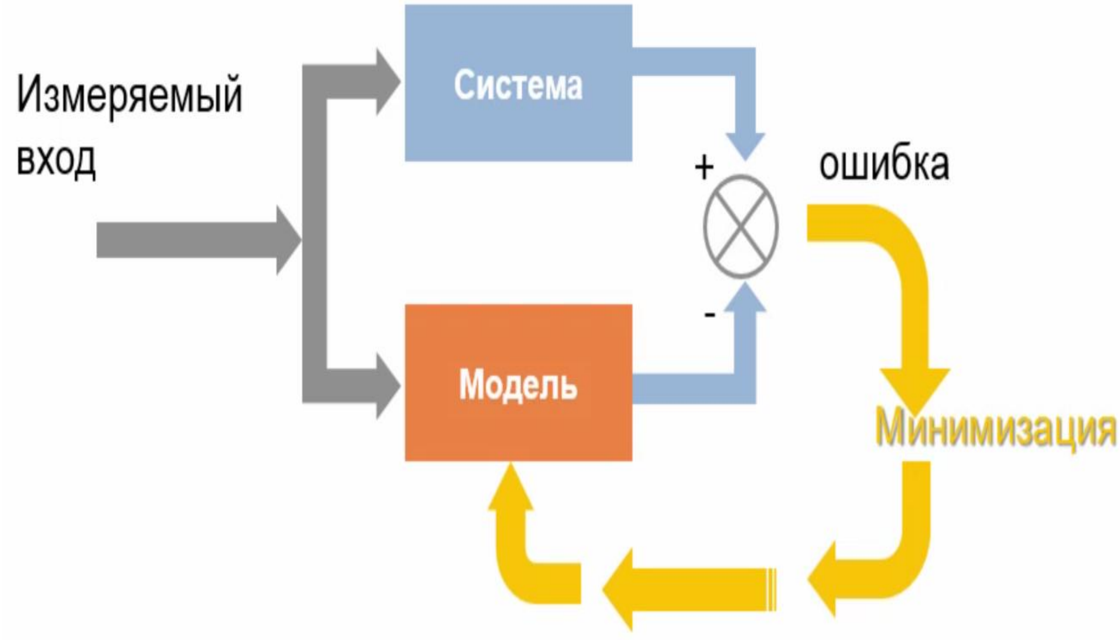


Рис. 9.3. Подход при создании модели

Процесс идентификации (построения) модели на основе данных можно представить следующими этапами (рис. 9.4 и 9.5):

1. Получение экспериментальных данных (сбор данных «вход-выход»). Данные могут быть представлены в различных формах: во временной или частотной областях. Интервал дискретизации должен позволять наблюдать динамику процесса. Если необходимо, то следует осуществить фильтрацию – удалить погрешности и шумы измерений («выбросы»).

2. Построение модели:

– выбор (или, возможно, синтез) структуры модели. Качество результата напрямую зависит от выбора начальной структуры. Знания физики объекта или процесса, математики, принципов работы, информация об аналогах (т.е. априорные знания) помогут в выборе начальной структуры модели. Могут использоваться модели типа «черный» (гибкая структура с возможностью настраивать аппроксимацию данных) и «серый» (структуры моделей на основе дифференциальных уравнений с неизвестными параметрами) ящик;

– определение (подбор и настройка) параметров модели для выбранной структуры. Данная задача может быть сформулирована в виде оптимизационной задачи.

3. Валидация на независимом наборе данных. Анализ ошибок и оценка модели.

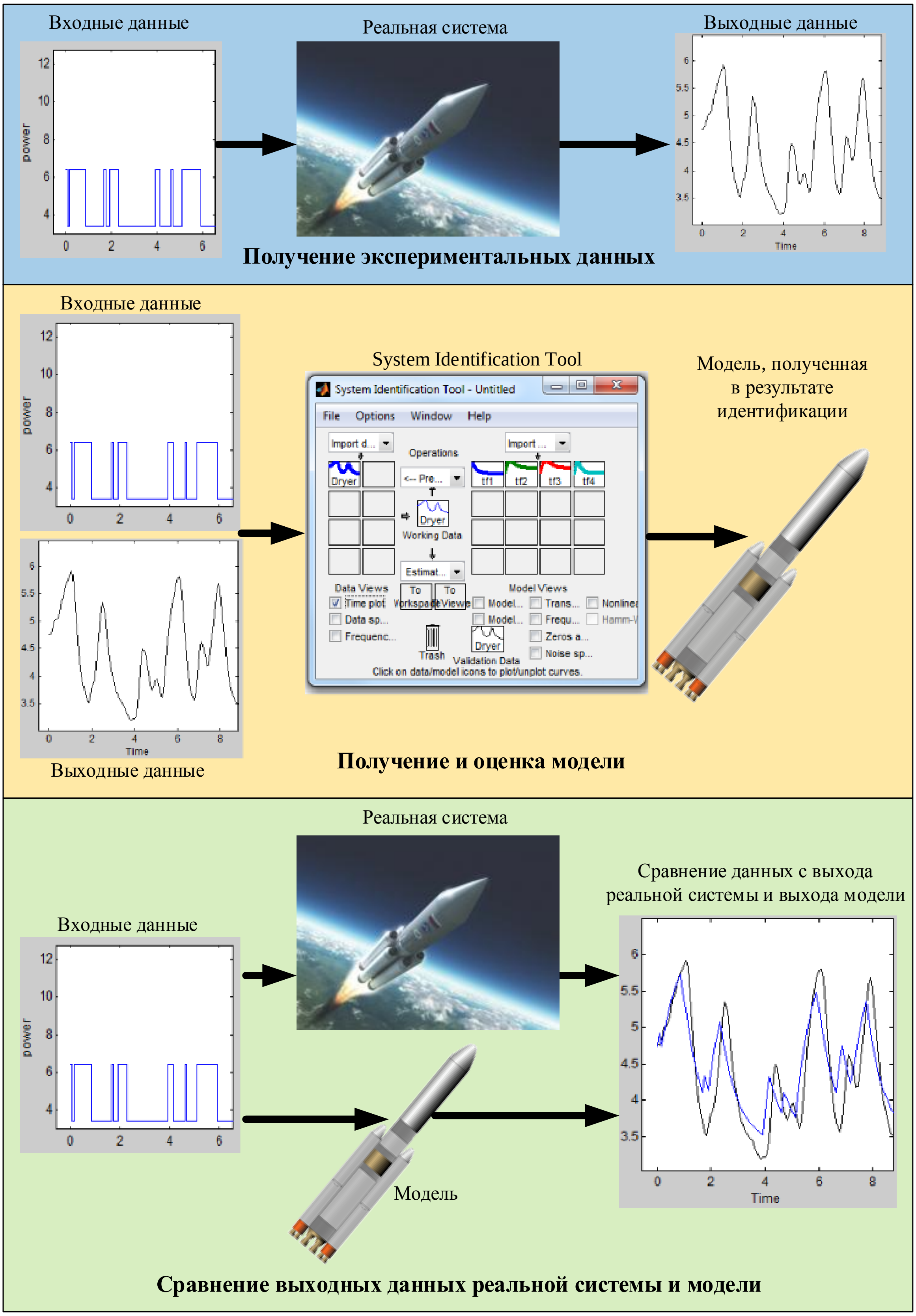


Рис. 9.4. Этапы идентификации

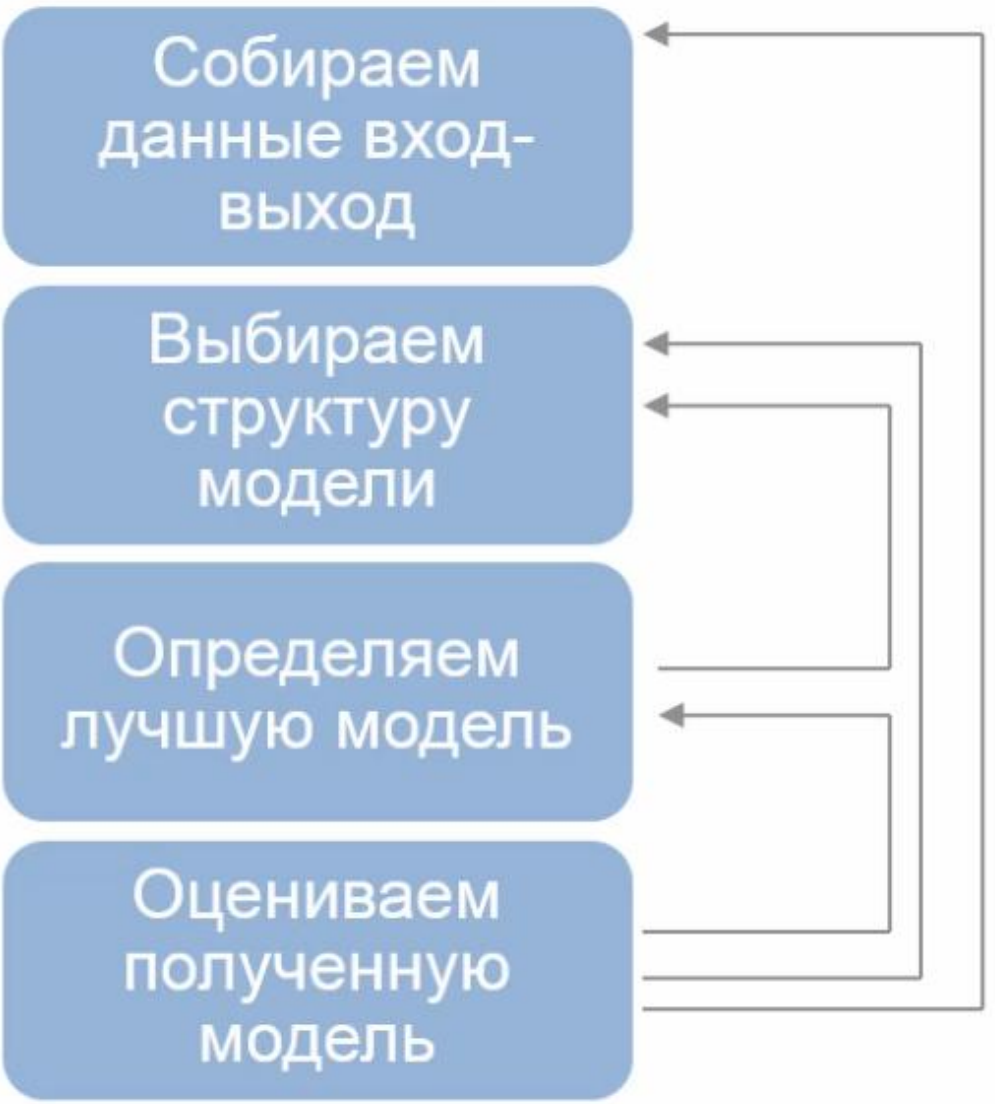


Рис. 9.5. Итерационный процесс идентификации

Ключевое свойство инструмента идентификации заключается в том, что необходимо производить оценку параметров модели объекта с тем, чтобы отклик модели совпадал с измерениями выходов объекта. Можно оценивать качество модели с помощью функции **COMPARE** (сравнение), которая позволяет:

– строить ***совместный график*** измерений и выходов модели;

– вычислять ***количественную оценку*** идентификации модели: процент отклонения отклика модели от измерений.

На рис. 9.6 выход модели обозначен как **«model»**, а экспериментальные данные – **«Validation data»**.

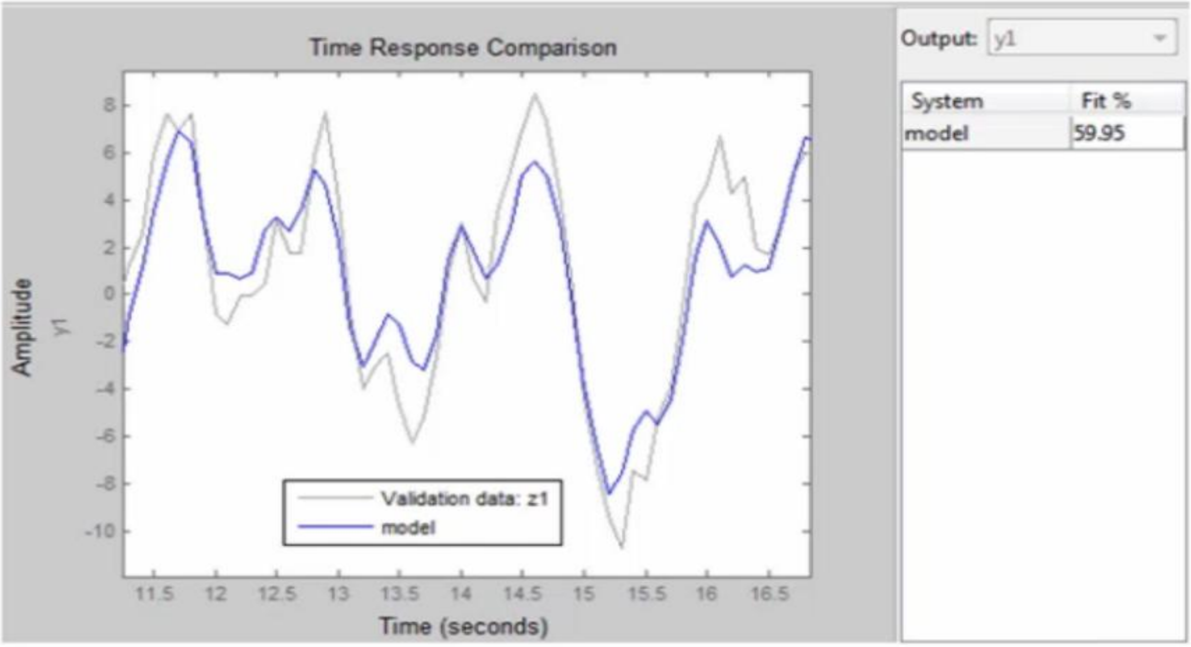


Рис. 9.6. Пример работы функции ***COMPARE***

Совместна оценка и валидация

Очень большой уровень совпадения выходных данных из модели и экспериментальных данных, которые использовались при идентификации, не означает, что получена требуемая модель. Слишком сложная модель будет точно воспроизводить данные («подстроится», «подгонится» под идентификационную (обучающую) выборку), но может потерять способность обобщать и прогнозировать. Это явление называется ***переобучением (overfitting)***. Следовательно, необходимо проверить релевантность модели на независимом наборе данных. «Настоящее» испытание для модели будет, когда в модель будут поданы данные, не участвовавшие в идентификации, т.е. будет проведена ***валидация*** модели. Как видно из графика (рис. 9.7) может наступить такой момент, начиная с которого усложнение модели приведет лишь к повышению качества (точности) подстройки модели по идентификационным данным, но для валидационных данных (не участвовавших в идентификации) качество модели будет резко ухудшаться. Это и будет тем указателем качества идентификации.

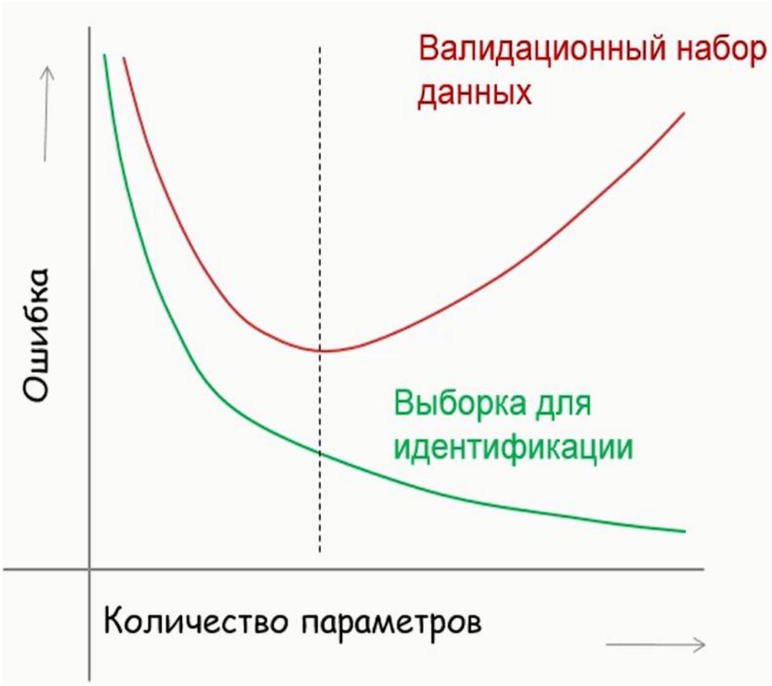


Рис. 9.7. Зависимость ошибки модели от её сложности

Для оценки качества идентификации необходимо также использовать анализ невязки. Под невязкой понимается ошибка моделирования на один шаг вперед – ***ошибка прогнозирования***, т.е. разница между измеренным значением и полученным из модели на основании предыдущих значений. Приемлемой ошибкой моделирования является ошибка, не имеющая корреляции с другими известными значениями, а именно, не должно быть связи с входным воздействием (для этого используются ***кросс-корреляционные функции***), и не имеющая корреляции с собственными значениями ошибки на других интервалах времени (***автокорреляционные функции***).

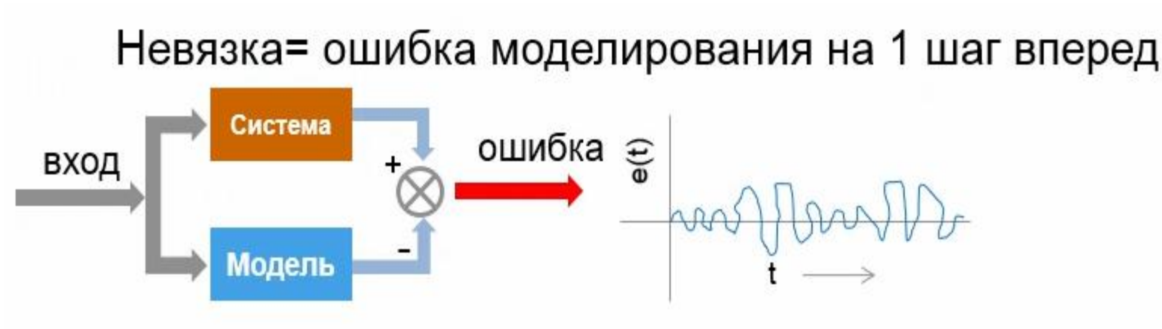


Рис. 9.8. Схема получения оценки невязки

Программное обеспечение

В работе используется программный пакет **Matlab** с **System Identification Toolbox**, которые позволяют осуществлять построение математических моделей динамических систем по измеренным данным входа и выхода реальной системы.

Содержание работы

1. Разделить выборку, выданную преподавателем в файле Excel, на две произвольные части. Для этого скопировать часть данных из нижней части массива и сохранить в другом файле Excel. В первом же файле удалить скопированные данные, при этом записав порядковый номер, начиная с которого удалены данные. Дать имена файлам **«Обучающая выборка»** и **«Валидационный набор данных»**.
2. Загрузить файл Excel **«Обучающая выборка»** в рабочую область Matlab. Для этого необходимо нажать кнопку **«Import Data»** и выбрать местоположение и сам файл с расширением xlxs (**«*.xlsx»**) (рис. 9.9).

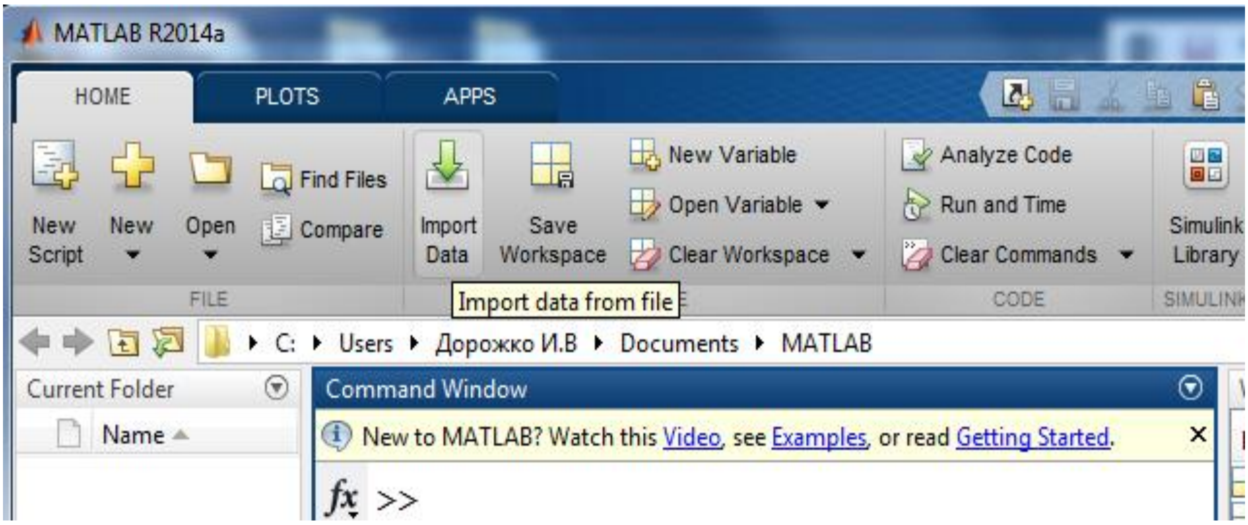


Рис. 9.9. Окно Matlab и кнопка «Import Data»

В появившемся окне нажать кнопку **«Import Selection»**. В рабочей области Matlab появятся два массива, которые можно просмотреть, дважды кликнув на их имена. Переименовать массивы данных на **«*u*1»** и **«*y*1»**.

3. Чтобы запустить пользовательский графический интерфейс **System Identification Toolbox**, нужно в командной строке Matlab ввести команду **«ident»** (рис. 9.10 и 9.11).

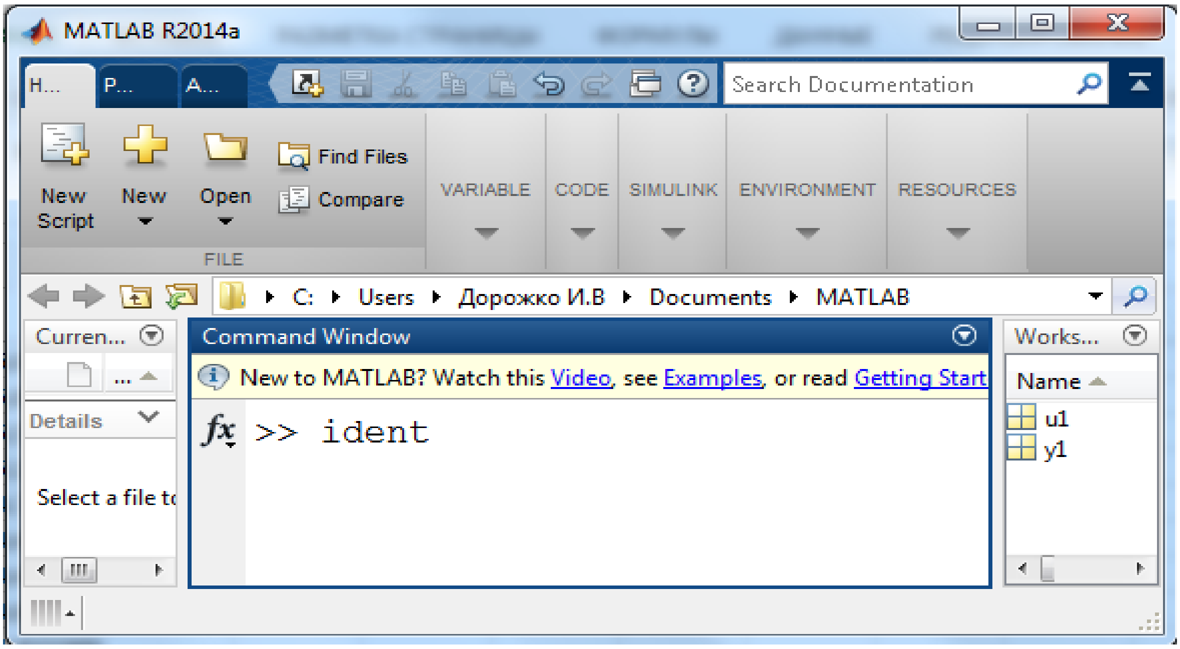


Рис. 9.10. Вид командного окна Matlab

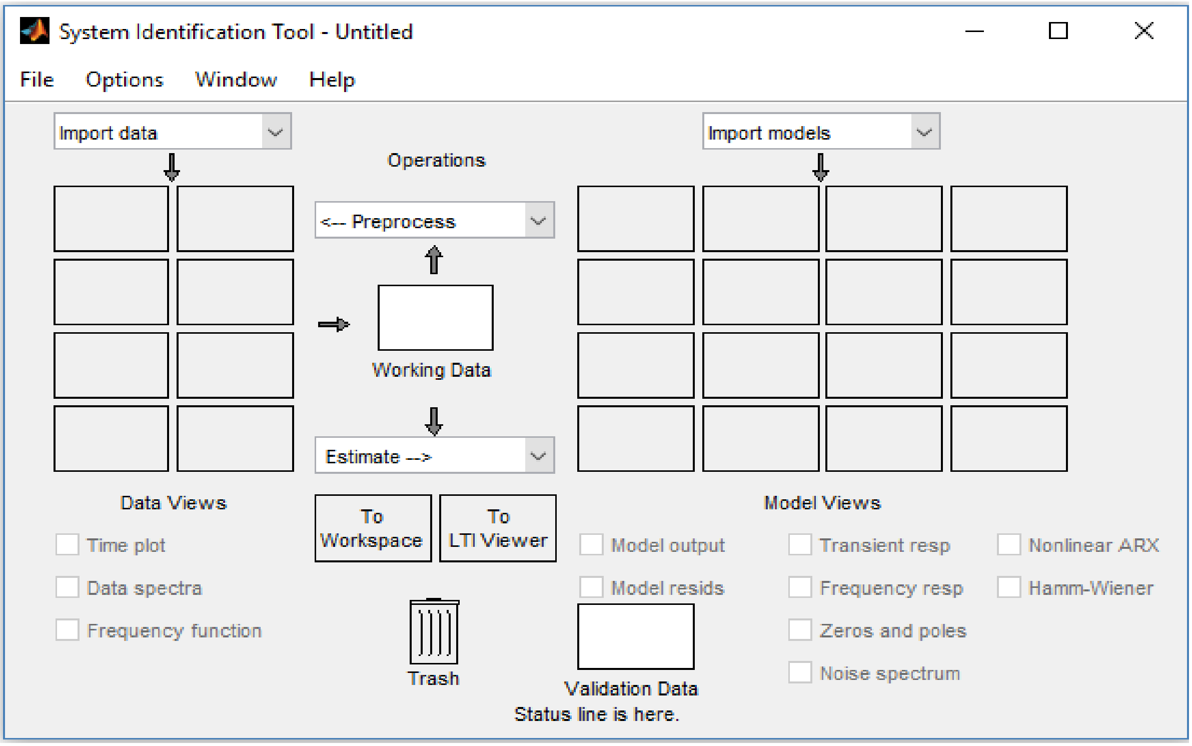


Рис. 9.11. Вид окна System Identification Toolbox

Интерфейс **System Identification Toolbox** содержит три области: область для работы с данными (**«Data Views»**), область выбора операций (**«Operations»**) и область для работы с моделью (**«Model Views»).**

4. Импортируем данные: в списке **«Import data»** выберем **«Time domain data…»** (рис. 9.12).

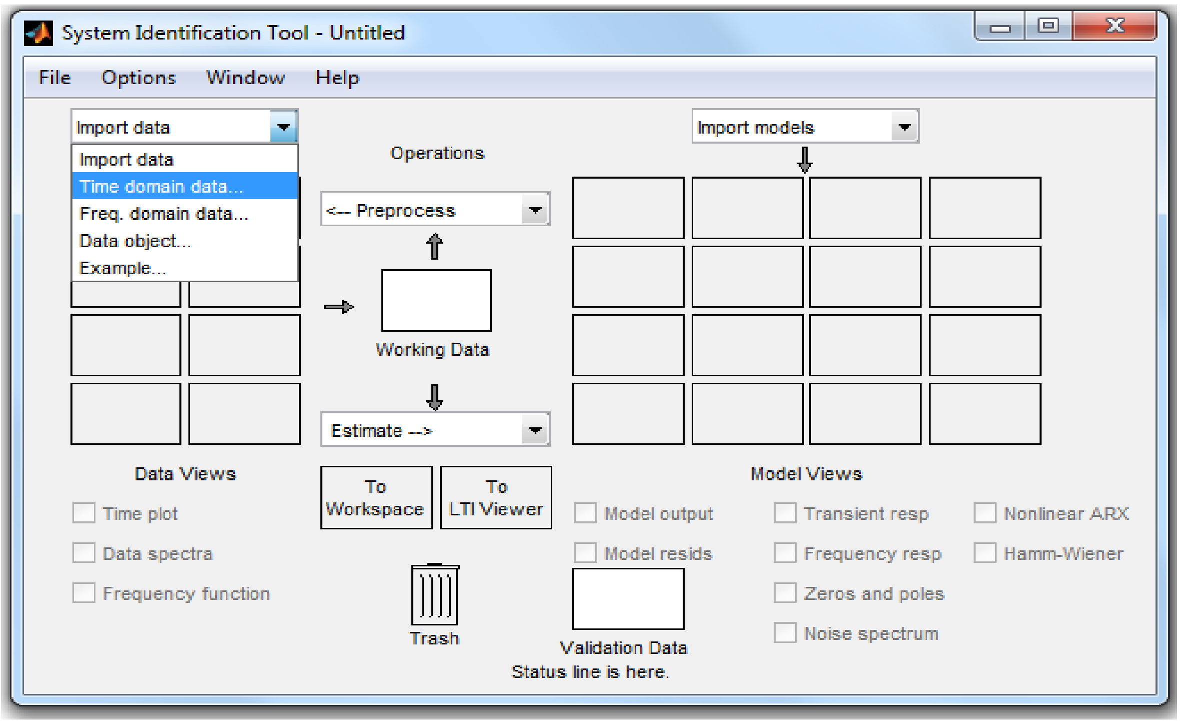


Рис. 9.12. Импорт данных

В поле **«Input»** вводим имя входной переменной ***«u1»***, в поле **«Output»** – ***«y1»***, в поле **«Starting time»** – **«0»**, в поле **«Sampling interval»** – **«1»** (рис. 1.13). Нажать кнопку **«Import»**.

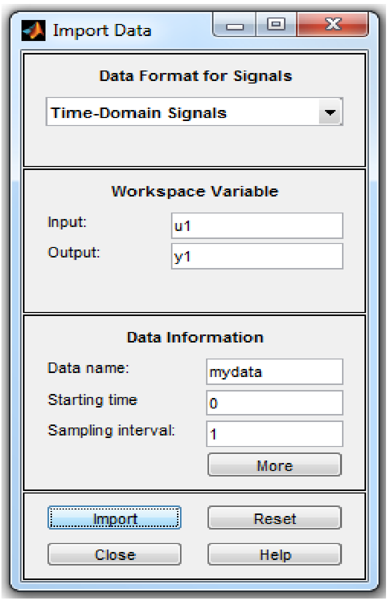


Рис. 9.13. Параметры импортируемых данных

Данные появятся в окне интерфейса **System Identification Toolbox**.

5. Для отображения содержимого импортированных данных в виде графика ставим галочку в поле **«Time plot»** (рис. 9.14).

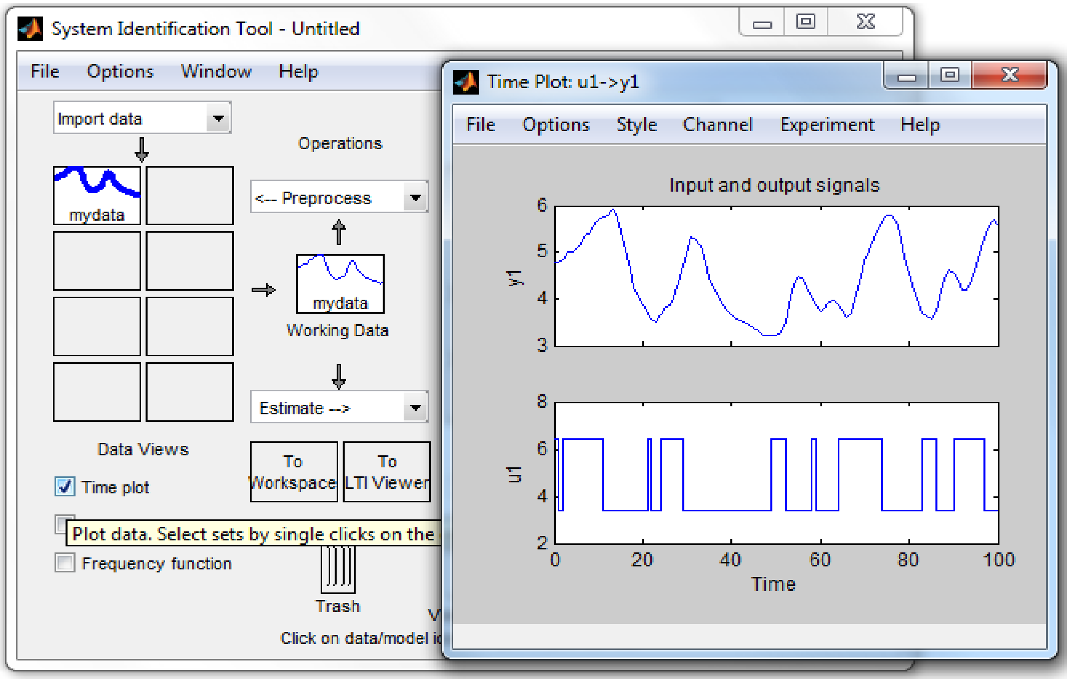


Рис. 9.14. Отображение импортированных данных в виде графика

Для отображения частотных характеристик необходимо поставить галочку в поле **«Data spectra»** (рис. 9.15). Изобразить полученные графики и периодограммы в отчетах.

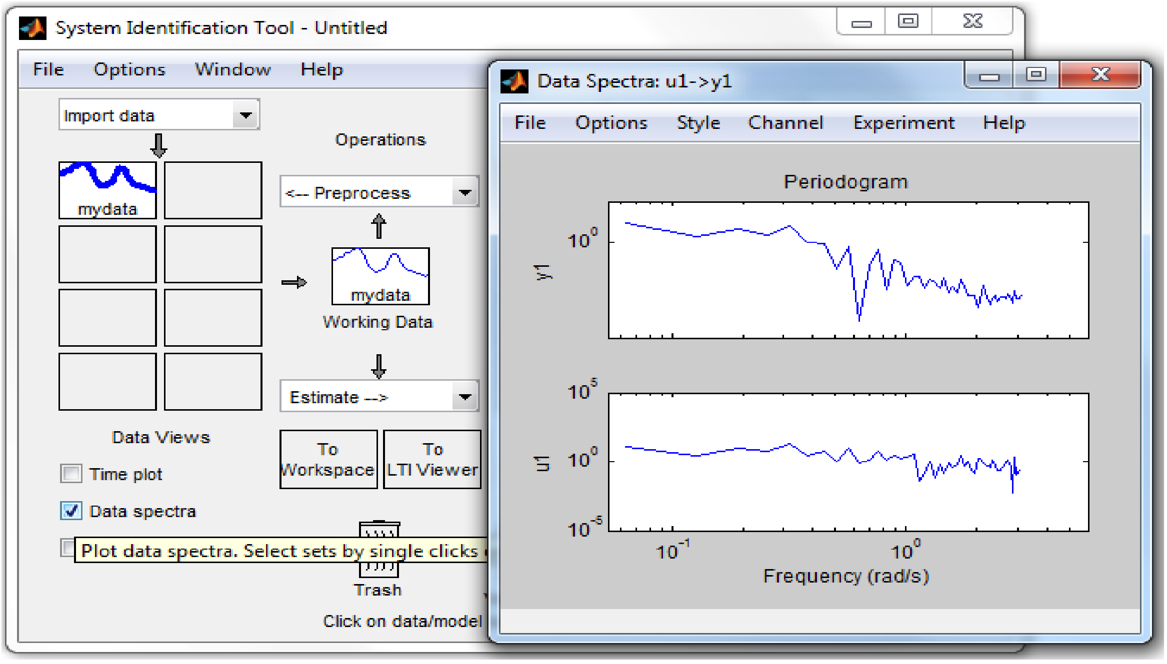


Рис. 9.15. Отображение импортированных данных в виде периодограммы

6. Далее данные передаем в функцию оценки модели.

Интерфейс **System Identification Toolbox** поддерживает следующие структуры моделей:

1) Линейные параметрические модели:

– модели вход-выход (передаточные функции): $Y(z) = \dfrac{B(z)}{F(z)} U(z)$;

– модели пространства состояний: $x(t+1) = A \cdot x(t) + B \cdot u(t)$, $y(t) = C \cdot x(t)$.

2) Линейные непараметрические модели:

– модели импульсного отклика;

– модели частотного отклика.

3) Модели процессов (достаточно часто применяются для описания технологических процессов). Используется комбинация усиления (K), задержки (T_d) и постоянных времени (T) для описания модели:

$$Y(s) = K \cdot \frac{e^{-s \cdot T_d}}{1 + s \cdot T} \cdot U(s).$$

4) Нелинейные модели.

Воспользуемся структурой модели в виде ***передаточной функции*** (рис. 9.16).

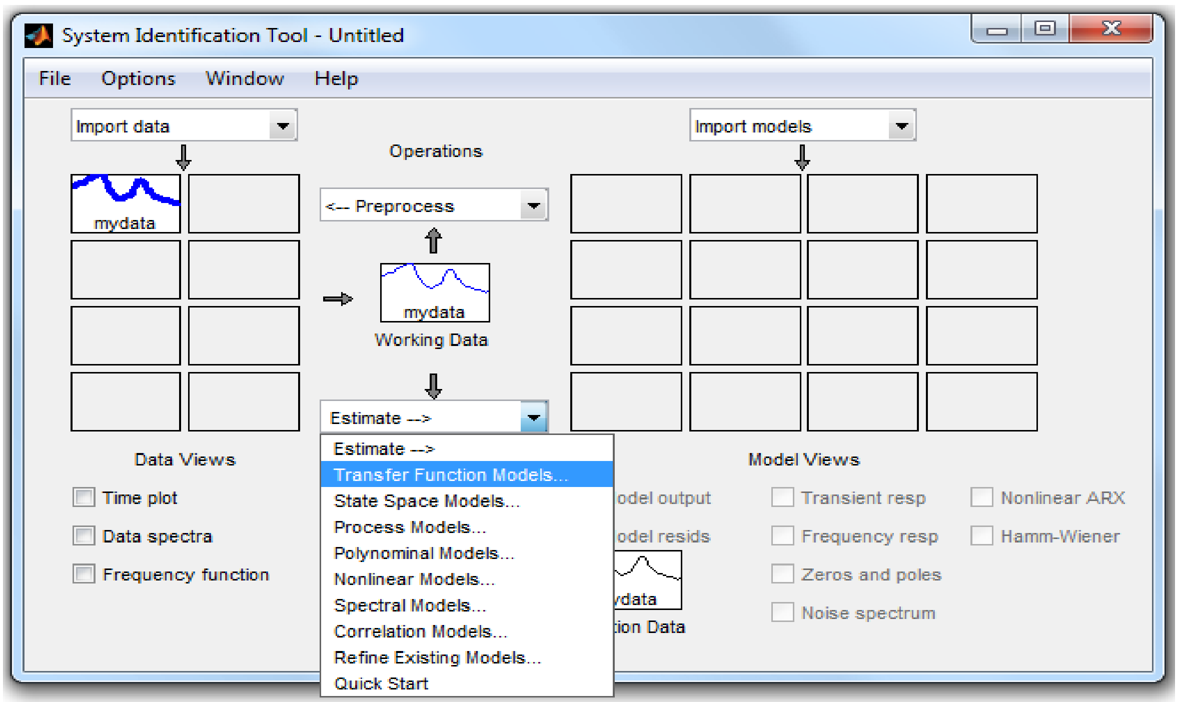


Рис. 9.16. Окно выбора модели

Выбираем порядок числителя и знаменателя: «**Number of poles**»=1, «**Number of zeros**»=0, а также непрерывное или дискретное время (рис. 9.17). Нажимаем на кнопку «**Estimate**».

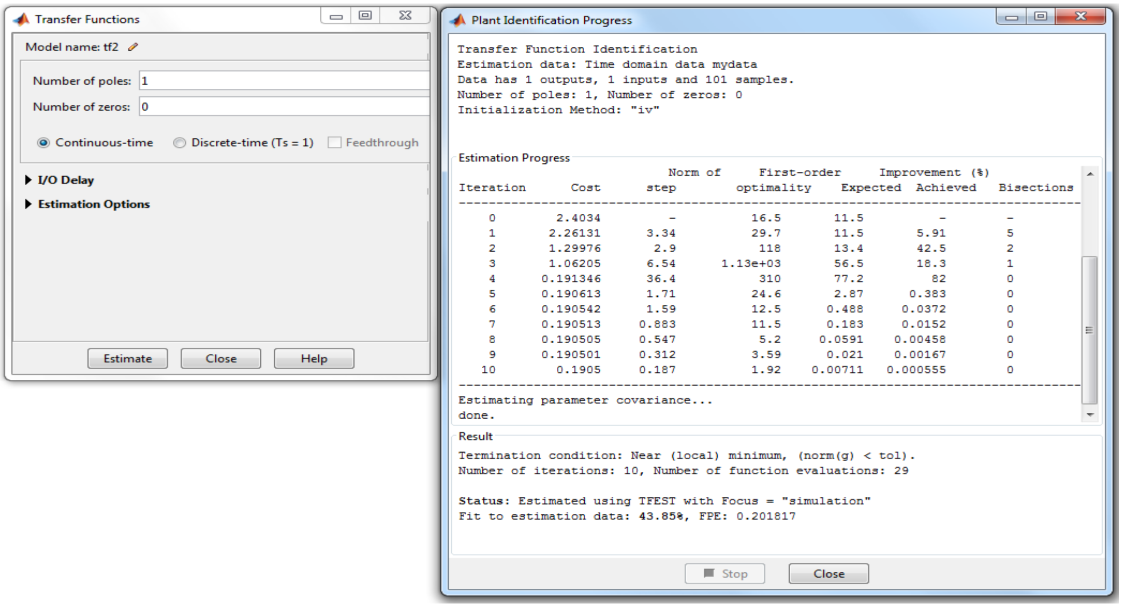


Рис. 9.17. Ввод порядка модели

Далее полученная модель исследуется. Просматриваем параметры модели, кликнув два раза на графике модели **tf1** (рис. 9.18). Передаточная функция имеет следующий вид:

$$W(s) = \frac{1.1072}{s + 0.1116}.$$

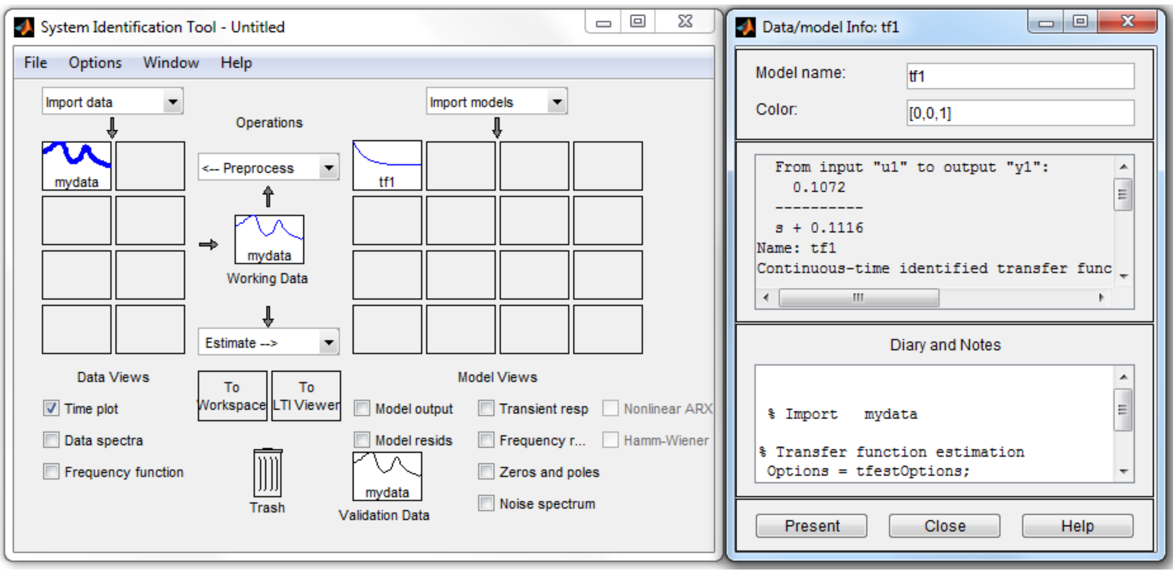


Рис. 9.18. Окно просмотра параметров модели

Записать полученную передаточную функцию в отчетах.

А также определяем насколько полученная модель точно описывает экспериментальные данные. Поставим галочку в поле **«Model output»** (рис. 9.19).

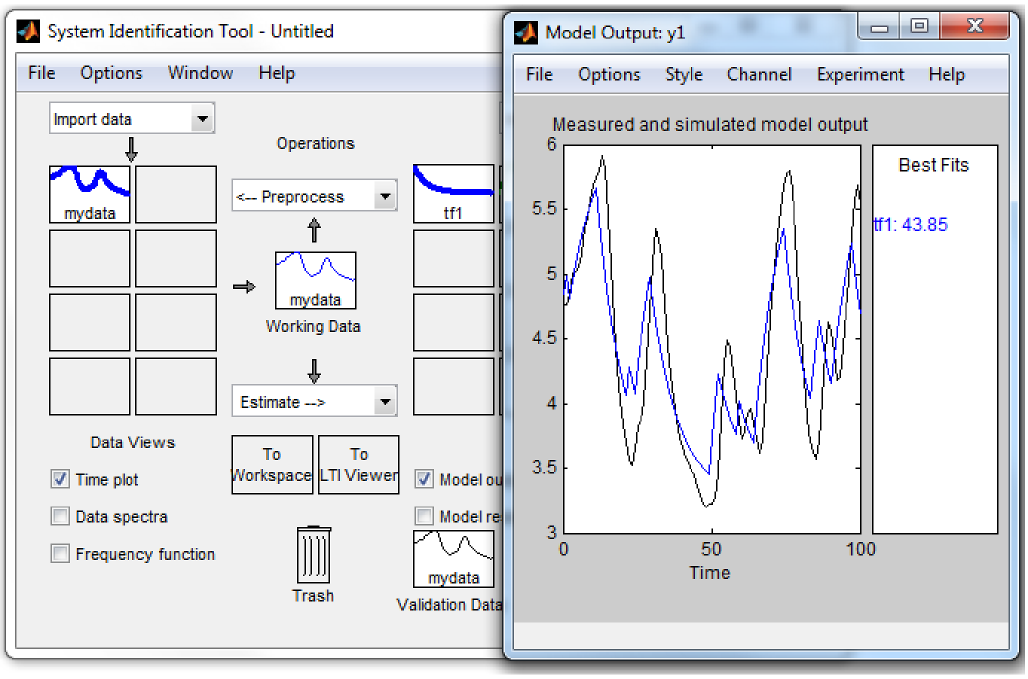


Рис. 9.19. Окно просмотра точности модели

Из анализа рис. 9.19 можно сделать вывод, что модель достаточно неточно описывает экспериментальные данные: процент совпадений составляет 43.85%. В отчетах изобразить графики и значение точности модели.

Усложним модель. В окне, изображенном на рис. 9.17 выбираем более высокий порядок числителя и знаменателя, например, **«Number of poles»**=2, **«Number of zeros»**=1.

Как видно из рис. 9.20 полученная модель **tf2** гораздо лучше описывает реальные экспериментальные данные: процент совпадений составляет 85.55%. Передаточная функция имеет следующий вид:

$$W(s) = \frac{-0.1144 \cdot s + 0.09706}{s^2 + 0.5509 \cdot s + 0.1014}.$$

Записать в отчетах полученные передаточную функцию, значение точности и изобразить графики.

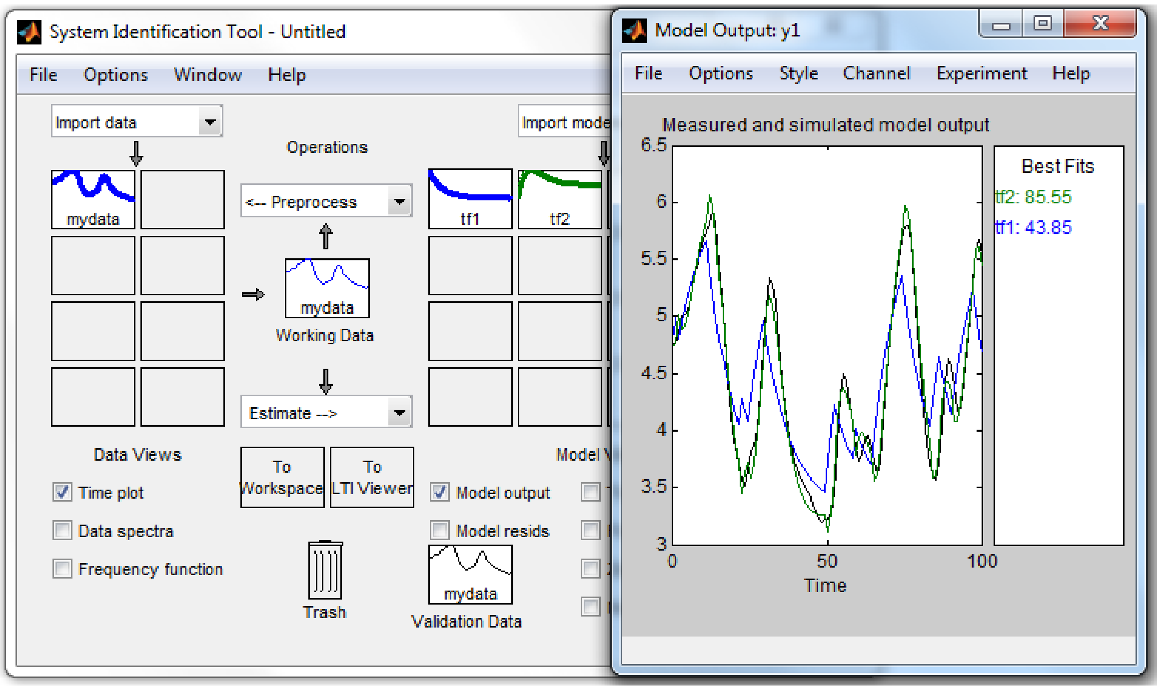


Рис. 9.20. Окно просмотра точности модели

Продолжим усложнение модели. Выбираем более высокие порядки числителя и знаменателя (**«Number of poles», «Number of zeros»**) (рис. 9.21).

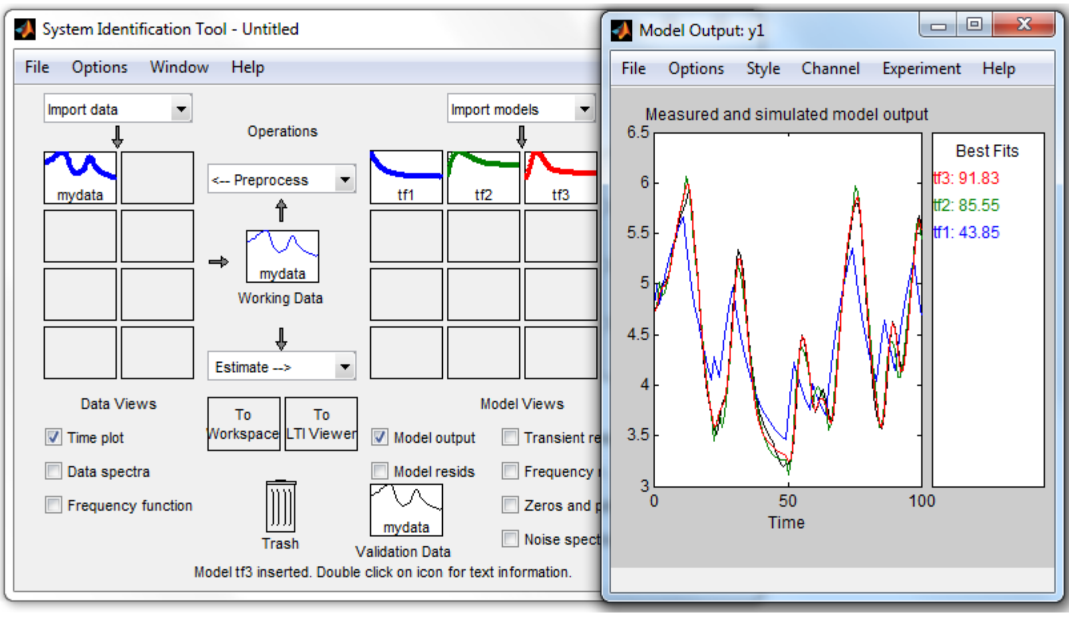


Рис. 9.21. Окно просмотра точности модели

Из рис. 9.22 видно, что полученная в результате идентификации модель **tf3** еще лучше описывает реальные экспериментальные данные: процент совпадений составляет 91.62%. Записать в отчетах полученные передаточные функции, значение точности и изобразить графики, проделав данную операцию несколько раз. Сделать вывод о возможности достижения 100% совпадений.

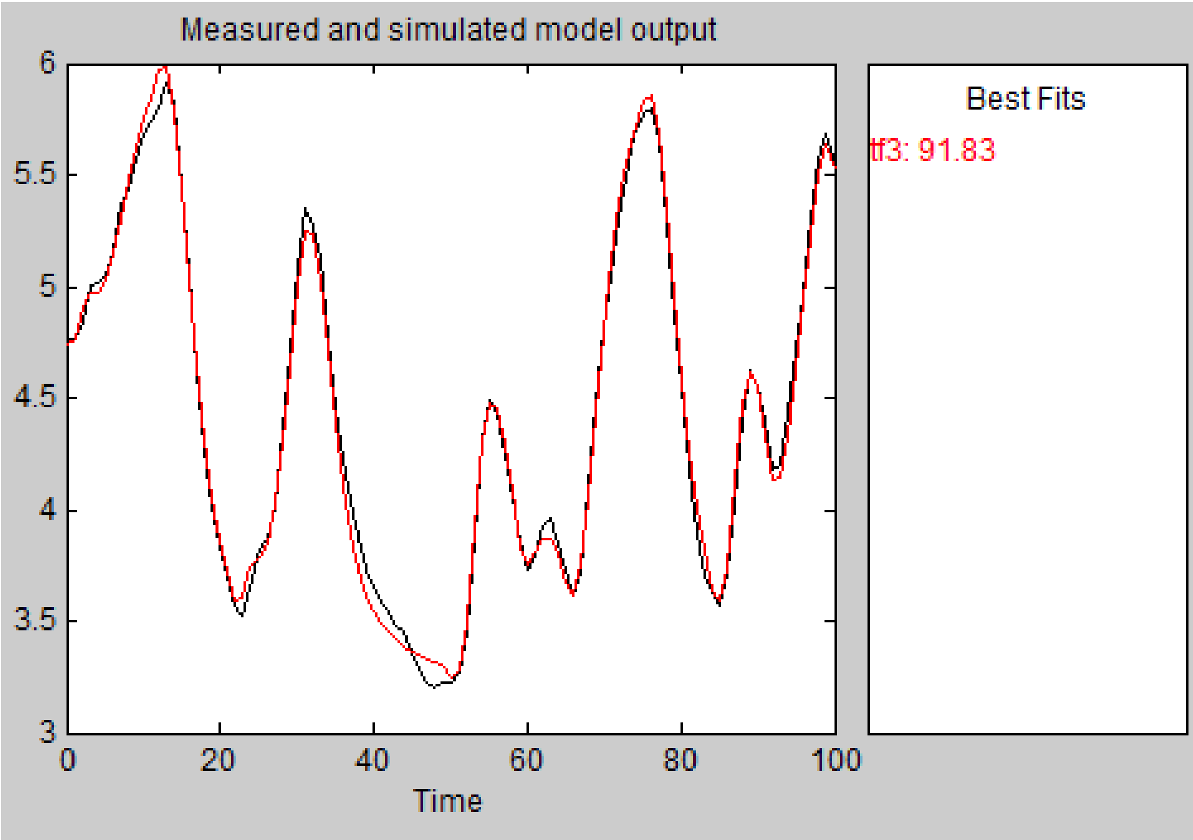


Рис. 9.22. Окно просмотра точности модели

Воспользуемся более сложным классом моделей – ***нелинейные авторегрессионные модели***, выбрав **Estimate-Nonlinear Models** (рис. 9.23). В надстройке **System Identification Toolbox** можно исследовать нелинейные динамические модели:

– авторегрессионная модель с внешним входом (Nonlinear ARX);
– модель Хаммерстайна-Винера (Hammerstein-Weiner);
– модель с вейвлетной сетью;
– модель с древовидным разделением;
– модель с нелинейной сигмоидальной сетью.

Выберем число входных параметров **«Input Channels»**, например, равное 10 и число выходных параметров **«Output Channels»**=10, а также во вкладке **«Model Properties»** вид нелинейности выбираем, например, **«Nonlinearity»=Sigmoid Network** и «**Number of units in nonlinear block»**=10 (рис. 9.24).

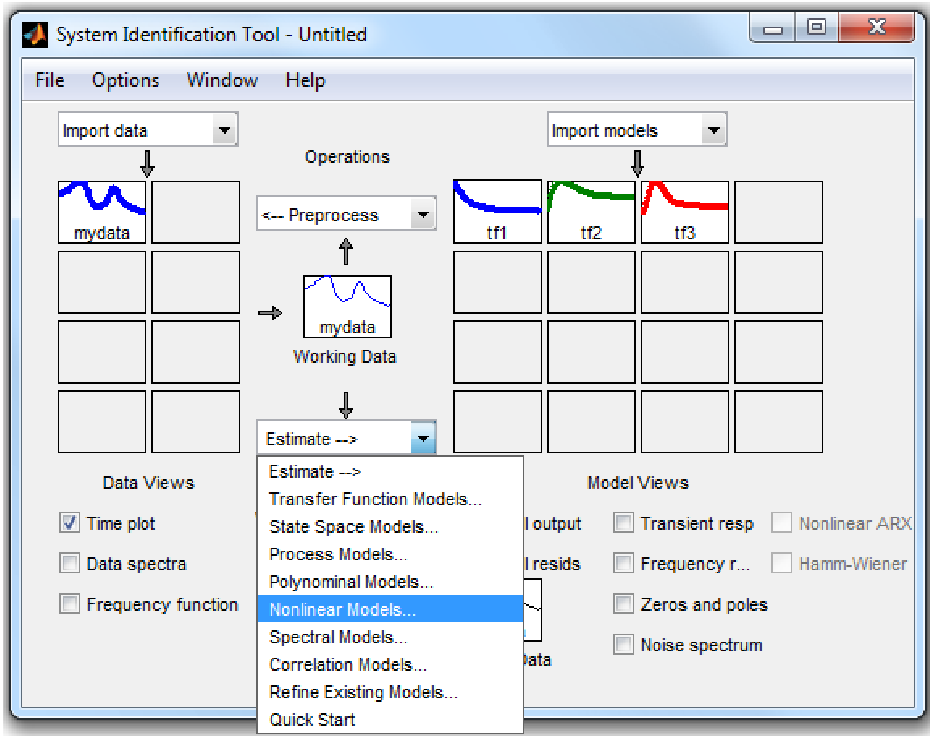


Рис. 9.23. Окно выбора структуры «Нелинейные модели»

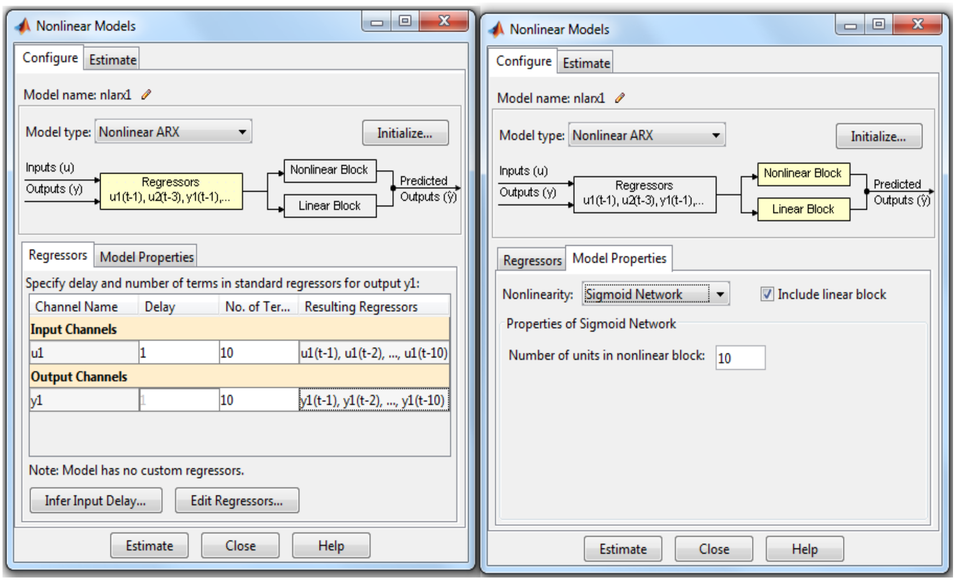


Рис. 9.24. Окно «Nonlinear Models» («Нелинейные модели»)

Нажимаем на кнопку **«Estimate»**. Процент совпадений данной нелинейной модели с реальными экспериментальными данными составляет 100.00% (рис. 9.25).

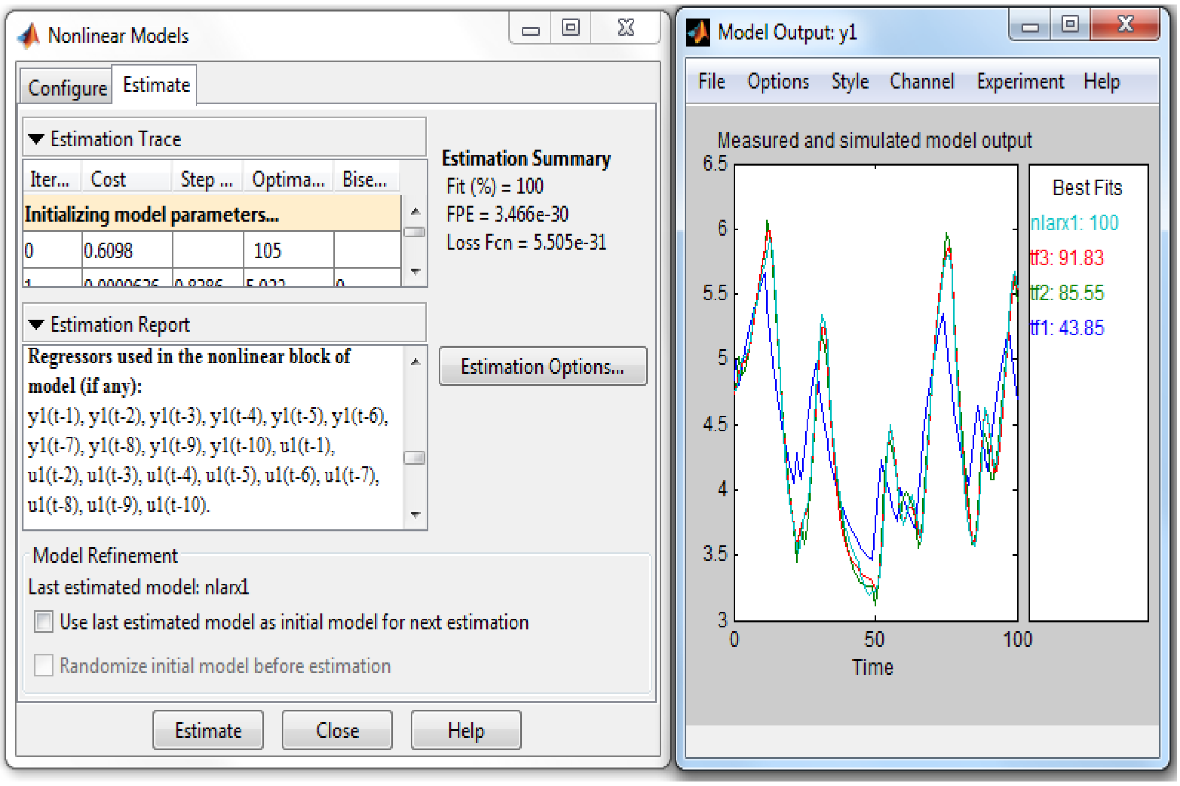


Рис. 9.25. Окно просмотра точности модели

Записать в отчет вид модели, значение точности и изобразить графики.

Исследовать другие виды моделей, например, полиномиальные с отображением результатов в отчете.

7. Осуществим ***валидацию*** полученных моделей. Загрузим, как указано в п. 2, файл **«Валидационный набор данных»**, содержащий данные не участвовавшие в обучении. Нажать кнопку **«Import Selection»**. В рабочей области Matlab появятся еще два массива, которые можно также просмотреть, дважды кликнув на их имена. Переименовать массивы данных на **«*u*2»** и **«*y*2»**.

Далее выбрать в окне **System Identification Toolbox** список **«Import data» – «Data object…»**. В поле **«Input»** вводим имя входной переменной **«*u2*»**, в поле **«Output»** – **«*y2*»**, в поле **«Data name»** – **«valdata»**, в поле **«Starting time»** указываем порядковый номер, начиная с которого были удалены данные в п.1, в поле **«Sampling interval»** – **«1»**. Нажать кнопку **«Import»**.

Данные появятся в окне интерфейса **System Identification Toolbox**. Совместные графики идентификационной и валидационной выборок, а также отдельно лишь валидационной выборки изображены на рис. 9.26. Изобразить графики в отчете.

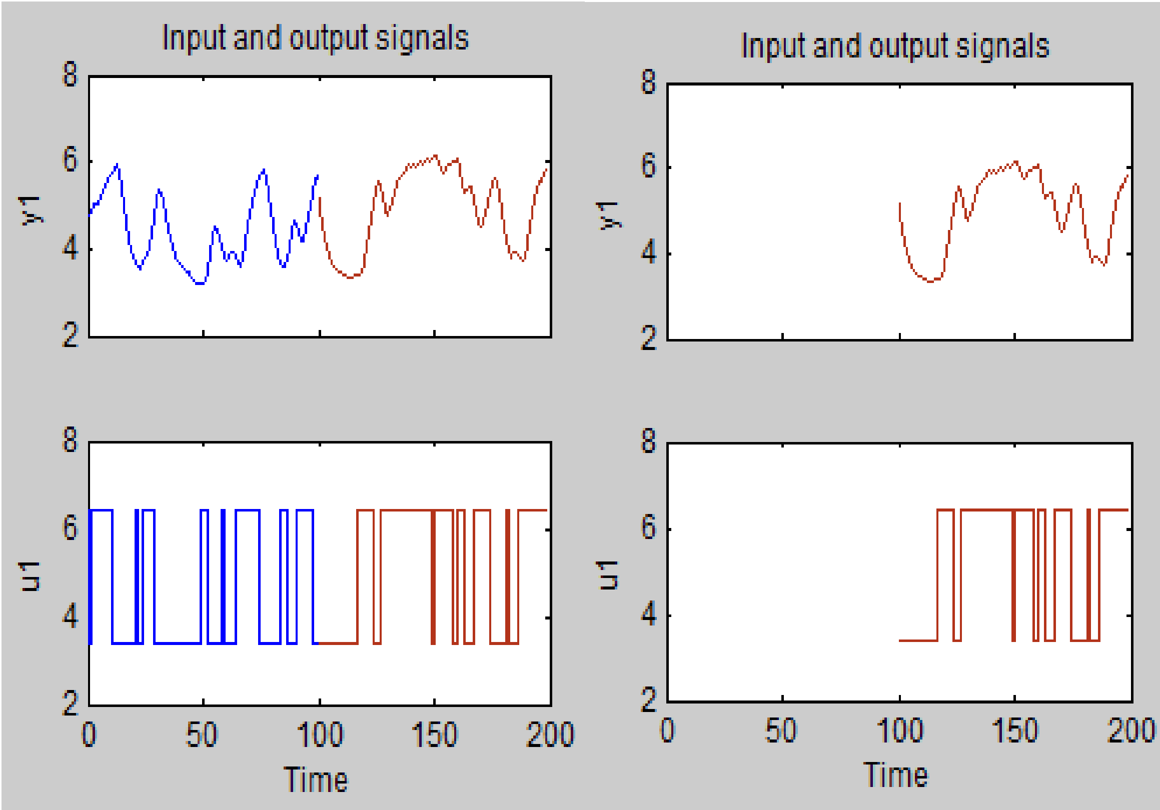


Рис. 9.26. Графики идентификационного и валидационного наборов

Сравним полученные модели, используя валидационный набор, для этого переместим данный набор в область **«Validation Data»**. Посмотрим как изменится выход модели (рис. 9.27).

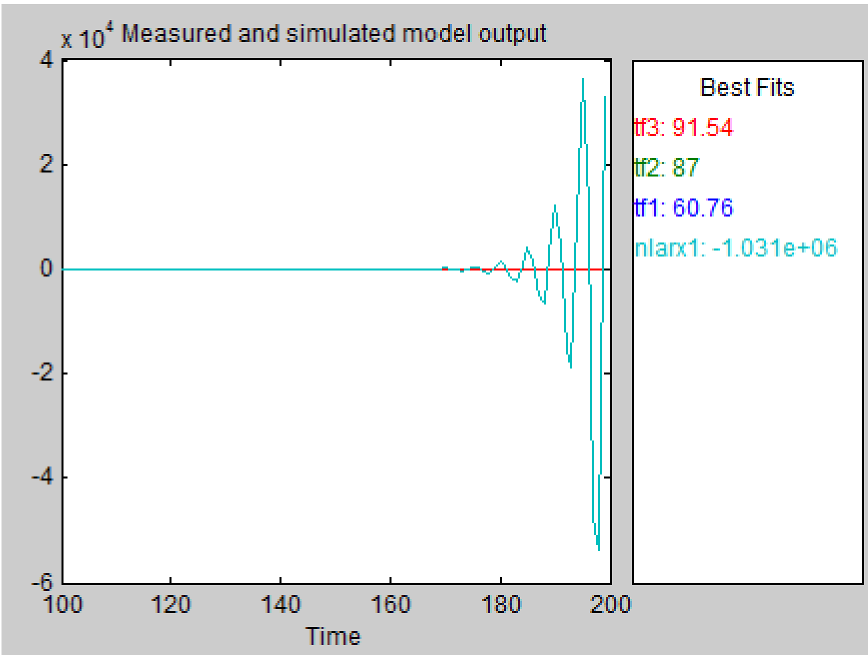


Рис. 9.27. Графики выхода модели с учетом валидации

Из рис. 9.27 видно, что лучшая модель, построенная по идентификационным данным, является наихудшей для валидационных данных, т.е. возникло явление ***переобучения*** (***overfitting*** или чрезмерная подстройка модели под обучающую выборку, в результате которой теряется возможность модели к обобщению). Это и является указателем на качество идентификации. Проанализировать полученные графики, изобразить их в отчете и записать выводы.

Уберем из окна отображения график наихудшей по результатам модели, кликнув на соответствующее окно этой модели (рис. 9.28).

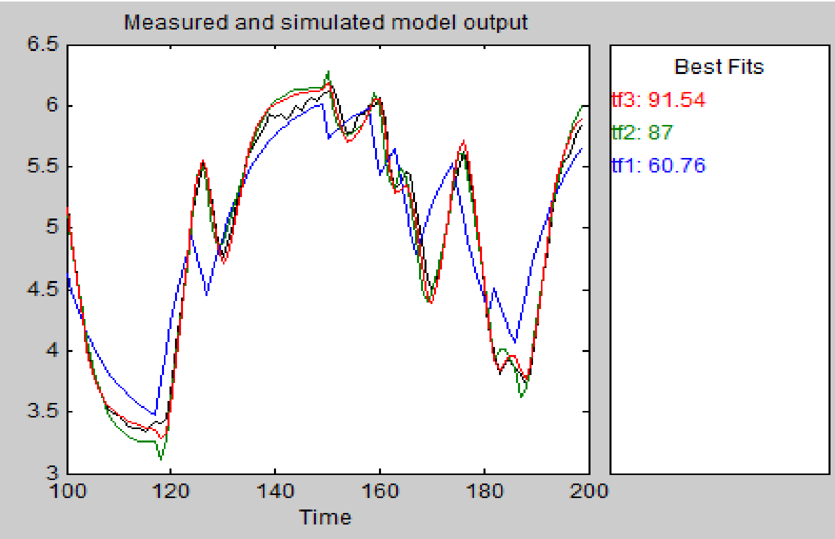


Рис. 9.28. Графики моделей

Также можно посмотреть графики ошибок моделирования, выбрав в окне отображения графиков вкладку **«Options»-«Error plot»** (рис. 9.29).

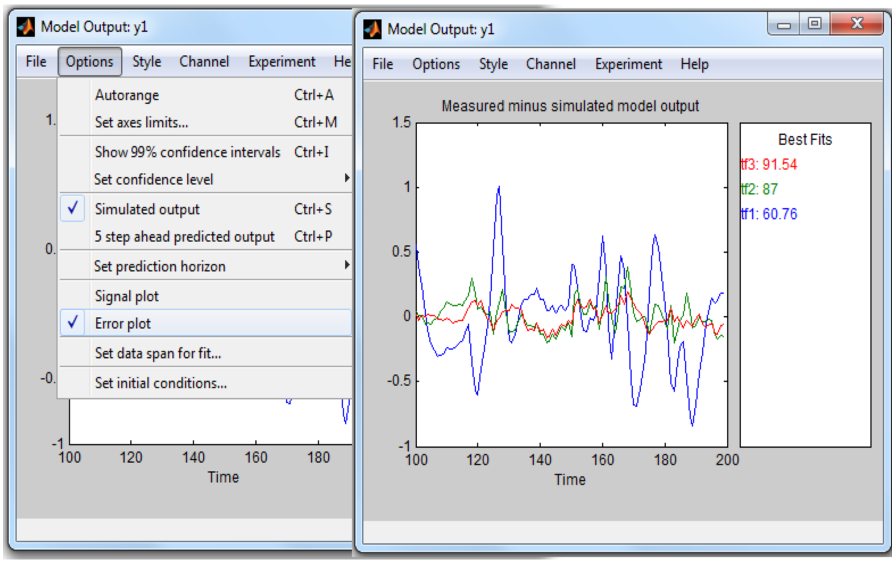


Рис. 9.29. Графики ошибок моделей

8. Проверить качество идентификации, при котором используется анализ невязки. Посмотреть как у оставшихся моделей выглядят ***кросс-корреляционные*** и ***автокорреляционные функции***. Для этого необходимо в окне интерфейса **System Identification Toolbox** поставить галочку в поле **«Model resids»** (рис. 9.30).

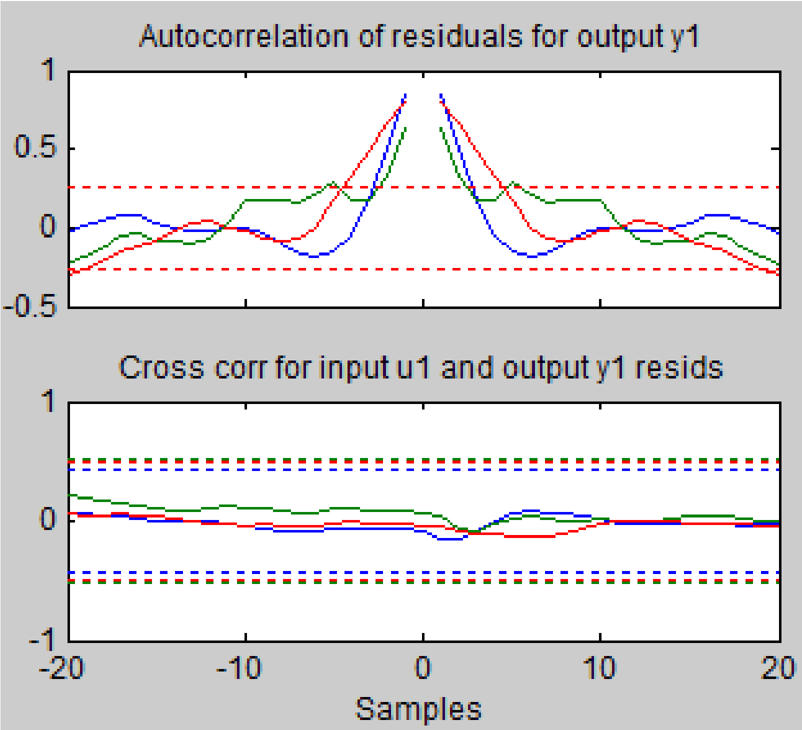


Рис. 9.30. Окно просмотра качества идентификации по кросс-корреляционным и автокорреляционным функциям

Как видно из рис. 9.30 ни одна из моделей не удовлетворяет установленным требованиям по качеству идентификации (графики выходят за границы интервалов). Сделать выводы. Если графики выходят за границы допусков, то качество идентификации неприемлемо и тогда необходимо выбрать другую структуру модели, проверить на наличие обратной связи в данных или учитывать возмущающие факторы (шумы).

9. Построим модель с учетом возмущений (шумов). Возмущение рассматривается как дополнительный вход в систему. Возмущение при этом не измеряется и характеризуется только статистическими свойствами. Измеренные входы содержат возмущения и если присутствует корреляция между возмущениями, то возможно лучше предсказывать будущие значения выходов через наблюдение и учет прошедших значений.

Воспользуемся структурой модели в виде «**State Space Models**» (рис. 9.31).

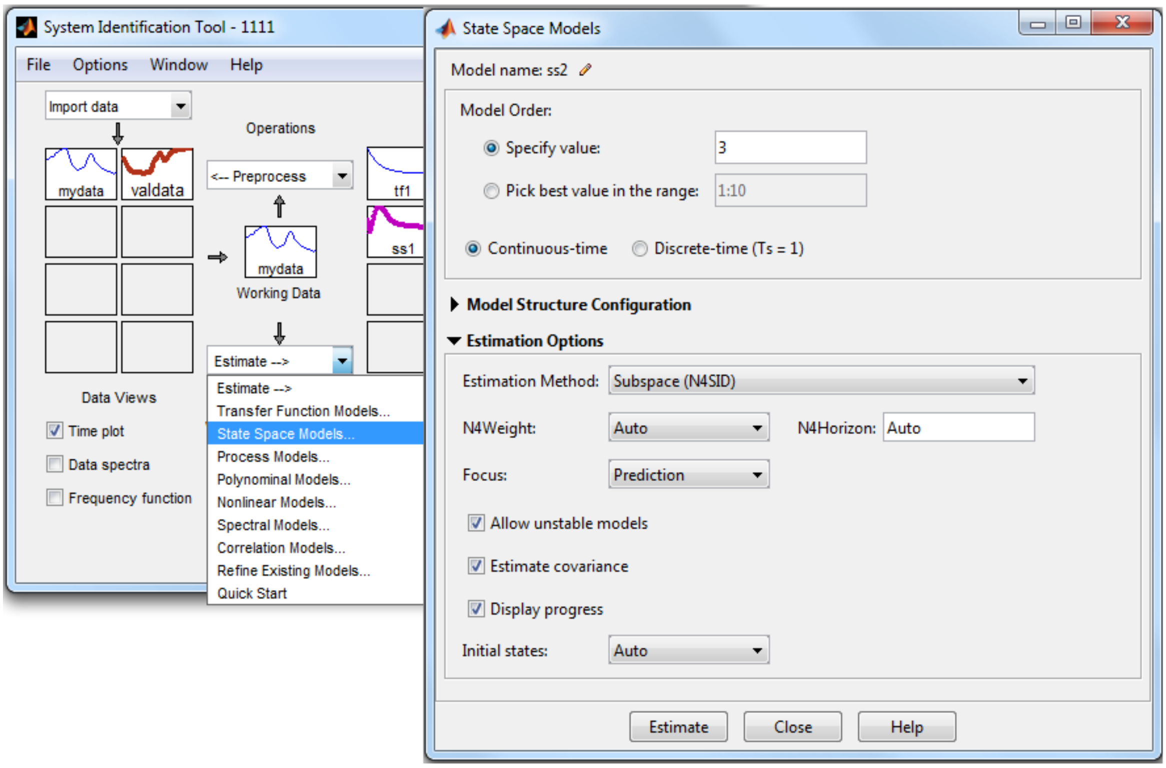


Рис. 9.31. Окно выбора модели

Посмотреть, как у модели выглядят ***кросс-корреляционные*** и ***автокорреляционные функции*** (рис. 9.32). Сделать выводы.

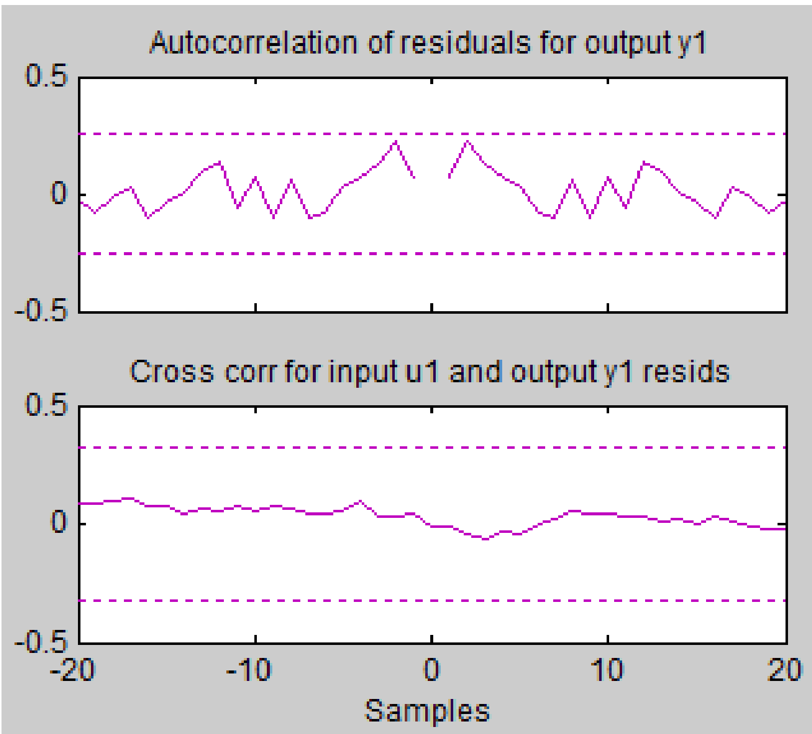


Рис. 9.32. Окно просмотра кросс-корреляционной и автокорреляционной функций

Отчетность по работе

После выполнения работы обучаемый представляет отчет, который должен содержать:

– название работы;
– цель занятия;
– исходные данные;
– графики импортированных данных, выхода моделей;
– передаточные функции и другие формы записи моделей;
– выводы.

Контрольные вопросы

1. Что понимается под идентификацией? Назовите цели и задачи.
2. Каковы основные этапы идентификации?
3. Перечислите основные структуры моделей.
4. Что понимается под валидацией модели?
5. Каким образом оценивается качество модели?
6. К чему приводит переобучение («переподгонка») модели?

Список литературы

1. http://matlab.ru.products/system-identification-toolbox.

2. Моделирование систем управления: учебное пособие / Н.А. Осипов, А.М. Барановский, И.В. Дорожко, А.Г. Тарасов. – СПб.: ВКА имени А.Ф. Можайского, 2015. – 197 с.